

Article

Urban Network and Regions in China: An Analysis of Daily Migration with Complex Networks Model

Wangbao Liu ¹, Quan Hou ^{2,*}, Zhihao Xie ^{3,*} and Xin Mai ¹

¹ Department of Geography, South China Normal University, Guangzhou 510631, China; wbliu@scnu.edu.cn (W.L.); maixin@m.scnu.edu.cn (X.M.)

² Shenzhen Institute of Building Research Co., Ltd., Shenzhen 518049, China

³ Dongguan Urban Planning and Design Institute, Dongguan 523000, China

* Correspondence: houquan@gmail.com (Q.H.); leoxie_scnu@163.com (Z.X.); Tel.: +86-136-9988-8365 (Q.H.); +86-156-2503-8341 (Z.X.)

Received: 27 March 2020; Accepted: 13 April 2020; Published: 16 April 2020



Abstract: This paper analyzed urban network and regions in China using a complex network model. Data of daily migration among 348 prefectural-level cities from the Baidu Map location-based service (LBS) Open Platform were used to calculate urban network metrics and to delineate boundaries of urban regions. Results show that urban network in China displays an obvious hierarchy in terms of attracting and distributing population and controlling regional interaction. Regional integration has become increasingly prominent, as administrative boundaries and natural barriers no longer have strong impacts on urban connections. Overall, 18 urban regions were identified according to urban connectivity, and the degree of urban connection is higher among cities in the same urban region. Due to geographical proximity and close interaction, several provincial capital cities form an urban region with cities from neighboring provinces instead of those from the same province. Identification of urban region boundaries is of significant importance for sustainable development and policymaking on the demarcation of urban economic zones, urban agglomerations, and future adjustment of provincial administrative boundaries in China.

Keywords: Chinese population flow; urban networks; complex networks model; boundary of urban regions; sustainable development; administrative boundaries

1. Introduction

A city is an open system, which regularly exchanges materials, products, services, information, and capital with external environments. As Taylor has pointed out, a city can only exist in the network formed by the interaction with other cities [1]. Castells believed that our society is formed by various flows, and that the accumulation of wealth and power in cities is not achieved through what is possessed in cities, but through various flows [2]. Cities cannot exist independently, thus to a large degree, cities are defined by their exchanges of inputs and outputs with other cities and rural areas. Therefore, with the world becoming increasingly urbanized, the relationship of a city with other cities has come to be the essential element determining the status of the city. In addition, urban systems are increasingly characterized by ‘networks’ in the era of globalization. Consequently, the urban network perspective has been one of the primary approaches to exploring the complex issue of urban systems.

Within the framework of the urban network approach, analysis of relational data between cities is the key focus. Characterized by traffic flows as well as the global expansion and spatial reorganization of transnational corporations among cities, the functional connections among cities have become an important means of interpreting urban linkages and are often used to study the organizational patterns and changes of the world urban system [1]. For example, the Globalization and World Cities (GaWC)

research group has used data of office locations for the global producer services industry to analyze the pattern of social and economic linkages among cities. In addition, traffic flows, such as air traffic, e.g., the number of flights [3,4] and road traffic, have been used to explore spatial patterns of urban linkages. In particular, journeys to work (commutes) data, as a major measurement of the functioning of local and regional labor markets, have been used to identify and delineate megaregions in the United States [5].

In recent years, the rapid development of mobile internet and geoinformation technologies has made it feasible to collect data with geotag, socioeconomic attributes, mobile trajectories, migration processes, and interaction patterns, which can be employed to examine spatial-temporal behaviors in a systematic way [6,7]. The so-called ‘big data’ on human spatial-temporal behaviors have also gradually found their use in studying urban linkages. In particular, social media data with location labels have been the major source of data to examine interactions among cities and regions. For example, keywords and spatial distribution data extracted from Twitter posts are analyzed to measure the level of network activity and the degree of network linkage between cities in the United States [8], while Cheng et al. used Foursquare and Gowalla data to analyze the spatial-temporal footprint among cities [9]. In China, Zhen et al. used Weibo’s friendship data to explore urban network structure and urban systems [10], and Liu et al. scrutinized urban linkages by utilizing check-in data from location-based social networking services (LBSNS) [11].

In contrast, studies of urban systems and urban regions based on population flow data are relatively rare in the literature (for example, [8–11]), although population flow is regarded as a critical indicator of urban productivity and a major measurement of urban linkages. This is partly due to the lack of reliable and timely data. In this regard, empirical studies of urban systems and urban regions using migration data can help to enrich this branch of urban and regional studies. This paper tried to analyze Chinese population flow data within the framework of social network theory and methodology. This paper measured connectivity characteristics among Chinese cities, identified urban regions (communities of cities) to reflect the strength of intercity linkages, and delineated boundaries of urban regions. Due to the availability of data, the analysis was confined to 369 prefecture-level cities in China.

Both theoretical and practical contributions were aimed at by this study. Theoretically, the paper offers an empirical approach to detect urban agglomeration communities, which can demonstrate the boundary-binding effect and its decay in an urban system. Practically, China has established a system of administrative divisions in formulating and implementing regional development strategies and plans. Our proposed methodology and findings can deepen the understanding of urban network development, the spatial pattern of urban linkages, and the structure of urban regions, which can inform adjustment of administrative divisions to achieve more efficient and sustainable policies for urban and regional development.

2. Data and Methodology

2.1. Daily Migration Data

Daily migration data were obtained by aggregating individual movements extracted from Baidu’s location-based services (LBS) open platform. Baidu, as an equivalent to Google in the United States, is the dominant Chinese internet search engine company. The LBS platform is the search provider’s positioning facility that serves more than 5 billion daily positioning requests sent through its various products, such as Baidu Maps and other apps. With its cloud computing capabilities and wide penetration of positioning technologies on mobile devices, Baidu is able to fully and instantly reflect population migration. Daily population movement data, which are recorded and constructed from the positioning information of users’ mobile phones, form the so-called ‘Baidu migration’ data. Data of people’s movement within each city were discarded, while daily movements between cities were extracted and recorded. Daily migration among China’s prefecture-level cities (excluding Sansha

(Sansha is a prefecture-level city officially established in 2012 to administer the island groups and their surrounding waters in the South China Sea. Its transportation connection with other parts of the nation mainly relies on voyages, which is quite different from other cities in China. Hong Kong, Macao, and Taiwan have special regulations for entering into and exiting from mainland China), Hong Kong, Macao, and cities in Taiwan) were constructed for about three months in 2015 (February 7 to May 16, 98 days in total). Average daily migration between each pair of cities was calculated and recorded in a 369×369 unsymmetrical square matrix. During this period, the population flows reached 121 million in total and 1.24 million per day on average.

The unsymmetrical square matrix represents a complex network of population migration among cities. Daily movement of people from a given city to a number of receiving countries forms a set of multidirectional connections. Movement of people between any pair of cities can occur in opposite directions in physical space, implying that the migration network is a weighted directed spatial network. The network constitutes of nodes (i.e., cities) connected by links representing the migrants to and from each city. In this perspective, the topology and behavior of the network can be examined through calculation of selected network metrics.

2.2. Data Applicability

The Baidu migration dataset is the collective result of individual human behaviors. This dataset is constituted of origins, destinations, and flow paths. Although the origins and destinations in a population flow network have specific physical locations, due to data collection limitations and constraints, each is often linked to a territorial area. For instance, origins and destinations may be countries, provinces, cities, or smaller administrative districts. In other words, population flows build networks between territorial areas. In this case, although the population flow network exists in a physical space, the often discussed population flows are just types of projections of a real network in certain spatial scales [12]. Population flows data can be divided into two basic types according to their spatial scales: those occurring at intracity level, such as commuting flow networks [13], and those taking place at intercity or regional level, such as migration flow networks [14,15]. This study applied the Baidu migration data to investigate regional-level migration flow networks. Due to the fact that migration flow networks could be regarded as complex networks, various kinds of network metrics (e.g., degree distribution) are tentatively applied to the structural analysis of migration flow networks [16]. Such metrics play a major role in uncovering the characteristics and nature of migration flow networks, and hence, the spatial pattern of urban connectivity.

2.3. Measurement

With recent development of graph theory and social network analytical tools, exploration of network organization has received increasing attention in academia. In particular, studies on spatial organization and patterns of urban interaction using the complex network model have become a major theme in transport planning and transportation geography research [17]. The basic approach within the complex network framework is to derive typical complex network metrics to investigate characteristics and the nature of the network. Among these metrics, centrality is the most widely used one, which consists of degree centrality (DC), betweenness centrality (BC), and closeness centrality (CC).

Communities in a network are the dense groups of the vertices, which are tightly coupled to each other inside the group and loosely coupled to the rest of the vertices in the network. Community detection, also known as community mining, denotes the identification of the community structure within the network through topological relationships and properties [14]. In our case, cities are well connected within communities and less connected among communities. Community structure is the most widely studied structural feature of complex networks. Consequently, community detection plays a pivotal role in understanding the functionality of complex networks.

2.3.1. Degree Centrality and Betweenness Centrality

Degree centrality is defined as the number of link incidents upon a node (i.e., the number of ties that a node has). DC is primarily used to measure the significance and influence of a node in the network. A higher value of DC indicates a greater distributing capacity of a certain node in the network. The calculation formulas for DC and BC are shown below:

$$DC(N_i) = \sum_{i=1}^g x_{ij}(i \neq j) \quad (1)$$

where $\sum_{i=1}^g x_{ij}$ is the total number of direct connections among the node i and other $g-1$ nodes, and denotes the total inflow and outflow of urban population flows. For example, each node of a line has a degree centrality of 2, as either node has one inflow connection and one outflow connection. In our case, DC was applied to measure the significance of a certain city in the urban network in distributing migrants and controlling interactions among cities.

As an indicator of a node's centrality in a network, betweenness centrality is defined as the number of shortest paths from all vertices to all others that pass through the node under concern. BC aims to measure a node's capacities of controlling information across the network. If a node falls on the shortest path between others, it means a higher degree of BC and a greater ability to control communication [18]. The calculation of BC is shown in the following formula:

$$BC(V) = \sum_{s,t \in V} \frac{\partial(i, j|v)}{\partial(i, j)} \quad (2)$$

where $\partial(i, j)$ is the number of shortest paths between nodes i and j ; $\partial(i, j|v)$ is the number of shortest paths passing through node v between nodes i and j . Also, taking a line as an example, of which either node has a betweenness centrality of 1, as all the paths passing through the node are also the shortest ones.

2.3.2. Community Detection

Recently, many algorithms have been proposed for community detection [19–22]. Among them, the fast unfolding algorithm has gained wide acceptance and has been widely used in the field of community detection [20]. This paper employed the fast unfolding algorithm to process community detection and this algorithm makes use of the function of modularity to characterize the network's distinctness among communities and reveal the characteristics of the local agglomeration within the network. A higher value of modularity indicates better division of communities [23]. In a weighted network, modularity (Q) is computed as follows:

$$Q = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \Delta(c_i, c_j) \right) \quad (3)$$

where m is the total number of edges in the network, A_{ij} is the weight of the edge linking node i and node j ; k_i and k_j are the sum of weights of all edges linking with node i and node j , respectively. C_i and C_j represent the communities to which node i and node j belong, respectively. Theoretically, the value of modularity (Q) falls between 0 and 1. The closer the value is to 1, the more obvious is the community structure. Empirically, the value of Q usually falls into the range from 0.3 and 0.7 in the complex networks representing a particular phenomenon in the real world, such as population flows.

3. Results

3.1. Degree Centrality

The spatial distribution of DC is shown in Figure 1a. The Jenks natural breaks classification method is used to describe the distribution according to the amount of daily average population

distribution for each city. The cities under investigation are classified into five categories with daily average population distribution ranging from 0–5952, 5952–16,441, 16,441–41,469, 41,469–74,881, and >74,881. The top 20 cities are as follows: Beijing, Shanghai, Suzhou, Tianjin, Shenzhen, Guangzhou, Langfang, Dongguan, Hangzhou, Chongqing, Nanjing, Baoding, Urumqi, Xi'an, Wuhan, Foshan, Xuzhou, Ganzhou, Ningbo, and Chengdu. It is obvious that the highest-level cities, namely, with the highest values of DC, are concentrated mainly in the coastal areas, especially China's four major urban agglomerations (i.e., Beijing–Tianjin–Hebei, the Yangtze River Delta, the Pearl River Delta, and Chengdu–Chongqing). In addition, an obvious belt-shape distribution of urban communities was identified along the Beijing–Shanghai high-speed railway. However, the high-level cities in the midwest of China are mainly provincial capital cities. The differences in the distribution hierarchy of provinces are also significant. There are no third- or higher-level distribution cities in Tibet, Ningxia, Hainan, Gansu, Fujian, Hunan, Yunnan, Guangxi, Inner Mongolia, Liaoning, Heilongjiang, and Jilin. Urumqi represents the highest level in northwest China, while Chongqing and Chengdu are the largest population concentration and distribution centers in southwest China. Shenzhen, Guangzhou, and Dongguan have the highest levels in Guangdong province, whereas the eastern, western, and northern parts of the Guangdong province do not have high levels of concentration and distribution centers. Fujian, Heilongjiang, Liaoning, and Jilin in northeast China, as well as Yunnan, Guizhou, and Guangxi in southwest China, do not have cities at the third level or above. Some provinces show exceptions, in which the provincial capital cities are not the highest-level cities. For example, the highest-level city of Jiangxi province is Ganzhou rather than its capital city, Nanchang. In addition, the level of Suzhou in Anhui province is slightly higher than the capital city Hefei, and the levels of Langfang and Baoding in Hebei province are higher than the capital city Shijiazhuang. Similarly, Shenzhen also has a higher level than the capital city Guangzhou in Guangdong province.

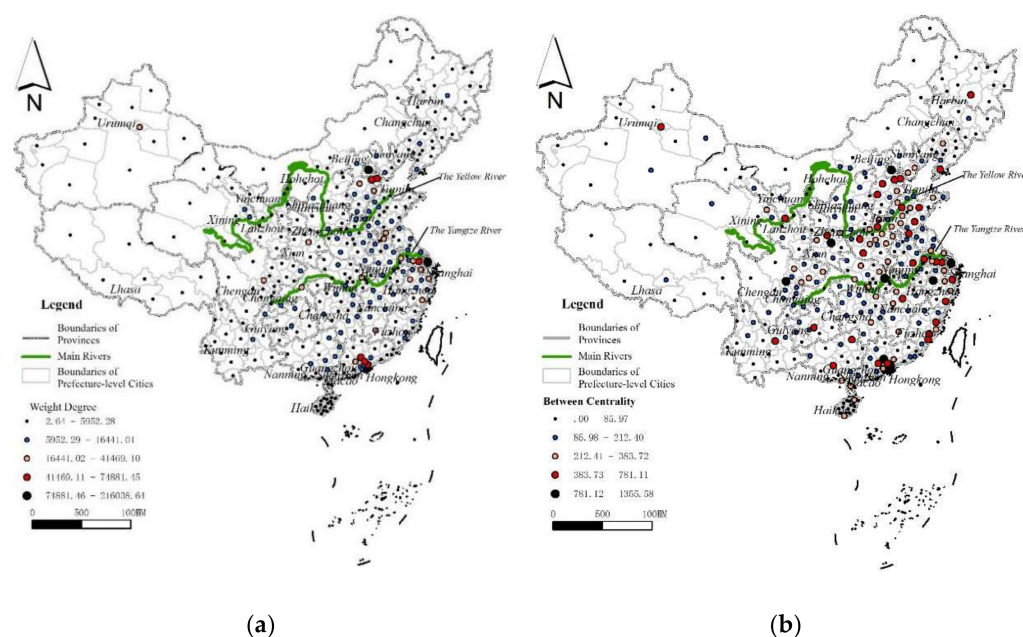


Figure 1. Spatial patterns of different levels of urban population flows: (a) degree Centrality; (b) betweenness Centrality.

3.2. Betweenness Centrality

Betweenness centrality (BC) measures the ability to control communication among cities in the global structure of a network, which indicates the probability of a node serving as a 'bridge'. Hence, nodes with high BC play the role of hubs in cities' interaction networks. Figure 1b illustrates spatial variation of BC, which can be classified into five levels using the Jenks natural breaks method;

namely, 0–85, 85–212, 212–383, 383–781, and >781. The top 20 centers controlling communication among cities are as follows: Beijing, Chengdu, Guangzhou, Shenzhen, Shanghai, Chongqing, Xi'an, Wuhan, Hangzhou, Suzhou, Tianjin, Nanjing, Shijiazhuang, Dongguan, Jinan, Nanning, Kunming, Fuzhou, Xiamen, and Foshan. Chongqing and Chengdu, Beijing and Tianjin, Shanghai and Hangzhou, Zhengzhou and Wuhan, and Guangzhou and Shenzhen are the communication control centers in western, northern, eastern, and southern China, respectively. However, there is a lack of strong control centers in northeast, southwest, and northwest China. In northwest China, the top three cities with the highest BC index are the provincial capital cities, namely Urumqi, Lanzhou, and Xining, whereas the top three cities with the highest DC index are Urumqi, Xining, and Haidong.

3.3. Community Detection

Community detection provides a new way to understand the functional relationships shaping the spatial organization of a region and the territorial structures derived from such shaping. The open-source network analysis and interactive visualization tool Gephi v0.9.2 was used to implement community detection. Using a 3D rendering engine to display graphs in real-time, Gephi supports almost all types of graphical networks, including complex networks, and speeds up the exploration of relationships within data. Resolution in the calculation was repeatedly adjusted until the modularity reached its maximum of 0.532, which falls into the range between 0.3 and 0.7, and then the community detection is considered the final result [21]. The distribution of the communities identified is shown in both Figure 2 and Table 1, revealing a significant pattern of spatial clustering. Eighteen clusters of cities, or communities, were identified. Cities within communities are well connected, while cities between communities are less connected. There were no geographical elements in the calculation of the fast unfolding algorithm to detect communities, yet the results reveal strong influences of geographical proximity on the structure of the clusters.



Figure 2. Distribution of communities detected.

Table 1. The scope of communities detected.

Communities	Distribution	Spatial Organization Models	Central Cities
Community1	Beijing, Tianjin, Hebei, Eastern Shanxi (Datong, Shuozhou, Jinzhou, Changzhi, Yangquan), Northern Shandong (Cangzhou, Dezhou, Jinan, Tai'an, Laiwu, Zibo, Weifang, Yantai, and Weihai), Liaoning (Anshan, Liaoyang, Jinzhou, Panjin)	Polycentric	Beijing, Tianjin, Langfang, Baoding, Luzhou
Community2	Ma'anshan and Chuzhou in Anhui, Nanjing in Jiangsu	Monocentric	Nanjing
Community3	Tibet, Ganzi in Sichuan, Diqing in Yunnan	Incompletely developed	/
Community4	Guangdong, Hainan, Guangxi (Guilin, Liuzhou, Hezhou, Laibin, Zhangzhou, Guigang, Nanning, Qinzhou, Fangchenggang, Beihai, Yulin), Hunan (Zhangjiajie, Huaihua, Yiyang, Shaoyang, Loudi, Xiangtan, Hengyang, Yongzhou, Quzhou), Jiangxi Zhangzhou, Fujian (Longyan, Zhangzhou), Guizhou (Southeast Guizhou)	Polycentric	Shenzhen, Guangzhou, Dongguan, Foshan, Luzhou
Community5	Shanghai, Jiangsu (except Nanjing, Lianyungang, Xuzhou), Anhui (except Suzhou, Huaibei, Ma'anshan, and Luzhou), Henan (Zhoukou, Xinyang), Zhejiang (Hangzhou, Jiaxing, Ningbo, Shaoxing)	Polycentric	Shanghai, Suzhou, Wuxi, Jiaxing, Hangzhou
Community6	Pingdingshan, Luohe, Zhumadian in Henan	Incompletely developed	/
Community7	Southern Gansu (Jinchang, Wuwei, Baiyin, Dingxi, Pingliang, Longnan, Tianshui, Qingyang), Ningxia, Inner Mongolia (Wuhai, Ordos), Shaanxi (Yulin, Yan'an, Xi'an, Baoji, Hanzhong, Xianyang), Shanxi (Zhangzhou, Luliang, Linyi, Jincheng), Henan (Luoyang, Jiaozuo, Xinxiang, Jiyuan)	Monocentric	Xi'an
Community8	Qinghai, Gansu (Lanzhou, Linxia), Sichuan (Aba)	Incompletely developed	/
Community9	Henan (Nanyang), Shanxi (Tongchuan, Shangluo, Ankang), Hubei (Shiyan, Xiangyang)	Incompletely developed	/
Community10	Henan (Zhengzhou, Yuncheng, Sanmenxia), Shanxi (Weinan)	Incompletely developed	/
Community11	Chongqing, Sichuan (Chengdu, Ya'an, Meishan, Leshan, Zigong, Luzhou, Neijiang, Deyang, Mianyang, Suining, Nanchong, Guang'an, Dazhou, Bazhong, Panzhihua), Guizhou (Tongren, Zunyi), Hunan (Zhangjiajie), Hubei (Enshi, Yichang, Shenlongjia), Yunnan (Lijiang, Nujiang, Dali, Baoshan, Dehong, Chuxiong, Kunming, Yuxi, Honghe, Xishuangbanna, Pu'er)	Dual-nuclei	Chongqing, Chengdu
Community12	Hubei (Jingmen, Suizhou, Xiaogan, Wuhan, Ezhou, Xianning, Tianmen, Qianjiang, Xiantao, Jingzhou), Hunan (Changsha, Yueyang, Zhuzhou), Jiangxi (Pingxiang, Ji'an, Xinyu, Yichun, Nanchang)	Monocentric	Wuhan
Community13	Xinjiang, Gansu (Jiuquan, Jiayuguan, Zhangye)	Monocentric	Urumqi
Community14	Shandong (Liaocheng), Henan (Puyang)	Incompletely developed	/
Community15	Shandong (Linyi, Jining, Zaozhuang, Heze, Rizhao), Jiangsu (Lianyungang, Xuzhou), Henan (Kaifeng, Shangqiu), Anhui (Suzhou, Huaibei)	Dual-nuclei	Suzhou, Xuzhou
Community16	Yunnan (Qujing, Wenshan, Zhaotong), Sichuan (Liangshan, Yibin), Guizhou (Guiyang, Bijie, Liupanshui, Weinan Southwest, Anshun, Weinan), Guangxi (Hechi, Bai, Chongzuo)	Incompletely developed	/
Community17	Fujian (Ningde, Nanping, Fuzhou, Putian, Quanzhou, Xiamen, Sanming), Zhejiang (Lishui, Wenzhou, Jinhua, Zhangzhou, Lishui, Taizhou), Jiangxi (Jiujiang, Shangrao, Jingdezhen, Fuzhou), Hubei (Huanggang, Huangshi)	Monocentric	Wenzhou
Community18	Heilongjiang, Jilin, Liaoning (Tieling, Fushun, Benxi, Dandong, Shenyang, Dalian, Yingkou, Fuxin, Huludao, Chaoyang), Hebei (Qinhuangdao), Inner Mongolia (Hulunbeier, Xing'anmeng, Tongliao, Chifeng)	Incompletely developed	/

The results of community detection of Chinese cities shed new light on the old yet important problem of regionalization. The delineation of urban regions is an important aspect of understanding spatial structure of urban system. Previous studies hold different views on the division of urban regions in China. For example, Gu claimed that China has 33 urban economic regions (two tier) [24], whereas Zhou and Zhang contended 11 urban economic regions (two tier) [25]. Chen et al. divided China into 19 urban economic regions according to highway passenger flows [26]. The prominent feature of detected communities is that the cities are well connected within each community, while less connected among communities. Population mobility can be taken as a measurement of economic connection between cities, and the boundaries of the detected urban communities have a high similarity to those of the urban economic zones. The number of detected urban communities matches the results of Chen et al. [26], while the spatial structure of urban economic regions is different. The main features of communities identified in this study are described as follows.

Firstly, urban linkages have obviously moved beyond administrative boundaries and natural geographical barriers. City regions show an increasing interprovincial integration. Cities in each province of Tibet, Xinjiang, Qinghai, Ningxia, Jilin, Heilongjiang, Jiangsu, Anhui, Guangdong, and Hainan belong to the province to which they administratively belongs, indicating a certain extent of an administrative district economy. However, almost all the communities' boundaries have moved beyond the provincial administrative boundaries, denoting that the restrictive effect on the regional economy by the administrative boundaries has been weakened. Cities in several provinces are separated into different communities. For example, Gansu is divided into three parts: the north part joins Xinjiang to form Community 13; the south part combines with Ningxia, central Inner Mongolia, Shaanxi, and parts of Shanxi to be Community 7; and the west part merges with Qinghai and Aba of Sichuan shape Community 9. In addition, due to geographical proximity and close urban linkages, the results reveal a phenomenon that some provincial capital cities may link up with their neighboring cities in adjacent provinces instead of those from the same province to establish a community. For example, the capital city of Nanjing in Jiangsu province joins hands with Ma'anshan and Chuzhou in Anhui province to form Community 2, while the other cities in Jiangsu province are under the realm of influence from Shanghai. Shanghai leads Community 5, consisting of the majority of cities from Anhui and Zhejiang provinces and part of Henan province. This phenomenon supports the argument to alleviate restrictive effects from the administrative boundaries on regional economic integration. Rapid development of transportation infrastructure has weakened geographical barriers, such as huge mountains and rivers, to urban linkages. The Wuyi Mountains and Taihang Mountains are the boundaries between Fujian–Jiangxi provinces and Shanxi–Hebei provinces, respectively. The Nanling Mountains meanders along the boundaries linking Guangdong, Guangxi, Hunan, and Jiangxi provinces. However, these mountains do not stand out in the community division map as dividing lines among city regions.

Secondly, divergences exist between the regional boundaries identified in this paper (e.g., the Yangtze River Delta (Community 5), Beijing–Tianjin–Hebei (Community 1), Pearl River Delta (Community 4), and Chengdu–Chongqing (Community 12)), with the boundaries of the urban economic zone or urban agglomeration defined by the government or academia. The scope of Beijing–Tianjin–Hebei (Community 1) consists of Beijing, Tianjin, Hebei, central and western Inner Mongolia, northern Shandong, and eastern Shanxi, which exceeds the scope of what is defined in the 'Beijing–Tianjin–Hebei Urban Agglomeration Plan' that states the Beijing–Tianjin–Hebei Urban Agglomeration only consists of Beijing, Tianjin, and 11 cities in Hebei province (Shijiazhuang, Zhangjiakou, Qinhuangdao, Tangshan, Baoding, Langfang, Xingtai, Handan, Hengshui, Cangzhou, and Chengde). The Yangtze River Delta agglomeration (Community 4) encompasses most parts of Jiangsu province (except Nanjing, Xuzhou, and Lianyungang) and Anhui province (except Suqian and Bozhou), Eastern Henan (including Xinyang and Zhoukou), and Northern Zhejiang province (Hangzhou, Shaoxing, Jiaxing, and Ningbo). This classification is also quite different from the scope of the Yangtze River Delta urban agglomeration as defined in the 'Yangtze River Delta Urban Agglomeration Plan' approved by the State Council in 2016. Furthermore, the radiation effects of the

cities in the north of Shanghai are obviously stronger than those in the south. The Pearl River Delta urban agglomeration (Community 4) covers Guangdong, Guangxi, Hunan, Jiangxi, most parts of Fujian, Southeast Guizhou, and Eastern Yunnan, and its scope is significantly larger than the existing definition of the Pearl River Delta agglomeration, which refers to nine cities within the Guangdong Province (Guangzhou, Foshan, Zhaoqing, Shenzhen, Huizhou, Dongguan, Zhuhai, Zhongshan, and Jiangmen). Nevertheless, Community 4 is smaller than the so-called Pan Pearl River Delta region (including Guangdong, Guangxi, Fujian, Jiangxi, Guangxi, Hainan, Hunan, Sichuan, Yunnan, Guizhou, Hong Kong, and Macao). The Chengdu–Chongqing agglomeration (Community 11) includes Chongqing, eastern Sichuan, northwest Guizhou, northern Yunnan, etc., which is larger than the scope of Chengdu–Chongqing urban agglomeration (including Chongqing and 15 cities in Sichuan, which are Chengdu, Deyang, Mianyang, Meishan, Ziyang, Suining, Leshan, Ya'an, Zigong, Luzhou, Neijiang, Nanchong, Yibin, Dazhou, and Guangan) as defined in the 'Regional Planning of Chengdu–Chongqing Economic Zone' endorsed by the State Council in 2011.

Lastly, a typology of urban regions in China was formulated based on the value of DC and its spatial distribution. Communities of cities can be classified into four types according to the spatial patterns: monocentric mode, dual-nuclei mode, polycentric mode, and incompletely developed mode. It is noted that cities with DC over the third level are defined as the central city in a community. In general, Communities 2, 7, 12, 13, and 17 fall to the monocentric category, while Communities 11 and 15 are identified as dual-nuclei mode; Communities 1, 4, and 5 are classified as polycentric mode, and Communities 3, 6, 8, 9, 10, 14, 16, and 18 belong to the incompletely developed mode. Beijing, Tianjin, Langfang, Baoding, and Cangzhou are central cities of Community 1 and are located mainly around Beijing and Tianjin. Shenzhen, Guangzhou, Dongguan, Foshan, and Ganzhou are the central cities of Community 4; Dongguan and Foshan have shown the strongest connections with Shenzhen and Guangzhou, while Ganzhou has the strongest connections with the Pearl River Delta region among cities outside the Guangdong province. Shanghai, Suzhou, Wuxi, Jiaxing, and Hangzhou are the central cities of Community 5, and these cities are located around the core city of Shanghai. Chongqing and Chengdu are the central cities of Community 11, while Chongqing has a higher level of distributing hierarchy than Chengdu. Suzhou and Xuzhou are the central cities of Community 15. Nanjing in Jiangsu, together with Ma'anshan and Chuzhou of Anhui, form the monocentric mode Community 2, of which Nanjing is the central city. Although Community 7 spreads across the provinces of Gansu, Ningxia, Inner Mongolia, Shaanxi, and Shanxi, only Xi'an, the central city, records a concentration distribution level above the third level. Community 13 involves the provinces of Xinjiang and Gansu, with Urumqi being the central city. The rest of the communities are generally less developed without obvious polycentricity.

4. Discussion

This study has revealed a prominent trend of interprovincial integration of regional development in China. The majority of urban regions identified in this study have moved beyond the provincial-level administrative boundaries and traditionally important geographic barriers, which is contradictory to relevant studies arguing that city communities often match well with administrative boundaries [27–29]. Administrative boundaries have been less important in shaping the spatial pattern of urban connectivity. The results of this study also confirmed the existence of administrative economy in China, yet it is limited to the relatively less developed regions. It is evident that, along with social and economic development, the boundary effects of administrative division and geographic barriers on urban and regional integration are weakening. This is argued to be the fundamental reason why a number of provinces have been separated into two or three urban regions. In particular, Gansu province is divided into northern, southern, and western parts, and Sichuan province is decomposed to be western and eastern parts. This finding has important policy implications as it can shed light on possible adjustments of provincial administrative boundaries in the future. Provincial capitals, such as Nanjing and Lanzhou, are closely connected to cities from adjacent provinces and constitute a de

facto urban region, which implies that there may be a need to adjust administrative boundaries to better cater for urban interactions. This argument even holds for the boundaries overlapping with major mountain ranges, such as the Taihang, Nanling, and Wuyi Mountains. Traditionally these boundaries have been a critical factor in delineating provincial territories. However, according to the city communities identified in this study, they may not constitute the boundaries of urban regions any more. This finding is different from previous arguments that they usually form a relatively closed or semi-closed regional economy [26]. Part of the reason for this difference is due to the difference of data. Highway flow data were used in Chen et al., on which the natural barriers could impose a significant impact [26]. While this study examined population flow data, which were less confined by natural geographic barriers. The measurement of socioeconomic intercity interactions, such as population flow data employed in this study, can provide a more systematic and comprehensive picture of the organization of urban network. Hence, we argue that the demarcation of urban regions—as urban economic zones or urban agglomerations—should pay more attention to socioeconomic linkages rather than geographic boundaries. This should be particularly emphasized in making regional policies and planning proposals.

Thus far, the boundaries and spatial patterns of urban agglomeration in China are still inconclusive and, to a large extent, dependent on the measurement used by researchers as well as major concerns of regional development plans. A systematic way of capturing the dynamics of diverse urban linkages is of fundamental importance for delineating urban agglomerations in China. This paper argues that urban linkage is critical to examine the mechanism of urban system and urban agglomeration, and thus is an essential aspect in formulating regional sustainable development strategies. However, location still matters, and the radiation capacity of cities in urban regions is an important factor in explaining the function, role, and relative position of a city in the urban network hierarchy. This study reveals that the composition of the four prominent urban regions is different from the well-documented urban agglomerations in China. The boundaries of Community 1 and Community 4 are distinct from the Beijing–Tianjin–Hebei and Pearl River Delta urban agglomerations as defined by the government or academia [30]. The application of big data, such as cellphone signaling data, has already been used to study urban agglomeration from the perspective of urban linkages [31]. Using the big data of population flow, this paper applied the method of complex network analysis to examine spatial pattern of urban regions in China, demonstrating this method is useful to identify the core city and unfold the polycentricity of regional configuration. The results show that complex network analysis may shed light on urban agglomeration studies and urban system planning.

5. Conclusions

By using Baidu migration data, this study adopted social network analysis to investigate the pattern and characteristics of daily population flow in urban China, leading to our demarcation of urban regions in China. A prominent hierarchy exists among cities in terms of their roles in gathering and distributing population and regional resources. Cities on the top of the hierarchy are mainly located in coastal regions and key transportation corridors. However, the capacity of a city as an intermediary control is largely determined by its location and the condition of its transport infrastructure. Apart from regional centers, such as Beijing, Shanghai, and Guangzhou, cities with advantaged logistic capabilities such as Jinan, Chengdu, Xi'an, Wuhan, Zhengzhou, Fuzhou, and Xiamen also have a high betweenness centrality (BC). Chongqing and Chengdu, Beijing and Tianjin, Shanghai and Hangzhou, and Zhengzhou and Wuhan, as well as Guangzhou and Shenzhen are the 'bridges' of the west, north, east, south, and middle of China, respectively. Natural and administrative boundaries are no longer the major barriers of urban linkages, implying a prominent trend of regional integration, as urban regions have obviously crossed over the administrative boundaries. Huge mountain ranges, such as the Taihang, Nanling, and Wuyi Mountains, are important interprovincial boundaries, but they no longer constitute the boundaries of urban regions as identified in this study. It is interesting that this study found the provincial capital city may form urban regions with cities from its neighboring

provinces due to geographical proximity and close urban linkages. For example, cities in Jiangsu province are divided into two different urban regions: the provincial capital city Nanjing leads the group of Ma'anshan and Chuzhou from Anhui province, while the rest of cities in this province fall to the group chaired by Shanghai. The spatial organization of urban regions in China can be classified into four types: monocentric, dual-nuclei, polycentric, and less developed mode. The identified regional boundaries of the Yangtze River Delta (Urban region 5), Beijing–Tianjin–Hebei (Urban region 1), the Pearl River Delta (Urban region 4), and Chengdu–Chongqing (Urban region 12) are quite different from the boundaries of the urban economic zones or urban agglomerations proposed by the government or academia. This finding can shed light on making policies on adjusting urban economic zones, regional sustainable development and urban agglomerations in China.

Two major limitations should be noted for the Baidu migration data used in this study. Due to the restrictions of data collection and preprocessing, travel behaviors by those who do not use Baidu Maps are not recorded in the dataset. In addition, for the recorded travels, a trip with a transfer station along the road was identified as two separate trips. Despite its inherited limitations, the Baidu migration dataset proves to be able to reflect the daily population flow dynamics in urban China, and reveal urban interactions and linkages in a comprehensive and systematic way. With regard to revealing dynamic relationships between cities, the dataset shows advantages over conventional data sources of census or transportation statistics. By exploiting the daily migration flows, this research sought to investigate the structure of urban systems in China. Yet the analysis was confined to the prefecture-level cities, due to the availability of such data. Future research may further investigate China's urban network with data at the finer scales, such as county level or even town level, which can enrich the details of the investigation and deepen our understanding.

Author Contributions: Conceptualization, W.L.; methodology, W.L., Q.H. and Z.X.; software, Z.X.; validation, W.L., Q.H., and X.M.; formal analysis, W.L. and Z.X.; data curation, Q.H.; writing—original draft preparation, W.L. and Z.X.; writing—review and editing, W.L., Z.X., X.M., and Q.H.; funding acquisition, W.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China, grant number 41001088 and the Fund of Social Sciences Research, Ministry of Education of China, grant number 17YJA840011.

Acknowledgments: The authors would like to thank Miss Lyu from The University of Hong Kong and Dr. Sigler from The University of Queensland for their valuable comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Taylor, P.; Hoyler, M. The spatial order of European cities under conditions of contemporary globalization. *Tijdschrift voor Economische en Sociale Geografie* **2000**, *91*, 176–189. [\[CrossRef\]](#)
2. Castells, M. Local and global: Cities in the network society. *Tijdschrift Voor Economische En Sociale Geografie* **2002**, *93*, 548–558. [\[CrossRef\]](#)
3. Matsumoto, H. International urban systems and air passenger and cargo flows: Some calculations. *J. Air Transp. Manag.* **2004**, *10*, 239–247. [\[CrossRef\]](#)
4. Zhou, Y.X.; Hu, Z.Y. Looking into the network structure of Chinese urban system from the perspective of air transportation. *Geogr. Res.* **2002**, *21*, 276–286. [\[CrossRef\]](#)
5. Nelson, D.G.; Rae, A. An economic geography of the United States: From commutes to megaregions. *PLoS ONE* **2016**, *11*, e0166083. [\[CrossRef\]](#)
6. Goodchild, M.F. Citizens as sensors: The world of volunteered geography. *GeoJournal* **2007**, *69*, 211–221. [\[CrossRef\]](#)
7. Ke, W.; Yu, Z.; Chen, W.; Wang, H.; Zhao, Z. Architecture and key issues for human space-time behavior data observation. *Geogr. Res.* **2015**, *34*, 373–383. [\[CrossRef\]](#)
8. Naaman, M.; Zhang, A.X.; Brody, S.; Lotan, G. On the study of diurnal urban routines on Twitter. In Proceedings of the 6th International AAAI Conference on Weblogs and Social Media, Dublin, Ireland, 4–7 June 2012.

9. Cheng, Z.; Caverlee, J.; Lee, K.; Sui, D.Z. Exploring millions of footprints in location sharing services. In Proceedings of the Fifth International Conference on Weblogs and Social Media, Barcelona, Spain, 17–21 July 2011.
10. Zhen, F.; Wang, B.; Chen, Y. China's city network characteristics based on social network space: An empirical analysis of sina micro-blog. *Acta Geogr. Sin.* **2012**, *67*, 1031–1043. (In Chinese) [\[CrossRef\]](#)
11. Liu, Y.; Sui, Z.; Kang, C.; Gao, Y. Uncovering patterns of inter-urban trip and spatial interaction from social media check-in data. *PLoS ONE* **2014**, *9*, e86026. [\[CrossRef\]](#)
12. Wei, Y.; Song, W.; Xiu, C.; Zhai, Z. The rich-club phenomenon of china's population flow network during the country's spring festival. *Appl. Geogr.* **2018**, *96*, 77–85. [\[CrossRef\]](#)
13. Patuelli, R.; Reggiani, A.; Gorman, S.P.; Nijkamp, P.T.; Bade, F.J. Network analysis of commuting flows: A comparative static approach to German data. *Netw. Spat. Econ.* **2007**, *7*, 315–331. [\[CrossRef\]](#)
14. Consterdine, E.; Everton, A. European migration network: Immigration of international students to the EU: Empirical evidence and current policy practice. *Science* **2012**, *290*, 1768–1771. [\[CrossRef\]](#)
15. Chun, Y.; Griffith, D.A. Modeling network autocorrelation in space–time migration flow data: An eigenvector spatial filtering approach. *Ann. Assoc. Am. Geogr.* **2011**, *101*, 523–536. [\[CrossRef\]](#)
16. Davis, K.F.; D'Odorico, P.; Laio, F.; Ridolfi, L. Global spatio-temporal patterns in human migration: A complex network perspective. *PLoS ONE* **2013**, *8*, e53723. [\[CrossRef\]](#)
17. Guimerà, R.; Mossa, S.; Turtschi, A.; Amaral, L. The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 7794–7799. [\[CrossRef\]](#)
18. Freeman, L.C. A set of measures of centrality based on betweenness. *Sociometry* **1977**, *40*, 35–41. [\[CrossRef\]](#)
19. Girvan, M.; Newman, M.E.J. Community structure in social and biological networks. *Proc. Natl Acad. Sci. USA* **2002**, *99*, 7821–7826. [\[CrossRef\]](#)
20. Clauset, A.; Newman, M.E.J.; Moore, C. Finding community structure in very large networks. *Phys. Rev. E* **2004**, *70*, e066111. [\[CrossRef\]](#)
21. Blondel, V.D.; Guillaume, J.L.; Lambiotte, R.; Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp. (JSTAT)*. **2008**, *25*, 155–168. [\[CrossRef\]](#)
22. Rosvall, M.; Bergstrom, C.T. Maps of random walks on complex networks reveal community structure. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 1118–1123. [\[CrossRef\]](#)
23. Lancichinetti, A.; Fortunato, S. Community detection algorithms: A comparative analysis. *Phys. Rev. E* **2009**, *80*, e056117. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Gu, C. A preliminary study on the division of urban economic regions in China. *Acta Geogr. Sin.* **1991**, *46*, 129–141. (In Chinese) [\[CrossRef\]](#)
25. Zhou, Y.X.; Zhang, L. China's urban economic region in the open context. *Acta Geogr. Sin.* **2003**, *58*, 271–284. (In Chinese) [\[CrossRef\]](#)
26. Chen, W.; Liu, W.; Ke, W. The spatial structures and organization patterns of China's city networks based on the highway passenger flows. *Acta Geogr. Sin.* **2017**, *72*, 224–241. [\[CrossRef\]](#)
27. Blondel, V.; Krings, G.; Thomas, I. Regions and borders of mobile telephony in Belgium and in the Brussels metropolitan zone. *Bruss Stud.* **2010**, *42*, 1–12. [\[CrossRef\]](#)
28. Ratti, C.; Frenchman, D.; Pulselli, R.M.; Williams, S. Mobile Landscapes: Using Location Data from Cell Phones for Urban Analysis. *Environ. Plan. B Plan. Des.* **2006**, *33*, 727–748. [\[CrossRef\]](#)
29. Tanahashi, Y.; Rowland, J.R.; North, S.; Ma, K.L. Inferring human mobility patterns from anonymized mobile communication usage. In Proceedings of the 10th International Conference on Advances in Mobile Computing & Multimedia, New York, NY, USA, 3–5 December 2012. [\[CrossRef\]](#)
30. Qi, Y.; Yang, Y.; Jin, F. China's economic development stage and its spatio-temporal evolution: A prefectural-level analysis. *J. Geogr. Sci.* **2013**, *23*, 297–314. [\[CrossRef\]](#)
31. Niu, X.Y.; Wang, Y.; Ding, L. Measuring urban system hierarchy with cellphone signaling. *Planners* **2017**, *33*, 50–56. (In Chinese) [\[CrossRef\]](#)

