Journal of Landscape Research 2024, 16(4): 47-50

Analysis of Urban Agglomeration Network Structure Based on Baidu Migration Data: A Case Study of the Guangdong-Hong Kong-Macao Greater Bay Urban Agglomeration

XIA Yuan[1], WANG Bin[2, 3*]

(1. Guangzhou Urban Planning and Design Co., Ltd., Guangzhou, Guangdong 510060, China; 2. Guangzhou Urban Planning & Design Survey Research Institute Co., Ltd., Guangzhou, Guangdong 510060, China; 3. Guangdong Enterprise Key Laboratory for Urban Sensing, Monitoring and Early Warning, Guangzhou, Guangdong 510060, China)

Abstract The inter-city linkage heat data provided by Baidu Migration is employed as a characterization of inter-city linkages in order to facilitate the study of the network linkage characteristics and hierarchical structure of urban agglomeration in the Greater Bay Area through the use of social network analysis method. This is the inaugural application of big data based on location services in the study of urban agglomeration network structure, which represents a novel research perspective on this topic. The study reveals that the density of network linkages in the Greater Bay Area urban agglomeration has reached 100%, indicating a mature network-like spatial structure. This structure has given rise to three distinct communities: Shenzhen-Dongguan-Huizhou, Guangzhou-Foshan-Zhaoqing, and Zhuhai-Zhongshan-Jiangmen. Additionally, cities within the Greater Bay Area urban agglomeration play different roles, suggesting that varying development strategies may be necessary to achieve staggered development. The study demonstrates that large datasets represented by LBS can offer novel insights and methodologies for the examination of urban agglomeration network structures, contingent on the appropriate mining and processing of the data.

Keywords Baidu migration data, Social network analysis, Urban agglomeration network structure, Greater Bay Area urban agglomeration

DOI 10.16785/j.issn 1943-989x.2024.4.011

An urban agglomeration can be defined as a system of towns and cities that exhibit a tendency towards regional integrated development. These agglomerations are formed with one or more central cities as the core, and are based on well-developed connection channels that facilitate organic links with neighboring cities through a process of division of labor. The spatial structure of an urban agglomeration refers to the geographic distribution characteristics and combination relationships of the functional elements of each city within the urban agglomeration. The spatial structure of urban agglomerations has undergone a gradual transformation in response to the forces of globalization. The traditional "center-hinterland" pattern has been superseded by a network-like spatial structure, which is increasingly evident in the urban landscape. The network structure of urban agglomerations is a network-like spatial pattern that is manifested by the interactions among cities within the urban agglomerations through their intrinsic mechanisms. The study of urban agglomeration network structure can reflect the interaction, linkage pattern, and evolution trend among cities within urban agglomerations. This has important theoretical and practical reference value for the synergistic combination development strategy within urban agglomerations, the spatial structure optimization organization of urban agglomerations, and the urban agglomeration network development strategy.

A substantial body of research has been conducted on the network structure of urban agglomerations, with a primary focus on the characterization of inter-city linkages. The conventional techniques for quantifying the spatial interconnections between cities encompass the gravity model and the urban flow model. A considerable body of research has been conducted on urban agglomeration network structure employing the aforementioned two models. Notable examples include the study on the measurement of the economic linkages of urban agglomerations in the middle reaches of the Yangtze River by Lao Xin et al.[1], the study on the measurement of the internal linkages of China's urban agglomerations by Zhu Xiaochuan et al.[2], the study on the network structure of the economic linkages in the Chengdu-Chongqing urban agglomeration by Cao Weiwei et al.[3], the study of the temporal and spatial evolution of the urban spatial linkages of the Dianzhong urban agglomeration by Yin Juan et al.[4], and the study of urban agglomerations in the middle reaches of the Yangtze River by Hu Ying et al.[5]. These studies can effectively promote the coordinated and integrated development of urban agglomerations, and provide important strategic recommendations for the development of urban agglomerations. However, since both the gravity model and the urban flow model are based on the intrinsic characteristics of cities, there is often a discrepancy between the theoretical and actual interconnections between urban areas. Consequently, an increasing number of scholars have been investigating the network structure of urban agglomerations, focusing on the specific interrelationships between cities, particularly those pertaining to transportation. For instance, Huang Jie et al.[6] conducted research and proposed spatial countermeasures for railroad passenger transportation links in the urban agglomeration situated in the middle reaches of the Yangtze River. Liu Zhengbing et al.[7] employed a combination of inter-city passenger transportation links and the gravity model to study the urban agglomeration in the Central Plains. The aforementioned research, based on the physical links between cities, can objectively reflect the specific links between cities. However, in the context of the digital age, where information is rapidly exchanged and interconnected, the inter-city links are no longer solely based on physical connections. Instead, they are increasingly shaped by information flow

Received: April 16, 2024 Accepted: June 23, 2024
* Corresponding author.

for communication and linkage. Additionally, there has been a notable increase in the number of studies examining urban networks based on information flow. These studies encompass a range of data sources, including Sina Weibo[8-9], Baidu index[10-11], telephone communication data[12], QQ group data[13], and other information flows. These studies, which are based on information flow, serve to complement the original research on urban networks. Furthermore, research based on LBS data, including Baidu migration and Tencent migration data, has emerged as a prominent area of investigation. Notable examples include the study of population mobility patterns between Chinese cities by Liu Wangbao et al.[14], and the study of urban space based on Baidu heat map data by Wu Zhiqiang et al. This study aims to introduce LBS data into the study of urban agglomeration network structure. To this end, the city linkage heat data from Baidu migration data is employed as a characterization reflecting the linkage between cities. The analysis of urban agglomerations employs a variety of techniques, including centrality analysis, to examine the structural characteristics of these networks.

1 Research methods and data

1.1 Research methods

The present study is concerned with the application of social network analysis (SNA). SNA is a method used to study social structure by analyzing relationships and interaction patterns between individuals or organizations to reveal group behavior and social dynamics. In the field of urban studies, SNA is a widely utilized method for understanding social networks, transportation flows, and community structures within urban contexts. Additionally, SNA is employed to examine various social phenomena in urban planning and governance. This study employs the SNA method to examine the maximum attractiveness and urban centrality within the Greater Bay Area urban agglomeration.

The maximum attraction linkage method is a technique employed to analyze and optimize the structure of a network. This is achieved by calculating the attraction between different nodes (cities, transportation hubs, etc.) in order to determine the strongest linkage.

Network density measures the ratio of the number of edges actually present to the maximum number of possible edges. The calculation of network density is as follows:

$$D=2E/N(N-1)$$

In this context, E represents the number of edges in the network, while N denotes the number of nodes in the network. A greater network density is indicative of a greater degree of interconnectivity and frequent migratory activity among the nodes.

Centrality analysis is a method of studying the importance of nodes in a network. It identifies which nodes occupy a central position in the network and play a key role in network connectivity and information flow. The most commonly utilized indicators for centrality analysis include degree centrality, closeness centrality, and betweenness centrality.

Degree centrality is defined as the number of edges that are directly connected to a given node. The degree centrality is calculated as follows:

$$C_D(v)=\deg(v)$$

where $\deg(v)$ represents the degree of the node v, which is defined as the number of edges connected to that node.

Betweenness centrality is a measure of the role played by a node in mediating the shortest path between other nodes in a network. The calculation of betweenness centrality is as follows:

$$C_B(v)=\sum_{s\neq v\neq t}\sigma_{st}(v)/\sigma_{st}$$

where, σ_{st} is the total number of shortest paths from the node s to the node t, and σ_{st} is the number of these paths that pass through the node v.

Closeness centrality is a measure of the average shortest path distance of a node from other nodes in the network. The calculation of closeness centrality is as follows:

$$C_C(v)=1/\sum_{t\in v}d(v,t)$$

where $d(v, t)$ denotes the shortest path distance from the node v to the node t, and V is the set of all nodes in the network.

1.2 Research scope

The research subject selected for this study is the Guangdong–Hong Kong–Macao Greater Bay urban agglomeration. In 2019, the Central Committee of the Communist Party of China (CPC) and the State Council published the *Outline of the Plan for the Development of the Guangdong–Hong Kong–Macao Greater Bay Area*. The document puts forth the establishment of a dynamic, globally competitive urban agglomeration within the Guangdong-Hong Kong–Macao Greater Bay Area, which would serve as a model for advanced and sustainable development. In light of the constraints imposed by the limited availability of data, the present study encompasses nine cities situated within Guangdong Province in the Guangdong-Hong Kong-Macao Greater Bay Area, including Guangzhou, Shenzhen, Zhuhai, Foshan, Huizhou, Dongguan, Zhongshan, Jiangmen, and Zhaoqing, but excludes Hong Kong and Macao. The extant research on the network structure of the urban agglomeration in the Greater Bay Area is primarily anchored in the following domains: transportation and economic links[15-16], science and technology innovation networks[17], etc. This study introduces a novel research perspective and employs Baidu migration data and SNA methods to investigate the network structure of the Greater Bay Area urban agglomeration.

1.3 Research data

The launch of Baidu Migration, a product developed by Baidu Inc. that employs LBS (Location Based Service) technology to track the movement of individuals between cities, generated significant online discussion during the Chinese New Year in 2015. The heat data in Baidu migration represents the intensity of the population flow along a given route. This is calculated by combining migration trips, transportation mode, and migration distance. It is currently being studied as a means of characterizing the strength of inter-city linkages.

The research employs the linkage heat between the two cities within the Guangdong–Hong Kong–Macao Greater Bay Area urban agglomeration on January 4, 2023 in Baidu migration as a measure of the strength of inter-city linkage. This is used to obtain the city linkage strength matrix, which is then subjected to analysis.

2 Analysis of inter–city linkage within urban agglomerations

The network density of the Greater Bay Area urban agglomeration is 100%, indicating that the cities within the Greater Bay Area urban agglomeration are highly interconnected.

3 Analysis of linkage structure within urban agglomerations

Community detection is a method used to identify a collection of nodes with similar characteristics in a complex network. In the case of the urban agglomeration of the Greater Bay Area (Fig.1), the greedy modularity maximization algorithm is employed to detect communities. The results of this analysis reveal the presence of three distinct communities, including Shenzhen–Dongguan–Huizhou, Guangzhou–Foshan–Zhaoqing, and Zhuhai–Zhongshan-Jiangmen.

The maximum attraction link diagram offers a comprehensive representation of the spatial linkage between cities. Cities with a greater number of maximum attraction links are perceived as more attractive within the region and exhibit a higher degree of spatial dominance. The analysis produces a maximum attraction link diagram of the Greater Bay Area urban agglomeration (Fig.2). As illustrated in Fig.3, the analysis reveals that within the Greater Bay Area urban agglomeration, Foshan City exhibits the most pronounced attractiveness, followed by Dongguan.

4 Analysis of urban division of labor within urban agglomerations

Prior to conducting the centrality analysis, the data should be binarized to streamline the analytical process. The mean value is taken as the threshold for the binary boundary line. It is assumed that a link strength exceeding the mean link strength indicates a significant linkage between the two cities, as indicated by the number 1. Conversely, a link strength below the mean value suggests a lack of significance, as indicated by the number 0. This process is employed to obtain the binarized link of the urban agglomeration in the Greater Bay Area. Subsequently, the centrality analysis is employed to obtain the results, as illustrated in Fig.4.

With regard to degree centrality (Fig.4), Guangzhou is identified as the city with the highest degree centrality within the Greater Bay Area urban agglomeration, followed by Foshan. This suggests that Guangzhou occupies a dominant position in external links and regional influence. With regard to betweenness centrality (Fig.5), Guangzhou and Foshan exhibit a higher value, signifying that these cities possess a greater degree of "intermediary" status and that their scheduling and transit functions within the regional network are substantial, serving as pivotal connectors between other cities. In terms of closeness centrality (Fig.6), Jiangmen exhibits the highest value, followed by Shenzhen and Zhaoqing. This indicates that these cities are the most proximate to other cities, suggesting that they can reach other cities in the region with the shortest possible time and possess the advantage of rapid connectivity and efficient services.

Concurrently, an examination of the Greater Bay Area urban agglomeration reveals the absence of a dominant and discernible center, with each city occupying a distinct and pivotal role. In accordance with the tenets of urban agglomeration theory, the nascent stage is typified by a radial spatial structure, the growth stage by a circle radial spatial structure, and the mature stage by a network spatial structure. In light of the preceding analysis indicating that the network density of urban agglomerations is 100%, it can be posited that the urban agglomeration within the Greater Bay Area represents a mature network-like spatial structure.

5 Conclusions and suggestions

First, the Greater Bay Area urban agglomeration represents a mature network-like spatial structure with a network linkage density of 100%. Second, three distinct communities are formed within the Greater Bay Area urban agglomeration: Shenzhen-Dongguan-Huizhou, Guangzhou-Foshan-Zhaoqing, and Zhuhai-Zhongshan-Jiangmen. Ultimately, the various cities within the Greater Bay Area urban agglomeration play distinct and pivotal roles, allowing them to pursue disparate development strategies to achieve a process of staggered development.

From a comprehensive standpoint, Guangzhou and Foshan exhibit the most extensive linkages and the strongest influence in the region. Concurrently, they serve as pivotal connectors between other cities, performing substantial scheduling and transit functions within the region. In light of the aforementioned evidence, it can be posited that these entities can be regarded as "leaders" and "intermediaries" within the urban agglomeration, serving to augment the influence of the industrial chain and reinforce the transportation infrastructure. Shenzhen, Jiangmen, and Zhaoqing are the most proximate to other cities and possess the

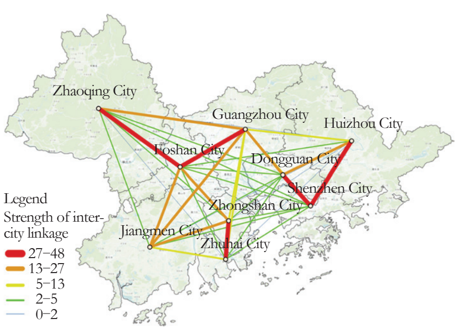

Fig.1 Network linkage of urban agglomeration

Fig.2 Urban maximum attraction link and urban community analysis results

Fig.3 Urban maximum attraction link and urban centrality

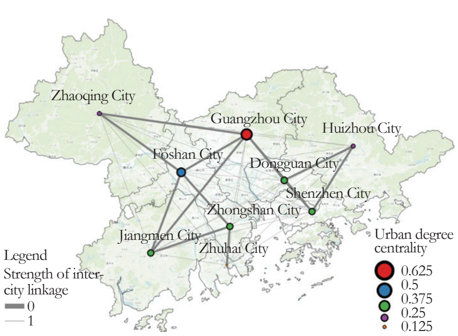

Fig.4 Degree centrality and strength of inter-city linkage

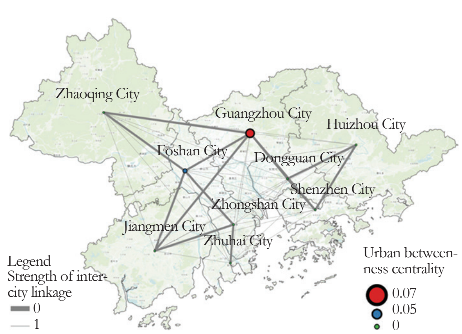

Fig.5 Betweenness centrality and strength of inter-city linkage

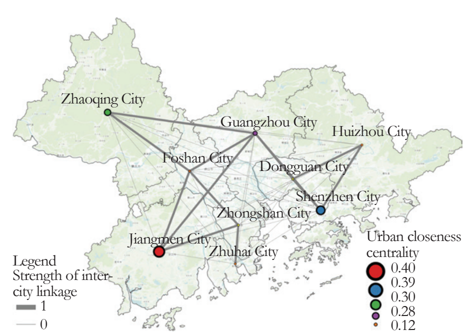

Fig.6 Closeness centrality and strength of inter-city linkage

advantage of rapid connectivity and efficient services, which affords them a competitive advantage in the regional market. In light of these considerations, it is possible to view these entities as "service providers" whose role is to facilitate the continued development of modern service industries, optimize the functional layout of the city, and enhance the speed and quality of the city's external links. The remaining cities have the opportunity to enhance their connections with other cities within the network, thereby boosting their energy levels.

The analysis of social networks based on LBS data from Baidu migration offers a novel perspective for the study of urban area structure. The present study employs a single day's data, which can be synthesized with data from multiple days to examine the spatio-temporal evolution of urban spatial structure. This approach can provide valuable insights that inform the study of urban spatial structure.

References

[1] Lao, X., Shen, T. Y. & Yang, Y. et al. (2016). A study on the economic network of the urban agglomeration in the middle reaches of the Yangtze River: Based on social network analysis method with gravity model. *Urban Development Studies*, (7), 91-98.

[2] Zhu, X. C., Wu, J. W. & Wu, P. P. et al. (2015). The expanded form of gravity model and its application in measuring inner-city linkage of Chinese megalopolies. *Urban Development Studies*, (9), 43-50.

[3] Cao, W. W., Yang, F. & Guan, Y. X. et al. (2016). Economic network structure of urban agglomeration in Chengdu–Chongqing economic zone. *Journal of Technology Economics*, (7), 52-57.

[4] Yin, J., Dong, S. H. & Chen, H. (2015). Temporal evolution of spatial connection among cities of Dianzhong urban agglomeration from 2004 to 2013. *Areal Research and Development*, (1), 65-70.

[5] Hu, Y., Zhang, J. & Liu, Z. H. et al. (2016). The research of the middle reaches of the Yangtze River's spatial contact-based on the gravity model and the urban flow. *Modern Urban Research*, (1), 52-57.

[6] Huang, J., Zhong, Y. X. (2016). Evolution of railway passenger traffic linkage and spatial pattern in Yangtze River middle reaches clusters. *World Regional Studies*, (2), 72-81.

[7] Liu, Z. B., Ding, Z. W. & Bu, S. P. et al. (2015). The network structure analysis of Zhongyuan urban agglomeration based on interactive and traffic flow relationship. *Human Geography*, (4), 79-86.

[8] Chen, Y. X., Zhen, F. & Wang, B. et al. (2012). A study of internet information asymmetry relations among Chinese cities based on the micro-blog platform. *Advances in Earth Science*, (12), 1353-1362.

[9] Zhen, F., Wang, B. & Chen, Y. X. (2012). China's city network characteristics based on social network space: An empirical analysis of Sina Micro-blog. *Acta Geographica Sinica*, (8), 1031-1043.

[10] Xiong, L. F., Zhen, F. & Wang, B. et al. (2013). The research of the Yangtze River Delta core area's city network characteristics based on Baidu index. *Economic Geography*, (7), 67-73.

[11] Zhao, Y. H., Gao, X. & Jiang, B. (2015). The urban network connection of three provinces in northeast China based on Baidu index. *Economic Geography*, (5), 32-37.

[12] Dong, C., Xiu, C. L. & Wei, Z. (2014). Network structure of "space of flows" in Jilin Province based on telecommunication flows. *Acta Geographica Sinica*, (4), 510-519.

[13] Zhao, Y. H., Chen, H. Q. & Yuan, F. et al. (2017). The characteristic and hierarchy structure of urban connection in northeast China based on QQ groups network. *Economic Geography*, (3), 49-54.

[14] Liu, W. B., Shi, E. M. (2016). Spatial pattern of population daily flow among cities based on ICT: A case study of "Baidu Migration". *Acta Geographica Sinica*, (10), 1667-1679.

[15] Yan, X., Liao, C. C. & Zhao, Z. G. et al. (2024). Network structure and coordination mechanism of the Guangdong–Hong Kong–Macao Greater Bay Area in terms of transportation and economic data. *Journal of Geomatics*, *49*(2), 114-121.

[16] Chen, S. D., Xiang, X. M. & Chen, Z. Q. (2024). From field space to flow network: A study on the spatial structure characteristics and optimization path of the urban agglomeration in the Guangdong–Hong Kong–Macao Greater Bay Area. *Lingnan Journal*, (1), 107-119.

[17] Geng, X., Li, H. H. & Qiu, H. T. (2023). Research hotspots and frontiers of scientific and technological innovation cooperation in Guangdong–Hong Kong–Macao Greater Bay Area based on the visual analysis of CiteSpace. *Strategy for Innovation and Development of Science and Technology*, *7*(5), 56-67.

****

(Continued from P46)

needs of rural development, analyze specific problems, prescribe targeted solutions, and avoid the phenomenon of one village for every village. Modern rural construction should fully tap into the regional and cultural characteristics of rural areas. While respecting and inheriting traditional architectural features, it should meet usage needs, introduce modern architectural design concepts and new technologies, and achieve the localization, practicality, comfort, and sustainable development of rural architecture.

References

[1] Zhang, L. (2018). Key agendas for rural revitalization strategy and rural development trends. *Urban and Rural Planning*, (1), 17-23.

[2] Huang, J. J. (2023). Inheritance and innovation of modern rural architectural design under the background of rural revitalization. *Theoretical Research in Urban Construction*, (26), 71-73.

[3] Hu, K. (2023). Renovation design of rural architectural space based on regional culture. *Beauty & Times*, (5), 76-78.

[4] Huang, X. (2023).Innovation in modern rural architecture design under the background of rural revitalization. *Beauty & Times*, (5), 28-30.

[5] Liu, M., Zhou, Z. K. (2023). The expression of defamiliarization of contemporary Chinese rural architecture integrated with digital. *Huazhong Architecture*, *41*(7), 11-15.

[6] Zhou, T. (2023). Design of rural buildings under the background of rural revitalization. *Jushe*, (13), 124-127.

[7] Huang, T. (2023). Design of characteristic rural scenery under the background of rural revitalization. *Designs*, (9), 69-72.

[8] Zhang, B. Q., Zhang, Z. H. (2021). Design strategies for traditional residential renovation in China from the perspective of rural revitalization strategy. *Industrial Construction*, *51*(2), 225.

[9] Liu, H. Y., Wang, X. F. (2023). Exploration of innovative strategies for traditional architecture under the background of rural revitalization. *Interior Architecture of China*, (12), 123-125.

[10] Song, L. P., Zhang, J. & Xu, Z. Y. (2023). Renovation and upgrading of idle rural buildings under the background of rural revitalization. *Housing and Real Estate*, (27), 42-45.