SPATIO-TEMPORAL CHANGE CHARACTERISTICS OF APATIAL INTERACTION NETWORKS: A CASE STUDY WITHIN THE 6th RING ROAD BEIJING

Jing Yang(1), Disheng Yi (1),Jing Zhang* (1)

[1] Capi Norm Univ., 105 W3th Ring Road North, Haidian district, Beijing, 100048, China
Email: 2173602025@cnu.edu.cn; 2180902094@cnu.edu.cn; zhangjings@mail.cnu.edu.cn

KEY WORDS: Spatial Interaction Network; change characteristic; snapshot; k-core decomposition

ABSTRACT: Explorations of human spatial-temporal behaviors are possible because of the occurrence of interactions. Furthermore, the hierarchical structure of urban areas, an important part of geography, can also be discovered. This paper analyzes the spatio-temporal change characteristics of the Spatial Interaction Networks of Beijing (SINB). To begin with, we construct 24 sequential snapshots of spatial population interaction on the basis of points of interest (POIs) collected from Dianping.com and taxi GPS data in Beijing. Then, we use the Jensen–Shannon distance measure and hierarchical clustering to integrate the 24 sequential network snapshots into four clusters. Finally, we improve the weighted k-core decomposition method combining the complex analysis method and weighted distance in geographic space. The results show that there are three layers discovered in the SINB, including the core layer, the bridge layer, and the periphery layer. The number of places varies greatly and the SINB shows an obvious hierarchical structure at the different periods. The core layer, between the 2nd Ring Road and the 5th Ring Road in Beijing, contains fewer places. The bridge layer is located from the 2nd Ring Road to the 5th Ring Road in Beijing, and the number of places in this layer is more than in the core layer. The distributions of places in the periphery layer are the most enormous and wide. Our research plays a vital role in understanding urban spatial heterogeneity and helps to support decisions of urban planning and traffic management.

1. INTRODUCTION

With the wide applications of spatio-temporal big data and improvements of the social sensing concept (Liu et al. 2015), analyzing spatially embedded networks and man–land relationship theory has gained popularity among many researchers. Sensing spatial interaction is one of the important aspects of social sensing. We can obtain traffic flows on the basis of aggregating an individual's or vehicle's trajectories at the collective level. In general, a trajectory begins with an origin point and ends at a destination point. The origin and destination (OD) interactive matrix consists of massive trajectories with the origin and destination points. Spatial assembly is an essential analytical step when aggregating individual-level geospatial data (Liu et al. 2015).

At present, the research on spatial interaction mainly involves the analysis of the spatial interaction network (Xu et al. 2017 and Kang et al. 2013), the spatial interaction pattern (Zhu et al. 2017; Takeuchi et al. 2017 and Tao et al. 2018), the spatial interaction strength model (Simini et al. 2012 and Liu et al. 2014), and the visualization of spatial interaction (Wood et al. 2010 and Yao et al. 2019). Research on the spatial interaction network depends on the approach of complex network analysis in order to analyze topological properties of the network, including the k-core value, centralization index, clustering coefficient, and average path length, etc. For example, Lordan and Sallan (2017) and Du et al. (2016) applied the k-core decomposition method to analyzing the hierarchical structure of the European airport network and Chinese airline network, respectively. Wang and her colleagues (2014) analyzed the evolution process of the air transport network of China from 1930 to 2012 by employing the network centralization index and the k-core network method. Previous studies have accomplished a large amount of meaningful work. Compared with them, this research takes a somewhat different approach: 1) Places are regarded as

[1] This work was supported by the Open Project Program of the State Key Laboratory of Virtual Reality Technology and Systems, Beihang University (Grant No.011177220010020 and No. 01119220010011)

the unit of spatial analysis and 2) we redefine the weighted k-core decomposition method in view of the weight of the node and edge to obtain the importance of nodes.

Our research, at the beginning, extracts the footprints of places using POI data as units. Next, the day is divided into 24 60-minute intervals and the 24 OD metrics are constructed by taxi GPS data accordingly. We merge them into four clusters for the following research because of their similarities. At last, we identify the evolution characteristics and hierarchy structure of Spatial Interaction Networks of Beijing (SINB) on the basis of the k-core decomposition method. The remainder of this paper is organized as follows. Section 2 introduces the data and processing. Section 3 presents the methods of the Jensen–Shannon distance measure, hierarchical clustering, weighted k-core decomposition, and additional methods. Section 4 analyzes the hierarchical characteristics of SINB at four periods of the day. Section 5 provides a summary, as well as a brief discussion of future relevance of the work

2. DATA DESCRIPTION and PROCESSING

2.1 Study Area and Data Description

We chose the area within the 6th Ring Road in Beijing as the study area, which covers 2,267 km2, accounting for 8% of the total area of Beijing (Figure 1). Taxi data is widely used in analyzing urban functions, urban structures, and human mobility patterns. This research applies a taxi dataset collected from Beijing, China, including more than 15,000 taxis from several anonymous taxi companies in consecutive weeks (6 June to 3 July 2016). In this paper, we only extract the taxi's ID, the time when the passengers are picked up and dropped off, and the location where the passengers are picked up and dropped off. Table 1 shows an example of the processed taxi data. POIs collected from Dianping.com are also used in this study to extract the footprint of places. We capture nearly 80,000 businesses in Beijing on 6 June 2016 from Dianping.com. Then, we preprocess these data into the following categories: longitude, latitude, name, places, and abnormal points. The data record of a processed POI sample is shown in Table 2.

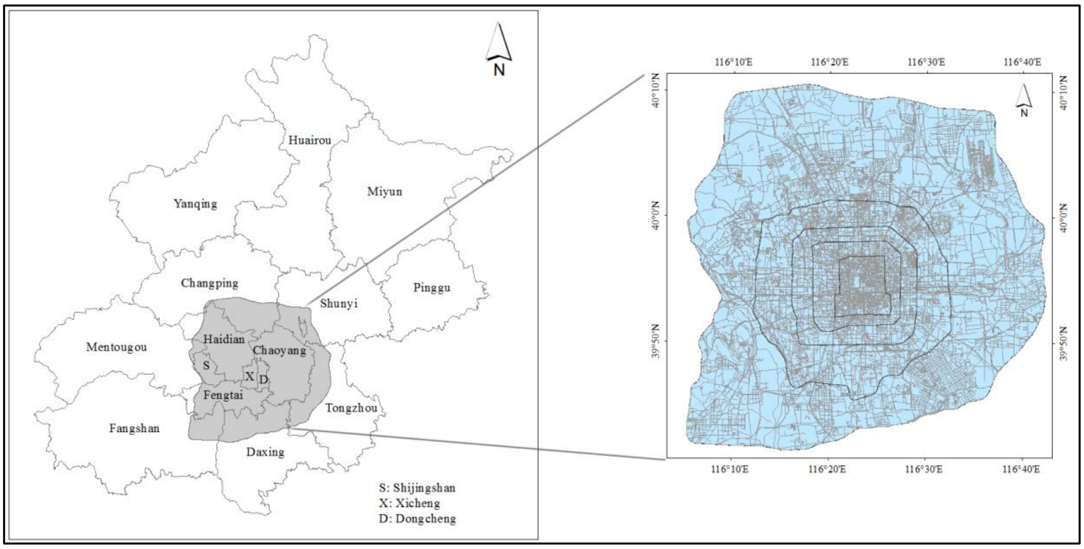

Figure 1. Study area in Beijing, denoted by blue.

Table1. Sample trip with pick-up and drop-off labels.

Taxi ID	Pickup time	Pickup coordinate	Drop-off time	Drop-off coordinate
158	2016-6-6 6:27:21	116.45806E 39.98764N	2013-6-6 6:44:5	116.40218E 39.94539N
2056	2016-6-10 0:2:44	116.58275E 40.07931N	2013-6-10 0:31:47	116.28463E 40.02774N

300 24	2016-6-29 6:40:2	116.45534E 39.94887N	2013-6-29 7:11:26	116.58095E 40.07179N

Table 2. The business data from Dianping.com data

Place name	Business name	Latitude	Longitude
Wudaokou	Yuye	39.99102N	116.3353E
Wangjing	Bafu	39.996445N	116.4815E
Sanlitun	Hema	39.93144N	116.4535E

2.2 Extracting the Footprints of Places

In this research, we applied a three-step workflow to extract continuous boundaries of cognitive places inside the 6th Ring Road of Beijing. In step 1, we apply the fuzzy set method based on adaptive kernel density estimation (Wang et al. 2014) for generating scores of each place using POIs. In step 2, the probability density of each place is converted into the membership degree by the membership function. In step 3, we constructed polygons from the footprints based on membership degree contours. As a result, we obtain 150 places within the 6th Ring Road of Beijing, as shown below.

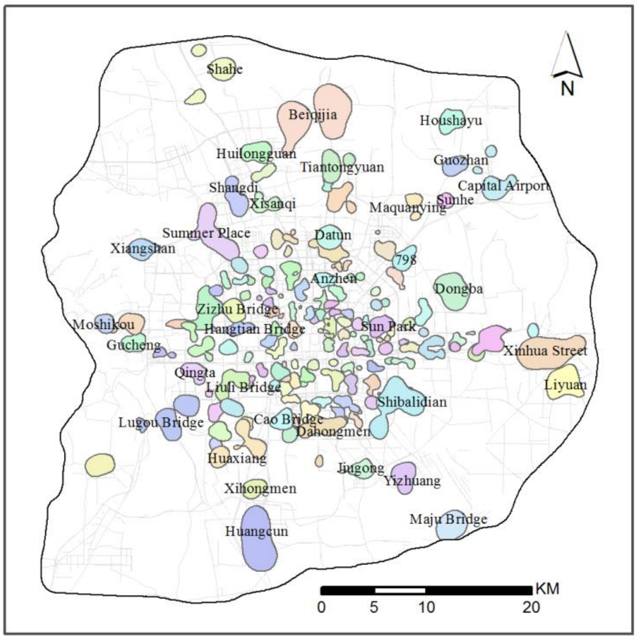

Figure 2. The distribution of the 150 places within the 6th Ring Road of Beijing.

2.3 Construction of Spatial Interaction Networks

We construct the Spatial Interaction Networks of Beijing (SINB) with places and their connections. The nodes of it are those places extracted before and the edges are the inter-place links connected by taxi trips. The SINB is defined as an undirected weight network $\mathrm{G} = (\mathrm{V}, \mathrm{E})$. The node set is defined as $\mathrm{V} = \{\mathrm{v}_1, \mathrm{v}_2, \cdots, \mathrm{v}_\mathrm{n}\}$, where n is the number of nodes. Additionally, the edge set is defined as $\mathrm{E} = \{\mathrm{e}_1, \mathrm{e}_2, \cdots, \mathrm{e}_\mathrm{m}, \}$, where m is the number of edges. In order to investigate the temporal characteristics of networks, we discretize the taxi data by hourly intervals. All trips extracted from the taxi trajectory data are aggregated based on the places. The 30-day taxi trip data are aggregated into one day to generate the 24 OD matrices from the original taxi data. The matrix is set as the one-hour taxi frequency running between nodes. We define the OD matrix from 0:00 to 1:00 as the initial network snapshot G_0, the second one G_1 is set from 1:01 to 2:00, and the last one G_{24} is set from 23:01 to 23:00.

3. METHODS

3.1 Jensen–Shannon Distance between Snapshots

Our goal is to examine layer similarity to choose the aggregation of a pair of similar layers rather than the aggregation of two very dissimilar layers. The quantum spectral Jensen-Shannon divergence between two layers was proposed as a similarity measure and clusters layers for networks (De Domenico and Biamonte, 2016; De Domenico et al. 2015). In general, the Jensen–Shannon distance is evolved from the Kullback–Liebler distance and is a more suitable quantity to measure the dissimilarity between two matrices than the Kullback–Liebler distance. The distance measure based on the Jensen-Shannon divergence is given by

$$D_{ij} = \sqrt{S_{(r)} - \frac{S_{(p)} + S_{(q)}}{2}}$$

where D_{ij} represents the Jensen–Shannon distance snapshots i and j, and takes values of [0,1]. The two density matrices p and q correspond to networks $G1$ and $G2$, respectively, and $r = \frac{p+q}{2}$. Neumann entropy is defined by $S_{(p)} = -\sum_{i=1}^{N} \lambda_i \log_2(\lambda_i)$, where λ_i is the ith eigenvalue of p. Neumann entropy $S_{(r)}$ and $S_{(q)}$ corresponds to matrices p and r, respectively.

3.2 Hierarchical clustering method

Hierarchical clustering merges the two nearest variables from bottom to top until all the variables are merged into an entire variable, while hierarchical divisive clustering is the reverse process. In this paper, the hierarchical clustering algorithm is applied to process the distance matrix between the snapshots. We use the shortest Euclidean distance to define the distance between clusters, and proceed by repeatedly applying three steps:
(1) Using the distance between snapshots as variables;
(2) Computing the similarity of the pair of variables based on a given distance measure and merging the highest similarity pair into a new cluster;
(3) Updating the similarity between the new cluster and the former existing variables, repeating the procedure until only one node is left.

3.3 The Weighted K-core Decomposition Method

The definition of k-core, firstly introduced by Seidman (Seidman, 1983), is of fundamental importance to detecting the hierarchical structure and finding the relationship between the substructures and a visual representation of the network. A k-core method is derived by recursively removing all the nodes with a degree that is greater than or equal to k until all nodes in the remaining network have a degree of at least k (Carmi et al. 2017).
Based on the weighted network decomposition proposed by Caras(Garas, Schweitzer and Havlin, 2012), this paper redefines the weight of the link considering the spatial weight (distance between nodes) in geographical space. Distance is an important factor affecting spatial interaction which determines the cost of time and interaction. If there are the same spatial interactions and different distances between two nodes, the further the distance, the more important the connection they have. As shown in Figure 3, if $S_{AB} = S_{AC}$, $D_{AB} > D_{AC}$, the weight of link AB is more important than link AC. Improving the weight of the link based on the interaction distance, the new weighted degree of a node i is defined as

$$k_i' = \left\{ k_i^{\alpha} \left[\sum_{j}^{k_i} \left(S_{ij} * D_{ij} \right) \right]^{\beta} \right\}^{\frac{1}{\alpha+\beta}}$$

where k_i is the degree of nodes; S_{ij} and D_{ij} represent the amount and distance of links between the origin node i and destination node j, respectively; and α and β are the adjustment parameters of node degree and node weights. We only consider the case when $\alpha = \beta = 1$ in this research. In this implementation, we normalize all the amount and distance of links. Similarly, our methods could be used to direct the network.

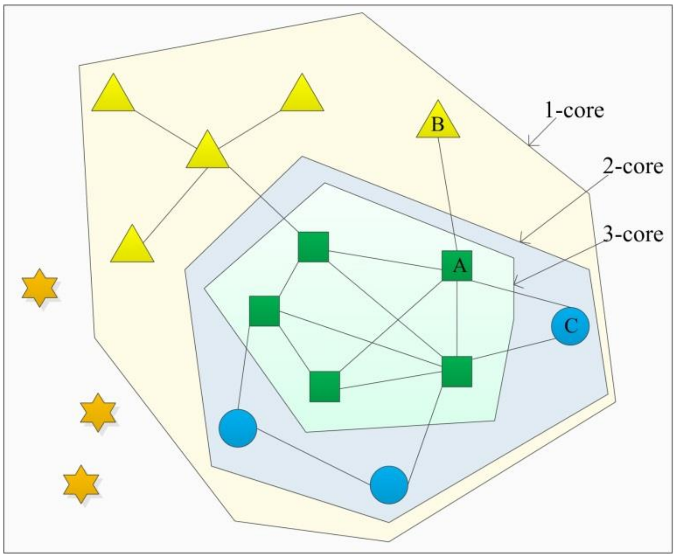

Figure 3. Illustration of the k-core decomposition method

4. RESULT

4.1 Layer Aggregation of SINB

The result of the similarity matrix and hierarchical clustering are shown in Figure 5(a) and Figure 5(b), respectively, on weekdays. The x-axis and y-axis represent the 24 snapshot networks in Figure 5(a), and the heat map represents the similarity between pairs of layers, and the darker the color, the higher the similarity. The x-axis represents the 24 snapshot networks, and the y-axis is the interclass distance in Figure 5(b). The result of this procedure is a dendrogram that is a hierarchical diagram where some snapshot networks are the original leaves. At each step of the algorithm, we iteratively aggregate these layers or the clusters of layers with a minimal distance as an internal node. Finally, all original leaves correspond to one root. We obtain a multilayer with fully aggregated graphs.

We calculated the ratio of between-cluster variance to the total variance for each possible k from 2 to 8 to determine the number of clusters k that best divide the data. And we find the variance ratio is small when k = 4 on weekdays and weekdays. As a result, we divide the 24 snapshot networks into four temporal characteristics: early morning (0-6AM), morning (7-10AM), afternoon (11AM-4PM), and evening (5-11PM). For the weekends, the heat maps of the similarity matrix and the hierarchical diagram are shown in Figure 5(c) and Figure 5(d), respectively. Here, we also choose a distance of 8 as the divider. Therefore, there are considerable differences between weekdays and weekends from the results of clustering. The periods of time are early morning (0-4AM), morning (5-11AM), afternoon (12AM-5PM), and evening (6-11PM).

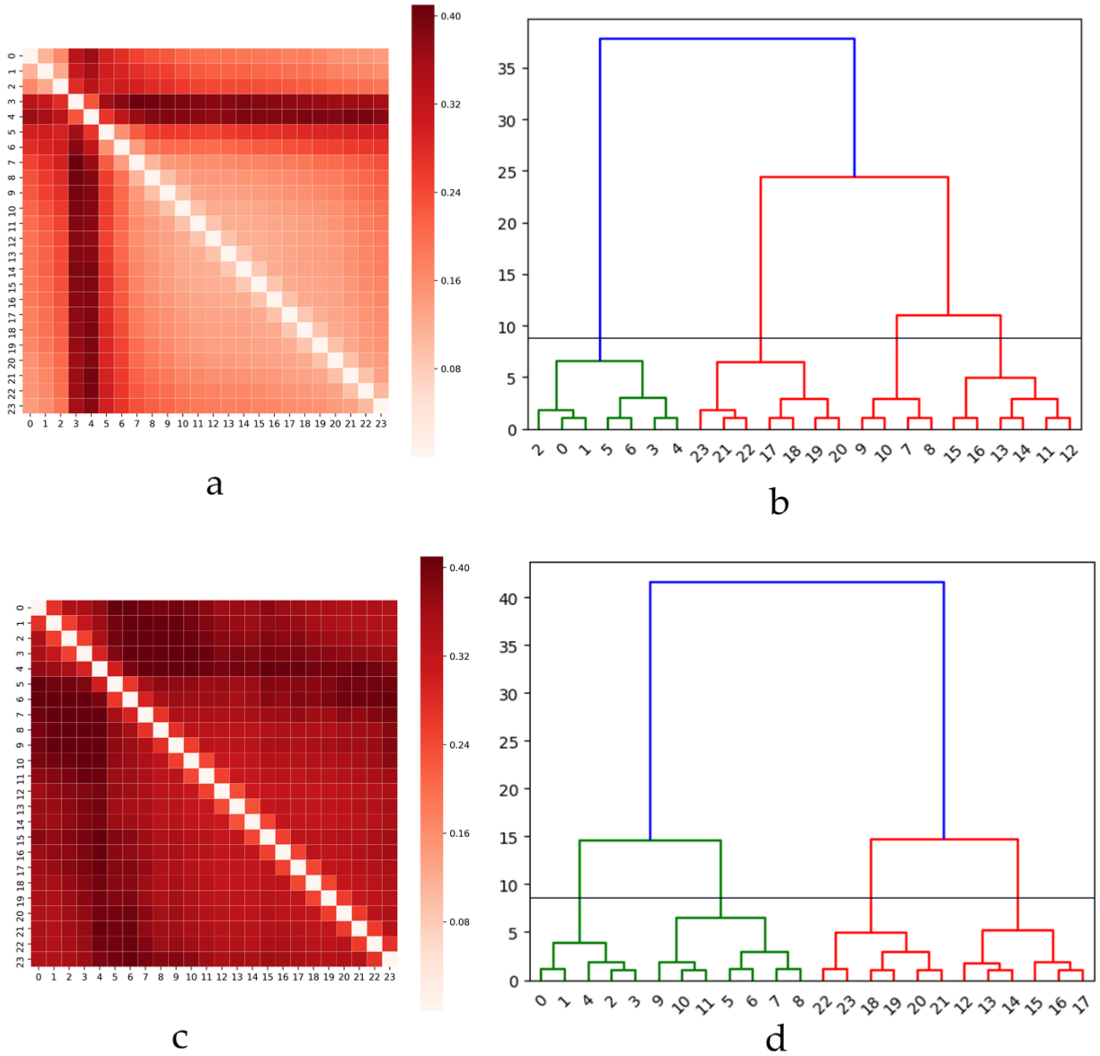

Figure 5. Similarity matrix and hierarchical clustering. (a) The similarity matrix between snapshots on weekdays; (b) the hierarchical clustering on weekdays; (c) the similarity matrix between snapshots on weekends; and (d) the hierarchical clustering on weekends;

4.2 The Evolving Properties of the SINB on Weekdays

Figure 6 illustrates the distribution of all the places we studied in the three layers by order of four periods on weekdays. The important places in the spatial interaction network exist in the core layer. As a result, the main places are not the same for the different periods. Specifically, there are three places extracted as the main places, including the Asian Games Village-Xiaoying area, Beijing Capital International Airport, and Datun in the early morning. The 14 main places are discovered in the eastern and northern areas of the city, such as Beijing Capital International Airport, Datun, the Asian Games Village-Xiaoying area, Zhongguancun, Wangjing, North Taipingzhuang, and Dawang Road, in the morning. In the afternoon, moreover, the 24 identified main places with a peak level are similar to those places at the last period. A wide distribution occurs in the other three directions between the 3th and 4th Ring, except in the south of the city. The number of main places drops at night. Only two places (Asian Games Village-Xiaoying area and Datun) showed up at this period. The interesting thing is that these two places appear durably in the core layer all day. This is because they have already become a comprehensive downtown area with offices, businesses, residences, and recreational areas since the Olympic Games and Asian Games were held there, respectively. At the same time, some new arrivals, like institutes which have a profound effect on the education industry, lead to a strong relation with other places. The bridge layer connects the core and the periphery as a transition. Places in this layer mainly distribute between the 2th and 5th Ring of the city. There are the same numbers of places in the early morning and in the morning, but different places. The main places include the Summer Palace, Xisi, Chaoyang Park-Tuanjie Lake, Beiyuanjiayuan, and Jinsong-Panjiayuan in the former period and then move to places like Sijiqing, Wukesong, Jianwai Avenue, Guanzhuang-Changying, and the Yansha-Agricultural Exhibition Center. A decreasing trend occurs in the afternoon, during which only 33 places appear, such as Shuangjing, Wudaokou, Xueyuan Bridge, Beiyuanjiayuan, and Zizhu Bridge. However, the number of places increases at night, with a value of 40. Those places, including Beiyuanjiayuan, Beijing Capital International Airport, Jinsong-Panjiayuan, Shibalidian, and Chaoyang Park-Tuanjie Lake, reflect the property of this layer. Some residential districts display repeated emergence in the four periods, like Zizhu Bridge, Xueyuan Bridge, Wudaokou, Qing River, Shuangjing, Dongzhimen, and Baiziwan.

The peripheral layer is by definition the periphery of the network with a relatively low importance. Compared to the other two layers, the magnitude of places is greater in order of time period (106, 95, 92, and 108 account for 71%, 63%, 61%, and 72%, respectively). There are constant places responsible for 53% of the total places in the peripheral layer. Also, their locations are variously distributed from the Ring 2th to the Ring 6th.

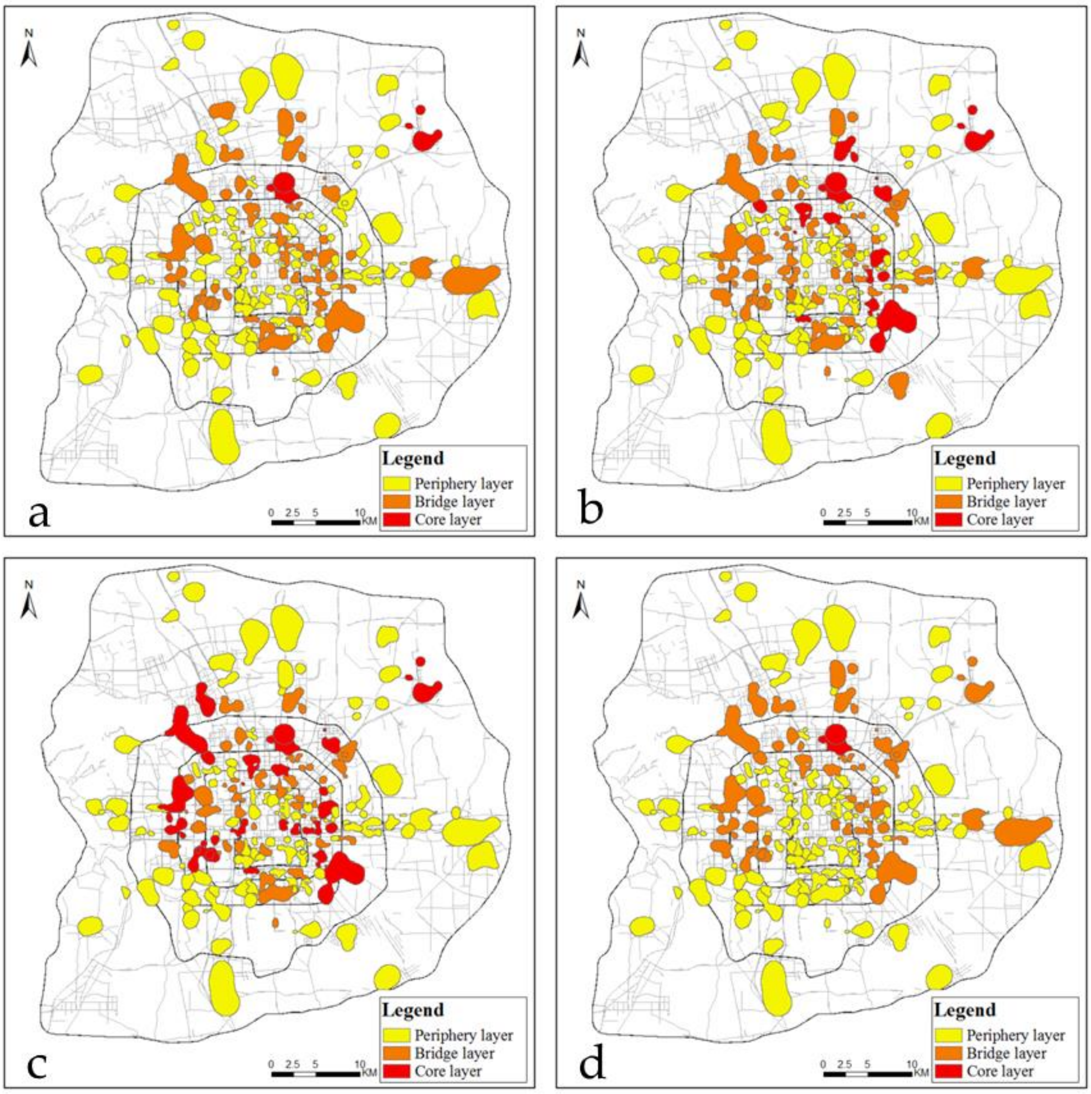

Figure 6. The distribution of all places at the different periods in three layers on weekdays. (a) In the early morning; (b) in the morning; (c) in the afternoon; and (d) in the evening.

4.3 The Evolving Properties of the SINB on Weekends

In contrast to the spatial portions of places on weekdays, they fluctuate in number remarkably in the core layer, with a quiet morning and a busy afternoon at weekends. There are 23 main places located with an interruption of south of the city in the early morning. Only two places, however, appear in this layer in the morning, including Asian Games Village-Xiaoying and Beijing Capital International Airport. The 23 increasing places spread across the regions between the 2nd and 5th Ring Road in the afternoon. In addition, northern and eastern Beijing contributes 15 places, which are all places in this layer that appear at night. The stable existence of two places (Asian Games Village-Xiaoying and Beijing Capital International Airport) take us by surprise in this layer at weekends.

Most places in the bridge layer are distributed popularly from the 2nd to 5th Ring of city and the locations of others are sprinkled throughout the external frontiers. The main places are discovered as Jianguomen-Beijing Railway Station, Wangjing, and the Summer Palace (in the early morning); Datun, Shuangjing, and Wangjing (in the morning); Jiuxian Bridge, Sanlitun, and Huangtian Bridge (in the afternoon); and Wukesong, Sijiqing, and Huangtian Bridge (at night). What is more, there are six constant places in this layer, including Jiuxian Bridge, Hangtian Bridge, Dewai Avenue, Dahongmen, the Yansha-Agricultural Exhibition Center, and Yang Bridge-Muxuyuan.

The number of places in each of the four periods in the peripheral layer is as follows: 84, 114, 76, and 90, comprising 56%, 76%, 51%, and 60% of total places, respectively. There is a large proportion of places in this layer in the morning. This is because residents do not walk about too early and prefer to choose haphazard trips at the weekends. Despite the fact that there is a time irregularity in the four periods, 40% of places like Beidadi, Beijing Eastern Railway Station, Caishikou, and Caofang always show up.

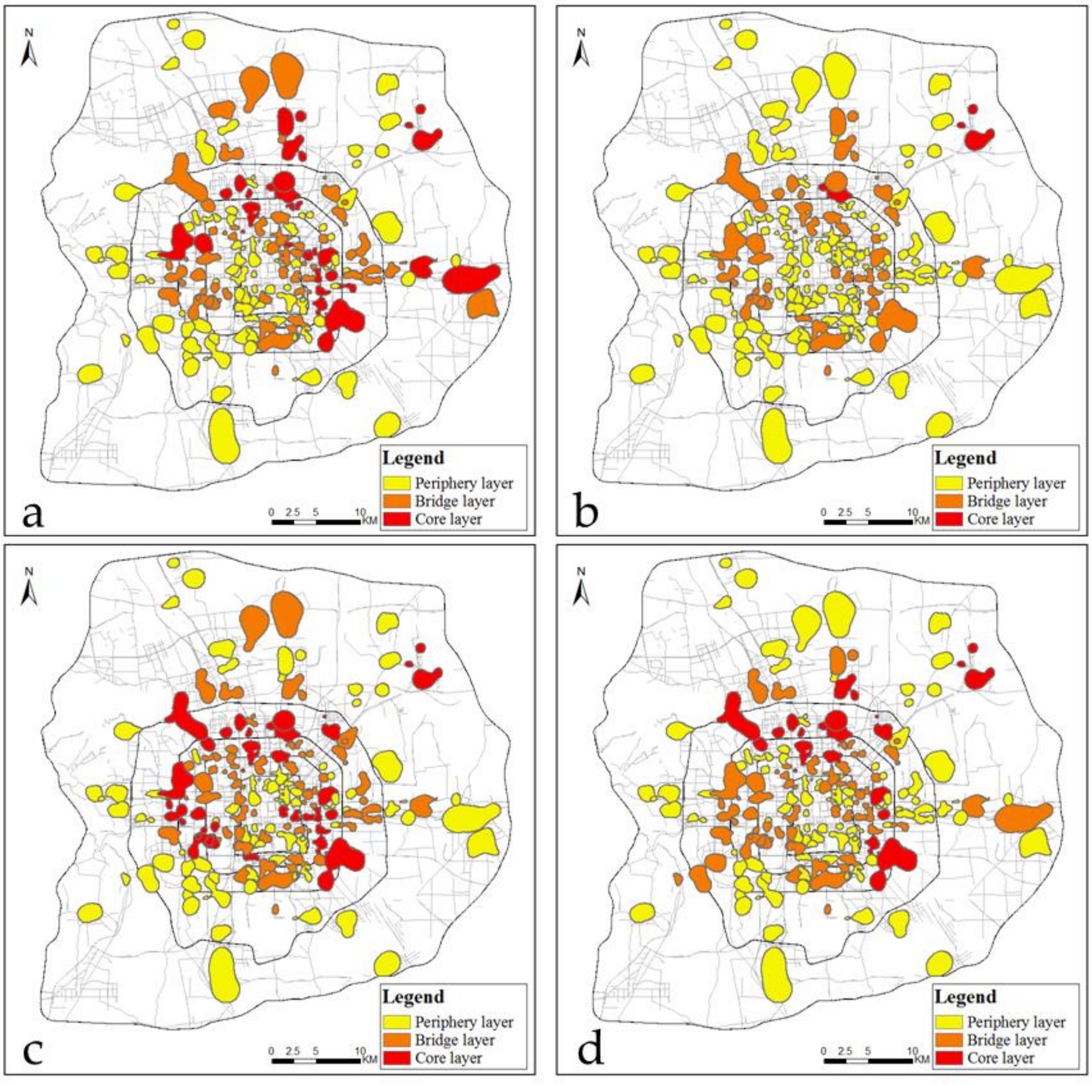

Figure 8. The distribution of all places at the different periods in three layers on weekends. (a) In the early morning; (b) in the morning; (c) in the afternoon; and (d) in the evening.

5. SUMMARY and DISCUSSION

Spatial interaction networks play a crucial role in the human mobility transportation of economics, culture, and the spread of transport diseases. The SINB is also dynamic, which reflects the passenger mobility daily and further reveals the spatial structure. In this paper, the SINB has been studied from the perspective of multi-layer temporal networks。

Spatial assembly is quite different between three layers in the different periods. A small and stationary distribution is found in the core layer, but the dispersal of places is throughout the area between the 2nd and 5th Ring, as well as Beijing Capital International Airport. Secondly, the number of places in the bridge layer is relatively higher than in at core, while their distribution is similar. However, the peripheral places are cosmic throughout the entire study area. At weekends, however, the situation changes. Gathering the connections in the peripheral layer in the morning becomes normal, and different layers attract internal connections between places in all the periods except for the morning.

Our research offers a new perspective for analyzing urban spatial structures based on the importance of nodes, yet there are limitations of our study. The SINB in the real world are changing over time. Choosing an appropriate time interval could help us to explore the characteristics of the SINB better. As a result, the first problem is our fixed time interval, which might cause a huge deviation of consequence. In future research, the length of the time span should change dynamically to adapt the temporal features in the network.

References

Carmi S, Havlin S, Kirkpatrick S, et al., 2007. A model of Internet topology using k-shell decomposition[J]. Proceedings of the National Academy of Sciences, pp.104(27): 11150-11154.

De Domenico M, Biamonte J., 2016. Spectral entropies as information-theoretic tools for complex network comparison[J]. Physical Review X, pp.6(4): 041062.

De Domenico M, Nicosia V, Arenas A, et al., 2015. Structural reducibility of multilayer networks. Nature communications, pp.6: 6864.

Du W B, Zhou X L, Lordan O, et al., 2016. Analysis of the Chinese Airline Network as multilayer networks . Transportation Research Part E: Logistics and Transportation Review, pp.89: 108-116.

Garas A, Schweitzer F, Havlin S., 2012. A k-shell decomposition method for weighted networks . New Journal of Physics, pp.14(8): 083030.

Kang C, Zhang Y, Ma X, et al., 2013. Inferring properties and revealing geographical impacts of intercity mobile communication network of China using a subnet data set. International Journal of Geographical Information Science, pp.27(3): 431-448.

Liu Y, Liu X, Gao S, et al., 2015. Social sensing: A new approach to understanding our socioeconomic environments. Annals of the Association of American Geographers, pp.105(3): 512-530.

Liu Y, Sui Z, Kang C, et al., 2014. Uncovering patterns of inter-urban trip and spatial interaction from social media check-in data . PloS one, pp.9(1): e86026.

Lordan O, Sallan J M., 2017. Analyzing the multilevel structure of the European airport network . Chinese Journal of Aeronautics, pp.30(2): 554-560.

Seidman S B., 1983. Network structure and minimum degree[J]. Social networks, pp.5(3): 269-287.

Simini F, González M C, Maritan A, et al., 2012. A universal model for mobility and migration patterns. Nature, pp.484(7392): 96.

Takeuchi K, Kawahara Y, Iwata T., 2017. Structurally Regularized Non-negative Tensor Factorization for Spatio-Temporal Pattern Discoveries//Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer, Cham pp.582-598.

Tao H , Wang K , Zhuo L , et al. Re-examining urban region and inferring regional function based on spatial–temporal interaction. International Journal of Digital Earth, pp.2018(3):1-18.

Wang J, Mo H, Wang F, 2014. Evolution of air transport network of China 1930–2012 . Journal of Transport Geography, pp.40: 145-158.

Wang S, Liu Y, Chen Z, et al., 2018. Representing multiple urban places' footprints from Dianping.com data . Acta Geodaetica et Cartographica Sinica, pp.47(8): 1105-1113.

Wood J, Dykes J, Slingsby A., 2010. Visualisation of origins, destinations and flows with OD maps. The Cartographic Journal, pp.47(2): 117-129.
Xu J, Li A, Li D, et al., 2017. Difference of urban development in China from the perspective of passenger transport around Spring Festival. Applied Geography, pp.87: 85-96.
Yao X, Wu L, Zhu D, et al., 2019. Visualizing spatial interaction characteristics with direction-based pattern maps[J]. Journal of Visualization, pp.22(3): 555-569.
Zhu D, Huang Z, Shi L, et al., 2017. Inferring spatial interaction patterns from sequential snapshots of spatial distributions. International Journal of Geographical Information Science, pp.(4):1-23.