

Article

Spatiotemporal Evolution Analysis of the Chinese Railway Network Structure Based on Self-Organizing Maps

Lingzhi Yin ¹ and Yafei Wang ^{2,*}

¹ School of Computer Science and Technology (School of Artificial Intelligence), Zhejiang Sci-Tech University, Hangzhou 310018, China

² Key Laboratory of Regional Sustainable Development Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

* Correspondence: wangyafei@igsnrr.ac.cn

Abstract: Delving into the spatiotemporal evolution of the railway network in different periods can provide guidance and reference for the planning and layout of the railway network. However, most of the existing studies tended to model the railway data separately and compare the network indices of adjacent periods based on the railway data of different periods, thus failing to integrate the railway network in different periods into a unified framework for evolution analysis. Therefore, this paper used the railway data from 2008, 2010, 2015, and 2019, and analyzed the spatiotemporal integration of the railway network evolution based on the complex network theory and the self-organizing maps (SOM) method. Firstly, this study constructed the geographical railway network in the four years and probed into how the network feature indices changed. Then, it used the SOM method to capture the spatiotemporal integration of the railway network evolution in multi-time series. Finally, it clustered the change trajectory of each city node and unveiled the relationship between the evolution of city nodes and the hierarchy of urban systems. The results show that from 2008 to 2019, the railway network feature indices showed an upward trend and that the expansion pattern of the railway network could be divided into the core-peripheral pattern, belt expansion pattern, strings of beads pattern, and multi-center network pattern. The evolution of the change trajectory of the city nodes was highly related to the hierarchical structure of the urban system. This study helps to understand the evolution process of the railway network in China, and provides decision-making reference for improving and optimizing China's railway network.

Keywords: Chinese railway network; complex network; self-organizing maps; spatiotemporal clustering; trajectory analysis



Citation: Yin, L.; Wang, Y. Spatiotemporal Evolution Analysis of the Chinese Railway Network Structure Based on Self-Organizing Maps. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 161. <https://doi.org/10.3390/ijgi12040161>

Academic Editors: Hartwig H. Hochmair and Wolfgang Kainz

Received: 8 February 2023

Revised: 5 April 2023

Accepted: 7 April 2023

Published: 10 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Transportation infrastructure has long been considered a core pillar of a country in underpinning its comprehensive development and social progress. As one of the most essential components of transportation infrastructure, railways play an irreplaceable role in long-distance passenger transport and communication between cities thanks to their advantages of a large capacity, high efficiency, low energy consumption, and low pollution [1–4]. With the rapid economic growth, an increasing number of railways are under planning, construction, and operation, thus continuously expanding the scale of the railway networks. By the end of 2021, China's railway operating mileage had reached 150,000 km, of which the high-speed railway operating length reached 40,000 km. As a result, the previously simple railway network was transformed into a complex one [5–7]. The layout of the railway network is an intuitive expression of interchange, communication, and connection between cities. The evolution and spatiotemporal characteristics of the railway network can indirectly reflect the development process of cities and their economic and social conditions [8,9]. Therefore, it is of great significance to clarify the trend of the railway network structure during its evolution for efficient development.

The complex network is a useful tool to abstract the physical railway network into a mathematical network, and to analyze the relationship between network features and their behavior [10–13]. Through mathematical network modeling, the complex railway stations and lines are converted into a mathematical model with a considerable number of connected edges and nodes, so as to describe the evolution mechanism of the railway networks [14–16]. There are a number of studies on the structural characteristics of the railway network based on complex network theory. Ghosh et al. explored the network properties of the Indian railways through a complex network [12]. Cao et al. used complex network theory to study the heterogeneity and hierarchical properties of China's high-speed railway network [17]. Lam et al. proposed a complex network analytical framework to identify the factors and effects of railway incidents [3]. Yin et al. built a geographical railway network using complex network theory and analyzed the vulnerability of the railway network under geological disasters [4]. However, these studies only reveal the static structural characteristics of the railway network. China's railway construction has developed rapidly, and the spatial pattern of the network has undergone dramatic changes. At the same time, policies are constantly adjusting and changing during the process of rapid railway construction. Exploring the evolution characteristics of the railway system can also serve as a reference and guidance for policy formulation, which can avoid short-term and short-sighted decision making. Therefore, it is necessary to explore the spatiotemporal characteristics of the railway network to make out a pattern of its evolution over time.

Currently, there are several studies on the spatial pattern changes of China's railway network structure in different development periods based on complex network theory. For example, Wang et al. analyzed the expansion of China's railway network, the evolution of its spatial accessibility, and the impacts on economic growth and urban systems over a time span of about one century (1906–2000) [18]. Huang et al. used complex network analysis and network indices to analyze the evolving network characteristics of China's railway network during each of the four main stages of China's high-speed rail (HSR) development over a 10-year period [19]. Lu et al. analyzed the hierarchical and spatial heterogeneity distribution of train flows based on the four main stages of China's railway development from 2008 to 2017 using a classical community detection algorithm [20]. Wei et al. used the HSR passenger flow data from the timeframe of 2014 to 2018 in the Yangtze River Delta to examine the evolution of the topological structure, hierarchical structure, and spatial structure of the HSR network [21]. He et al. used the dwell time to simulate the passenger flow between stations and complex network analysis to investigate the evolution of the network's spatial structure at regional and local scales [22]. However, most of these studies modeled the railway data in different periods separately, and compared the network feature indices of two or more adjacent periods for the evolution analysis, thus failing to integrate the railway network in different periods into a unified framework for evolutionary analysis.

Self-organizing maps (SOM), as a spatiotemporal analysis approach, can produce a two-dimensional plane picture of the spatiotemporal change information through dimension reduction to visualize the spatiotemporal pattern [23,24]. Currently, SOMs have been widely used in spatiotemporal evolution studies such as those on socio-economic changes [25], epidemics [26], crime [27], and land use [28], yet those on railway networks are relatively scarce. Therefore, this paper used the SOM method to analyze the spatiotemporal evolution pattern of the railway network structure. By setting a sufficiently large size of the SOM output panel, with the help of SOM topology retention characteristics, the multi-time series data of railway network features were simultaneously input into the SOM network for training and output, so as to integrate railway networks in different periods and conduct structure evolution analysis.

2. Data

With the rapid development of the national economy, China's railway network has become increasingly perfect. During its rapid development, China's railway system generally experienced the following stages: the opening of the Beijing–Tianjin Intercity Railway in

2008, which marked China's official entry into the era of high speed as it owned the first railway operating at a running speed of more than 300 km/h; in 2011, the 7.23 Ningbo–Wenzhou Railway Accident precipitated a trough period in China's railway system, where both the speed and standards declined; in 2016, the revision of the Medium and Long Term Railway Network Plan (2016–2030) signaled China's entry into a “take-off” period along with the proposal of the eight vertical and eight horizontal high-speed railway blueprint. However, the outbreak of COVID-19 in late December 2019 suspended the operation of trains in many cities, so the operation of trains in the ensuing years is not representative of the real operation status of China's railways. Under such circumstances, this paper selected the train operation data from 2008, 2010, 2015, and 2019, each corresponding to a different stage of development, to analyze the spatiotemporal evolution of China's railway network.

The train operation data from the above-mentioned years were collected from the China's Railway website (<https://www.12306.cn> (accessed on 20 June 2019)), Ctrip.com International Ltd. (CTRP) and the Shikai Railway Timetable. The range of the data only includes mainland China, meaning that the data from Taiwan Province are excluded. The data of the railway timetable mainly include the train number, departure and arrival station, arrival time, and departure time and stop time, as shown in Table 1. The types of passenger trains mainly include high-speed EMUs (G), intercity high-speed trains (C), EMUs (D), direct express passenger trains (Z), express passenger trains (T), fast passenger trains (K), and regular passenger trains (four-digit train number).

Table 1. An example of a railway timetable.

Train No.	Sequence	Station Name	Arrival Time	Departure Time	Stop Time (min)
C1215	1	Changchun	-	12:28	-
	2	Longjia	12:42	12:44	2
	3	Jilin	13:14	-	-

3. Methods

The overall framework diagram is presented in Figure 1. First, the railway timetable data for the four years 2008, 2010, 2015, and 2019 were collected. Then, supported by the complex network theory, the space P modeling method was used to model the railway networks for the four years, and the network feature indices, such as the degree, weighted degree, betweenness centrality and closeness centrality, were selected and analyzed for each year. Finally, the SOM method was used to further study the spatiotemporal evolution pattern of the railway network.

3.1. Construction of the Geographical Railway Network

The space P modeling method treats the city where the railway station is located as the node, and the connection between any two nodes as the edge if there is a direct train between them. All nodes with the same train number form a fully connected graph, while the nodes on different railway lines have no directly connected edges. The space P model focuses on the accessibility and service efficiency between nodes, which accords with people's cognition and travel habits [29]. Therefore, the space P modeling method was used to construct the geographical railway network in 2008, 2010, 2015, and 2019. The construction process is shown in Figure 2. First, according to the names of the stations in the railway timetable, the actual geographical coordinates of the train stations and the cities and provinces to which they belong were obtained from Baidu Maps, Google Maps, and a world map. The selection of city nodes was based on 337 cities in 2019 in administrative regions at the prefecture level and above, mainly including 293 prefecture-level cities, 4 municipalities directly under the central government and 40 states, leagues, and regions. Then, the space P method was used to model the undirected network based on the railway timetable from the four years. The final geographical railway network was obtained by integrating the nodes and edge weights according to the cities where the stations are located.

The geographical railway network can be viewed as an undirected graph $G = (V, E)$, where V is the set of city nodes, $V = \{1, 2, 3, \dots, i, \dots, n\}$, E is the set of edges between the stations, $E = \{\delta_{ij}\}$, and δ_{ij} indicates the connection relationship between node i and node j . When $\delta_{ij} = 0$, it means that node i and node j are not connected. Meanwhile, when $\delta_{ij} = w$, it means that the passing train number (i.e., weight) of nodes i and j is w .

3.2. Characteristics of the Railway Network Topology

The characteristics of the railway networks are usually described by network feature indices [17,30]. The commonly used feature indices include the degree, weighted degree, betweenness centrality, clustering coefficient, network density, and closeness centrality [31,32], among which the following feature indices were used for the complex network analysis in this paper.

1. Degree and Weighted Degree

The degree of node i is denoted by k_i , defined as the number of other nodes connected to node i , and the average value of the degrees of all the nodes in the network is called the average degree. In the railway network, the average degree denotes the average number of cities in the network that can be reached directly from any given city. Additionally, $\langle k \rangle$ denotes the weight degree, which represents the number of trains passing through city node i .

$$\langle k \rangle = \sum_i w_i k_i \quad (1)$$

2. Betweenness Centrality

Betweenness centrality is defined as the number of shortest paths between pairs of nodes that pass-through node i , denoted by B_i . It is used to estimate how a node contributes to the communication within the whole network, reflecting the transfer and linkage functions of the node in the network. In the railway network, the cities with numerous pairs of trains passing by and those as the ‘arteries’ are prone to have a high value of betweenness centrality, and such cities with a high betweenness centrality hold a high place in the operation of the railways, with control over and influence on the surrounding area.

$$B_i = \sum_{j,k \in v} \frac{n_{jk}(i)}{n_{jk}} \quad (2)$$

where n_{jk} represents the number of shortest paths between nodes j and k , and $n_{jk}(i)$ is the number of shortest paths between nodes j and k passing through node i .

3. Closeness Centrality

Closeness centrality, denoted by $C(i)$, is the inverse of the sum of the length of the shortest paths between node i and all other nodes. It demonstrates the ability of a city to choose from multiple paths when it reaches out to other cities and crystallizes the accessibility service level of a node in the network. The higher the closeness centrality, the higher the travel convenience of the node, and the more destination options in railway travel. It is calculated as follows:

$$C(i) = \frac{n-1}{\sum_{j \neq i} d_{ij}} \quad (3)$$

where d_{ij} is the length of the shortest paths between node i and node j , and n represents the number of all nodes.

3.3. Spatiotemporal Evolution of the Railway Network

The SOM method, as a competitive neural network method, can map complex and nonlinear high-dimensional data into a low-dimensional space with a simple geometric structure and interactive relationship. It employs unsupervised learning to train the input sample data and generate a low-dimensional (usually two-dimensional) representation

in a discrete form of the high-dimensional data in the input space. It is often used to explore the clustering patterns of spatial objects or as a visual data mining tool to map high-dimensional data to two-dimensional planes, so as to visually express and identify potential patterns [33–35]. The SOM model is composed of two layers of the network, including the input layer and the output layer, wherein the input layer is used to receive the input training samples, and the neurons of the output layer are generally arranged in a two-dimensional array. With the two-way connection between the two layers of neurons, the input set is classified by finding the optimal weight vector, that is, the best matching unit (BMU).

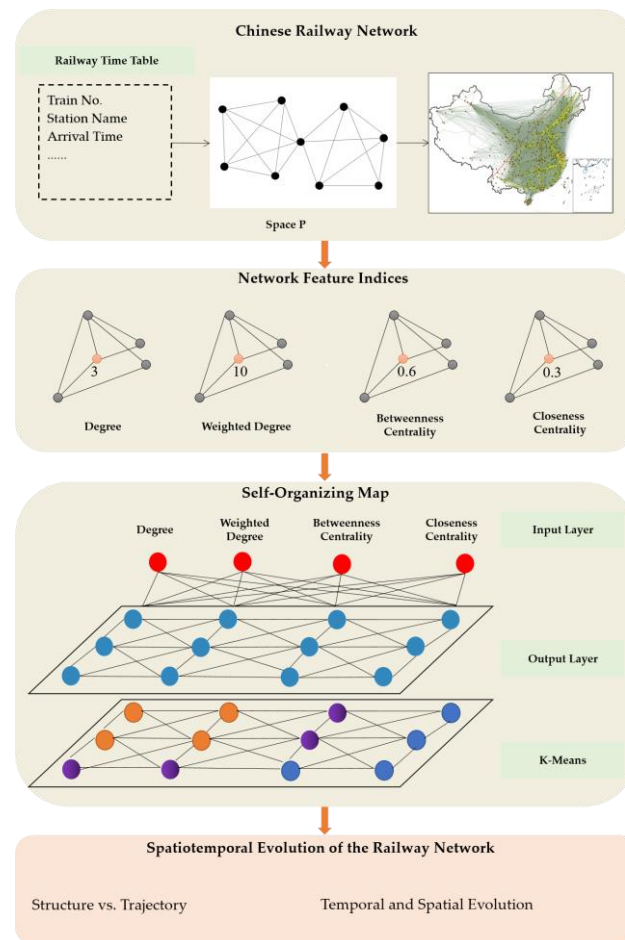


Figure 1. Overall research framework.

The SOM method classifies $y(s, t)$ data into a pre-specified number $L = m \times n$ of classes, where s and t represent spatial points and time points, respectively, for example, the network feature value of Shanghai City (s) in 2008 (t). Additionally, m and n are two positive integers whose values are close to each other. Each class is denoted by a neuron $z(k, t)$, $k = 1, 2, \dots, L$ types, placed on an $m \times n$ two-dimensional panel. In the first step, $y(s, t)$ is normalized to a value between $[0, 1]$, and $z(k, t)$ is assigned to a random number between $[0, 1]$. In the second step, the distance $d(s, k)$ is calculated as follows:

$$d(s, k) = \sum_i ||y(s, t) - z(k, t)|| \quad (4)$$

where $||\cdot||$ is the distance measure, and Σ represents the sum of the terms. In the third step, for a given s^* , the neuron k^* wins s^* . In the fourth step, for neuron k^* and its surroundings $z(k^*, t)_{\tau+1}$, the equation is shown as follows:

$$z(k^*, t)_{\tau+1} = z(k^*, t)_{\tau} + \delta[y(s^*, t) - z(k^*, t)_{\tau}] \quad (5)$$

where δ is a pre-given learning speed between (0, 1). In the fifth step, the remaining s all repeat the third step and iterate until $\forall s$. The sixth step goes back to the second step to start the next round of iterations, until τ_{max} .

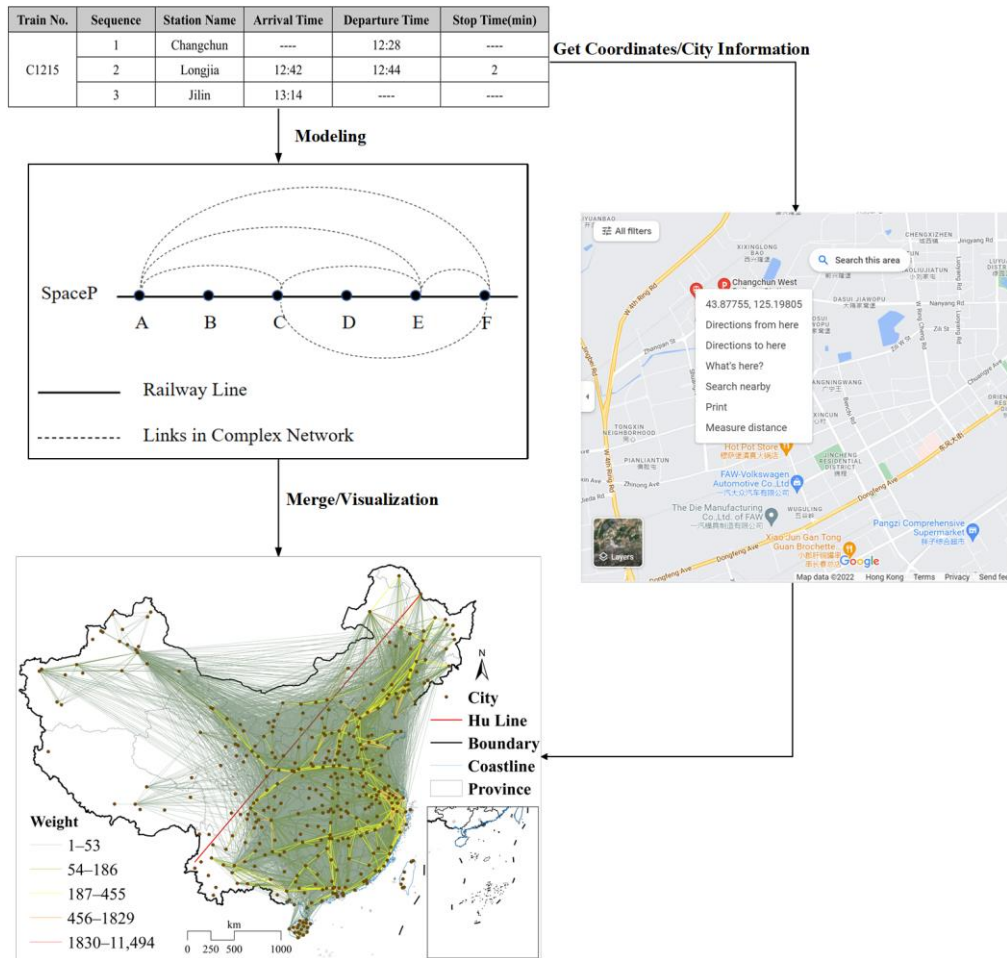


Figure 2. The construction process of the geographical railway network.

The input data in this paper are the network feature indices of 337 city nodes in four years, with a data volume of $4 \times 337 = 1348$ pieces of data. Each datum includes four network feature indices, namely, degree, weighted degree, closeness centrality, and betweenness centrality, as shown in Table 2. The normalization processing was performed before inputting the data into the SOM model. Due to the need to analyze the evolution pattern of the railway network and draw the change trajectory of each city node, the output panel size of the SOM model in this study was set to be large enough to ensure that each input node had a single winning neuron. Based on previous research experience [30,36], the size of the SOM output panel in this study was set to $100 \times 100 = 10,000$; that is, 10,000 output neurons. Due to simultaneously inputting and training the data from the four years, the clustering results showed a consistent temporal and spatial contrast. On this basis, the 10,000 neurons in the output panel were clustered using k-means based on their weight vectors, and the clustering results were visualized in the SOM panel and the geographical space. The optimal number of clusters was determined by considering both the silhouette coefficient method and the elbow method.

Table 2. Input data.

ID	City	Year	Degree	Weighted Degree	Closeness Centrality	Betweenness Centrality
1	Shanghai	2008	199	3876	0.75	0.0204
2	Shanghai	2010	206	5146	0.75	0.0199
3	Shanghai	2015	235	12,222	0.80	0.0158
4	Shanghai	2019	237	12,693	0.78	0.0101
...	...	...	...	...	...	...

4. Results

4.1. Analysis of the Feature Changes of the Railway Network Structure

Table 3 shows the values of the number of trains, the number of nodes, the number of edges, the degree, and the weighted degree of the geographical railway network in 2008, 2010, 2015, and 2019. The number of trains in operation increased from 2501 in 2008 to 8005 in 2019, and the number of nodes increased from 302 in 2008 to 331 in 2020. The number of edges also increased from 10,736 in 2008 to 17,533 in 2019, indicating that, as more and more trains are in operation, more and more cities are gaining access to railways, and the relationship between the nodes of the cities is becoming closer and closer. The degree increased from 71 in 2008 to 105 in 2019. The weighted degree increased from 1591 in 2008 to 2981 in 2020. Overall, from 2008 to 2019, the scale of the railway network expanded, and the network's structural feature indices showed an upward trend, which implies that cities are becoming increasingly connected, and people may find it more convenient to travel.

Table 3. The feature index values of the geographical railway network.

	2008	2010	2015	2019
Trains	2501	3782	5716	8005
Nodes	302	307	316	333
Edges	10,736	12,434	15,944	17,533
Degree	71	81	101	105
Weighted Degree	1591	2091	2728	2981

Figure 3 illustrates the betweenness centrality changes in different years. As time went by, the betweenness centrality value of each city declined. This indicates that, as more railways are constructed, the influence of each city in the railway network is gradually weakened. In 2008 and 2010, there were a few cities with high betweenness centrality values; these cities functioned as core hubs in the railway network, and their controllability over the railway network was strong. With the gradual perfection of the high-speed railway network, in 2015, the cities with high betweenness centrality values were increasingly dispersed, meaning that the distribution of railway network resources became increasingly reasonable. After the implementation of the “eight vertical and eight horizontal” high-speed railway framework in the Medium and Long Term Railway Network Plan (2016–2030), in 2019, the cities with relatively high betweenness centrality values tended to concentrate in the eastern and central areas, and southern areas of Hainan, indicating that, with the improvement of the high-speed railway network, the cities with strong controllability over the network were distributed in the central area of the national network.

Figure 4 exhibits the changes in the closeness centrality values in different years. The closeness centrality value of each year on a national scale showed a balanced distribution, where cities with relatively high closeness centrality values were scattered in all regions. In the early stage of the development of high-speed railways (2008–2010), the types of railways were dominated by the normal-speed railways, and the travel convenience embodied by the closeness centrality was mainly determined by the status of the city. The cities where the railway lines intersected had high travel convenience, such as Beijing, Zhengzhou, and Wuhan. In the era of the high-speed railway (2015–2019), the cities with high closeness

centrality values were mainly distributed in the central and eastern provincial capital cities showing a polarized pattern.

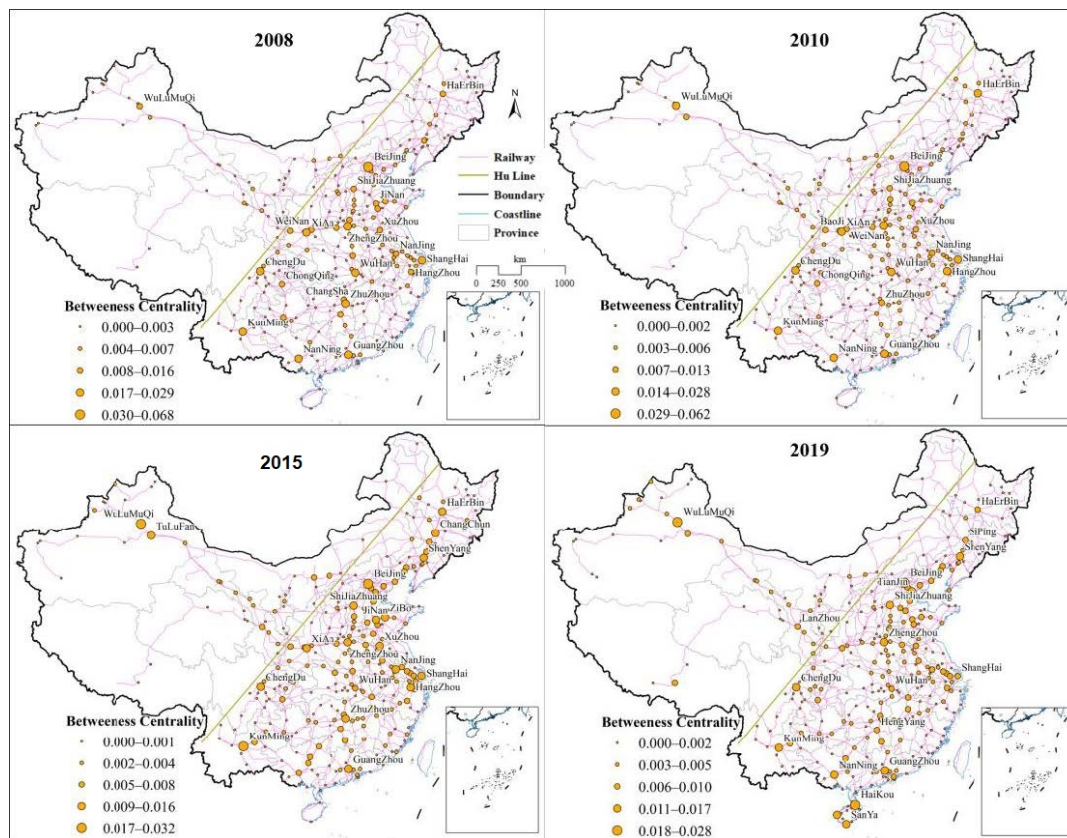


Figure 3. The maps of betweenness centrality values in different years.

4.2. Analysis of the Spatiotemporal Evolution of the Railway Network Structure

4.2.1. Output Panel of the Network Structure

The network feature indices of the railway network, such as the degree, weighted degree, betweenness centrality, and closeness centrality, represent the connection degree, traffic convenience, geometric center, and hub center of the nodes, respectively. Each unit in the panel had a weight vector corresponding to the four network feature indices after training. Therefore, the corresponding network feature index value of each panel unit can be calculated through the weight vector and displayed in a visual manner. Each network feature index was divided into five categories by natural breaks, as shown in Figure 5. These distributions carry no geographical meaning but can help us to understand the position of feature indices indicators in the output panel. The red area indicates a high concentration, while the green area suggests low aggregation. The spatial distribution patterns of the four network feature indices are identified and expressed in the SOM output panel. The high-concentration area of the degree and closeness centrality gradually clustered towards the upper right corner area from the bottom area of the output panel. The clustering patterns of the degree and closeness centrality shared certain similarities, which indicates that they were correlated to a certain extent. The high-concentration area of the weighted degree, originally located in the lower right corner of the output panel, gradually clustered towards the upper left corner. The high-concentration area of the betweenness centrality was relatively small, mainly located on the right side of the bottom, and gradually clustered towards the upper left corner.

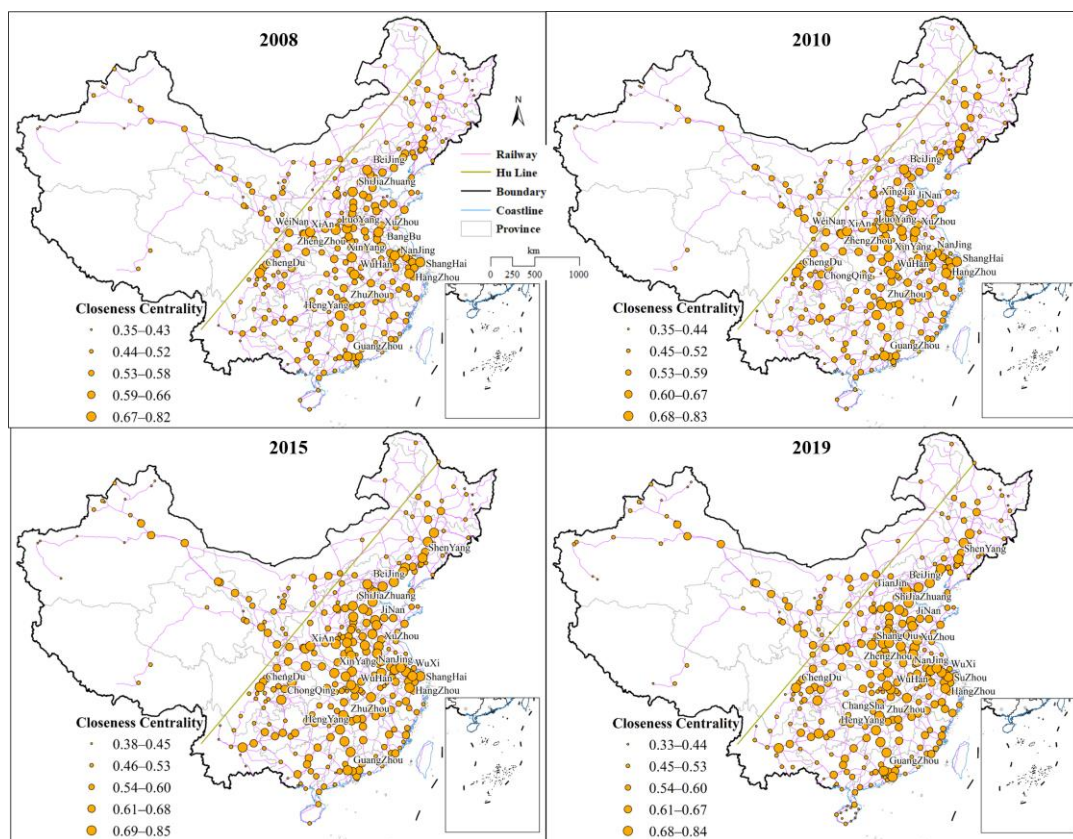


Figure 4. The maps of closeness centrality values in different years.

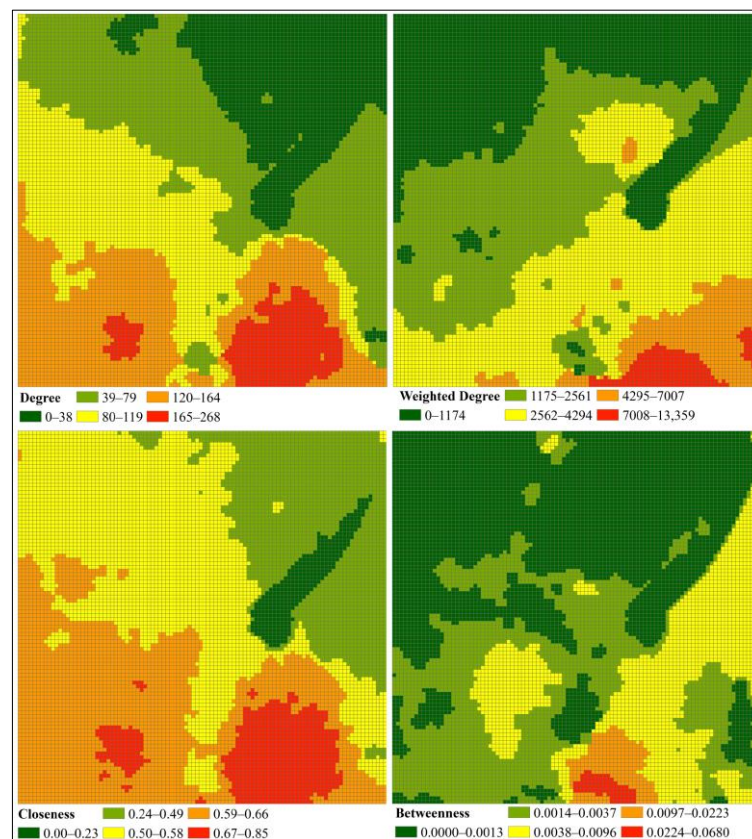


Figure 5. Output panel of the network structure.

4.2.2. Secondary Clustering

After training the SOM network, the k-means algorithm was used to perform a secondary clustering of the 10,000 weight vectors in the output panel. Then, the entire panel was divided into different types based on the clustering results, and the type of each city node was obtained from the type of the output neuron where the city node is located. By using the silhouette coefficient and elbow methods, it was found that when the number of categories was seven, the clustering effect was sound. The seven types of clusters are represented by different colors in Figure 6a. To facilitate the determination of the meaning of the seven types of clusters, a line graph of five hierarchical regions of the four network feature indices and the seven corresponding types of clusters was drawn, as shown in Figure 6b. As the four network feature indices reflect the importance of the nodes in the railway network from different aspects, the seven types of clusters were further divided into four classes of nodes, used to reflect the importance of the nodes, that is, the hub nodes (types I and II in the output panel), important nodes (type III), general nodes (types IV and V), and periphery nodes (types VI and VII).

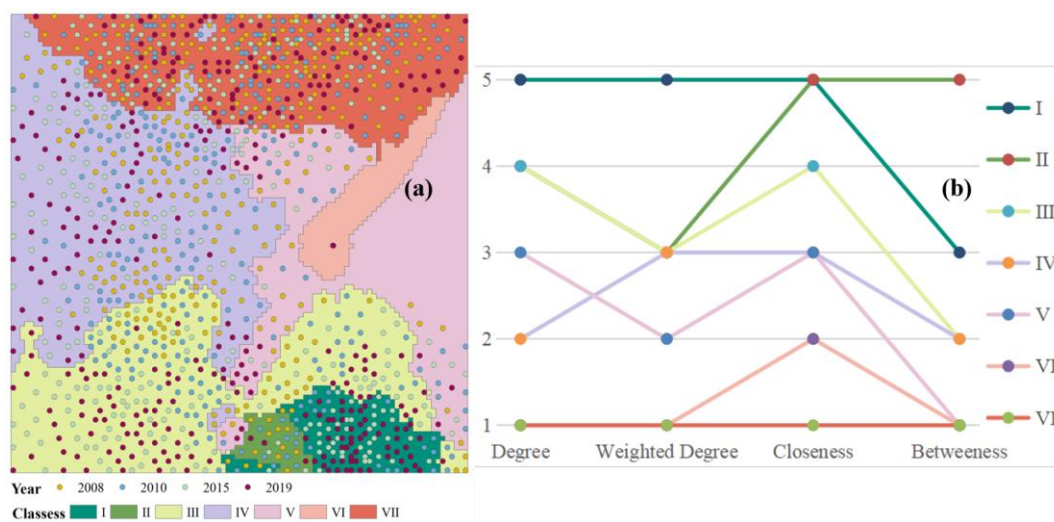


Figure 6. Results of the secondary clustering: (a) 7 types of clusters; (b) network feature types of the 7 clustering results.

4.2.3. Evolution Analysis of the Hierarchical Structure of City Nodes

Firstly, by calculating the Euclidean distance between the input nodes and all output neurons, the output neuron that lies closest to the input node was selected as the winning neuron. Then, according to the positions of the winning neurons in the SOM output panel, we could identify which hierarchical classification each city node belongs to in different years, and the city nodes of different years were mapped to the corresponding positions in the panel with different colors. Finally, the hierarchical classification type of each city node was visualized in the geographical space, and the changes in the network features of each city node in different years were intuitively exhibited and compared.

The clustering results of the city nodes in different years are shown in Figure 7. In 2008, due to the small number and short length of railway line facilities, low network density and loose connections, the railway lines showed a typical core-peripheral pattern in space. The provincial capital cities, and the cities with a large population and high levels of economic development, served as the hub nodes in the railway network, mainly including Beijing, Shanghai, Wuhan, and Harbin, with a total number of 15. They were connected with important nodes and general nodes, but the degree of connection was low. There were 48 important nodes, which were mainly distributed along the railway lines of Beijing–Harbin, Beijing–Guangzhou, and Beijing–Shanghai, all of which are located east of the Hu Line (which is an important national dividing line of China’s population density

and socio-economic development level), serving only a few areas. The nodes to the west of the Hu Line were general nodes and periphery nodes. The development of the east and west was extremely unbalanced. In this period, the construction of high-speed railways had just been initiated, the railway network was dominated by normal-speed railways, and the importance of most city nodes was relatively low.

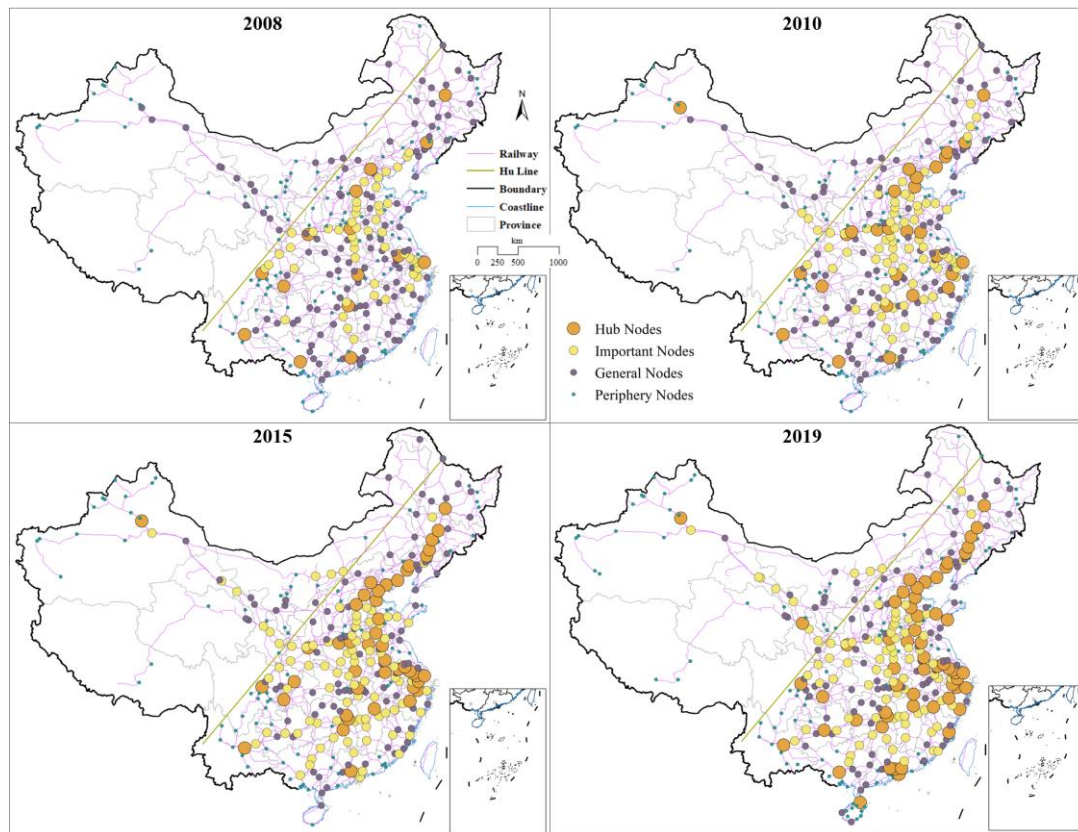


Figure 7. Clustering of the city nodes in different years.

In 2010, the railway lines presented an axial belt expansion pattern in space, showing a salient belt distribution based on the high-speed railway lines. The start and end nodes of the connective belts were the hub nodes and included high-population cities, economically developed cities, provincial capital cities, and national transportation hubs, with a total of 28. Urumqi City, located to the west of the Hu Line, became a new hub node. These nodes had a greater demand for external connections. With the connection and extension of more railway lines, such as the opening and operation of the Wuhan–Guangzhou section of the Beijing–Guangzhou high-speed railway, more cities along the railway line were incorporated into the network. Meanwhile, the number of nodes in the network grew rapidly, and connections between the nodes were established, but the degree of connection was relatively low. At this time, the stations, lines, and other facilities in the railway network were relatively stable, and the main changes were characterized by stronger connection.

In 2015, the railway lines displayed a pattern of ‘strings of beads’ in space, in which the expansion of the belts was still based on the high-speed railway lines, yet the connection became increasingly stronger. The status of hub nodes in the railway network continued to strengthen. The number of hub nodes increased to 45, and the new hub nodes were mainly distributed along the Beijing–Shanghai high-speed railway and Beijing–Harbin high-speed railway lines. There were many hub nodes in the Yangtze River Delta. The high-speed railway lines connected in series were no longer a single node, but a ‘domain’ formed centering on the hub nodes and important nodes. The spatial pattern of the railway network was a ‘core–core’ connection. With the opening of more regional and

intercity lines, the connection between important nodes, general nodes and periphery nodes was strengthened, and the status and influence of important nodes in the network were enhanced. In particular, the nodes at the intersection of main railway lines reached more nodes and played an increasingly important role in the network.

In 2019, the railway lines presented a multi-center network pattern in space, and the network connectivity and connection strength were greatly improved. The influence and status of the original important nodes and general nodes in the network were continuously improved, the whole network was featured with a multi-center pattern, and the development of the railway network was increasingly balanced. The number of hub nodes increased to 59, and the new nodes were mainly located in the southeast coastal, southwest and southern regions, especially in the Pearl River Delta region. Through the construction of the new railway lines and the extension of the main lines, China completed the construction of the “four horizontal and four vertical” fast track and built a high-density, high-speed railway network in the Yangtze River Delta, the Pearl River Delta, and the Bohai Sea. The high-speed railway network was promoted from the areas with a large population and high level of economic development to the remote areas, thereby connecting all major and medium-scale cities, and completing high-speed railway connectivity among the eastern, central, western, and northeastern regions of China. However, due to the influence of the regional economic development level, topography, population density, and politics, the network density and connection intensity varied greatly in the spatial distribution.

4.2.4. Evolution Trajectory of the Network Structure

The change trajectory of each node was obtained by connecting the positions of the nodes in the railway network each year in the SOM output panel according to the time sequence. At the same time, the change trend of each node was obtained by combining the different clustering categories in the SOM output panel. On this basis, the k-means algorithm was used to cluster the change trajectory of each node, which was divided into five types. Figure 8 shows the clustering results of the node change trajectories in the SOM output panel, where the blue and red dots represent the start and end points of each change trajectory, respectively, that is, the positions of the winning neurons at each node in the railway network in 2008 and 2019.

Figure 9 shows the clustering categories of each node’s change trajectory in the actual geographical space, and the correspondence between the cluster types and the colors in Figure 8 also applies to Figure 9. By combining the clustering results of the node change trajectory, the trajectory direction, and the clustering type in the SOM output panel, we found the following.

Trajectory 1 is composed of nodes with very high network feature index values, which will continue to increase over time. These nodes played a decisive role in the early stage of the construction of the railway network, and occupied a dominant position, mainly including the municipalities directly under the central government, provincial capital cities, and sub-central cities, such as Beijing, Shanghai, Guangzhou, Zhengzhou, Harbin, Wuhan, Chengdu, Nanjing, and Hangzhou. These cities have a large number of connections with other cities, serving as the core nodes between the main cities in the north and south, and forming the skeleton of the railway network. Meanwhile, they are also the agglomeration centers of the national economy and are the cities with the largest scale in the national urban system. With the construction of the railway network, the connectivity between these nodes and the surrounding cities has been continuously enhanced, and their central role in the network is still improving.

Trajectory 2 consists of nodes that are located at important geographical positions and gradually evolve into hub areas, and the importance of the nodes is greatly improved. These nodes were not hub nodes in the early stage of the construction of the railway network, but due to the expansion of the railway network and the deviation of the railway construction center, these nodes have gradually become the connection center of the region, with stronger connectivity in the network. These nodes are mainly located to the east of the

Hu Line and are important nodes in the railway network connecting the south and north areas, as well as the west and east areas, such as Shenyang, Tianjin, Changsha, Chongqing, Nanchang, Changchun, Jinan, Xuzhou, and Suzhou. These cities directly link with most of the prefecture-level cities in the country, mostly the relatively developed provincial capital cities.

Trajectory 3 shows the nodes changing from general nodes to important nodes, most of which are located along main railway lines. These nodes have many connections with other nodes and play an important role in the network, such as Xiamen, Liuzhou, Guiyang, Huaihua, Guang'an, Suining, and Taiyuan. These cities maintain close contact with prefecture-level cities with the same sphere of influence and have close direct contact with major cities outside the sphere of influence.

Trajectory 4 presents the nodes that remain as general nodes, mainly regional cities, with very limited influence, such as Baise, Beihai, Ziyang, Haikou, Chengde, Jiaozuo, Jingdezhen, and Jingzhou. These nodes maintain close and direct contact with important cities within the first-level sphere of influence of the national railway hub but have less contact with cities outside the sphere of influence.

Trajectory 5 mainly illustrates the periphery nodes that had no trains opened at the beginning of the construction of the railway network, and generally had low values of the network structure feature indices, even after the construction was completed. With a few trains passing through these nodes, they lack contact with most nodes in the network. These nodes are mainly located in more remote areas, with relatively underdeveloped economic development, such as Karamay, Lijiang, Wuchang, Danzhou, Bazhong, and Xiantao.

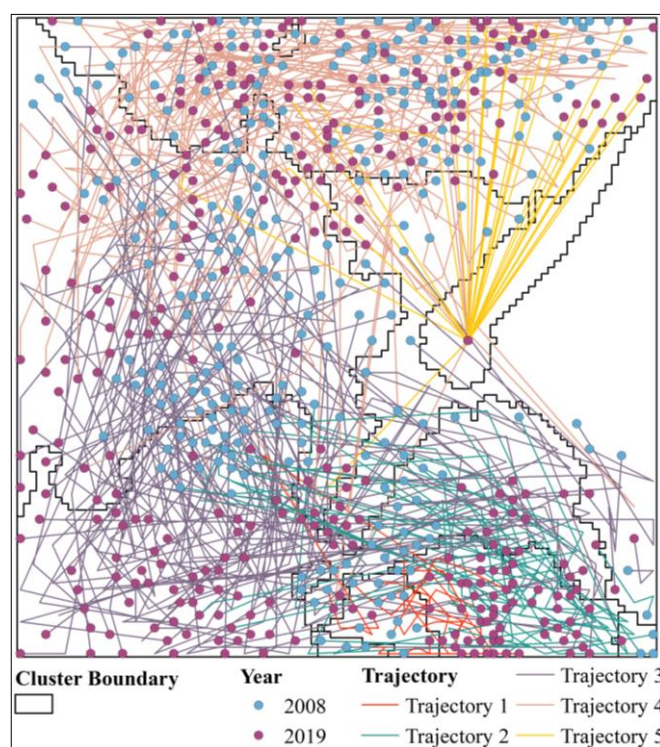


Figure 8. Clustering results of node change trajectory.

Overall, how the change trajectory of the nodes has evolved is highly related to the scale and hierarchical structure of the urban system, and the position of a city in the railway network reflects its level in the national urban system to a certain extent. Theoretically, the larger the population of the city, the more likely it is to be connected to other cities, and the stronger the railway service capacity. The urban scale structure can be clearly reflected in the railway network structure, and the spatial pattern of the railway network represents the characteristics of the spatial structure of the urban system to a certain extent.

In the future, taking both efficiency and fairness into consideration, more efforts should be made to build a railway network with a reasonable layout, wide coverage, high efficiency, and convenience, with an emphasis on both economic and social benefits according to the pattern of its evolution.

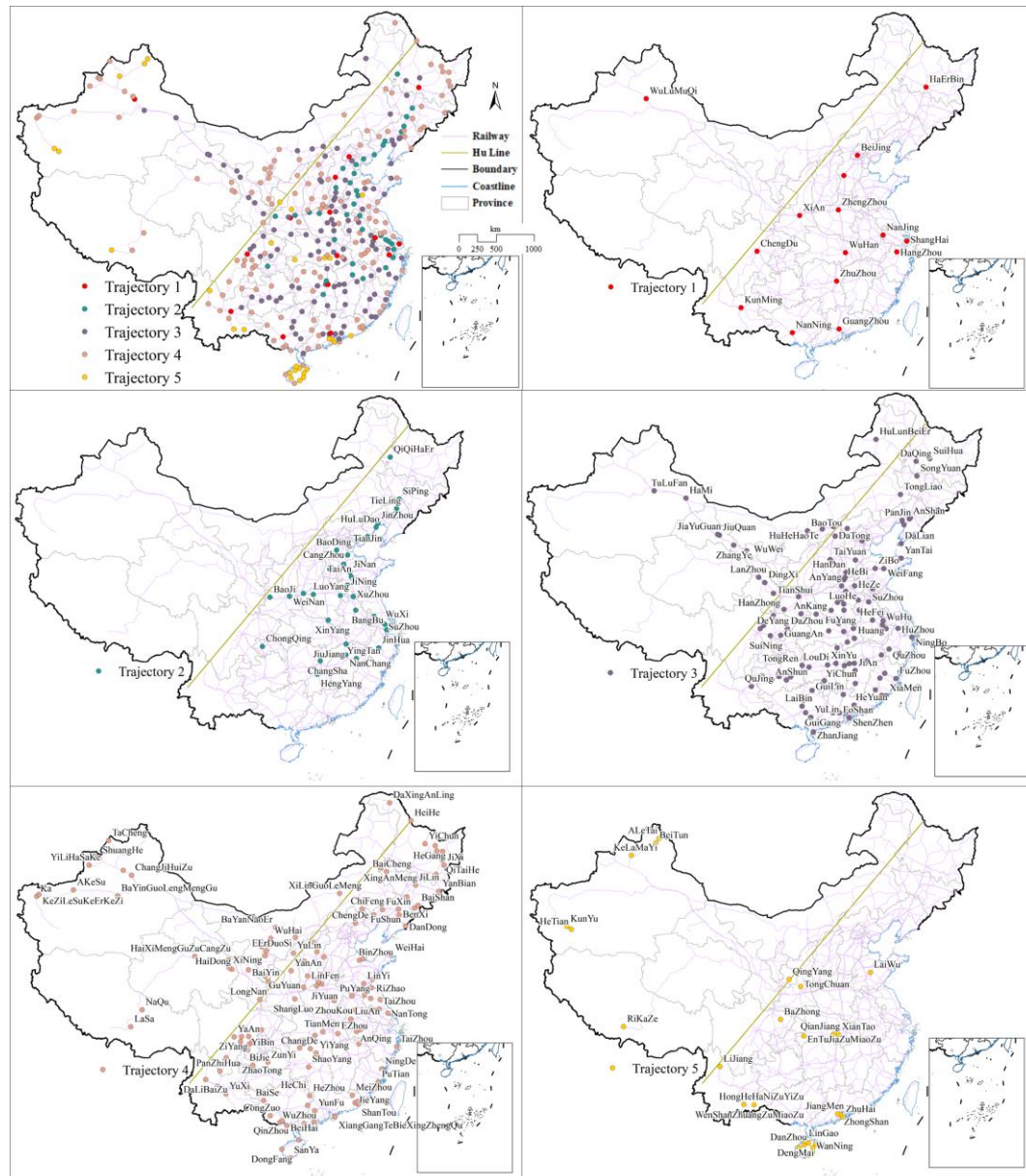


Figure 9. Clustering categories of the city nodes' change trajectories.

5. Discussion

In the past decade, China has entered the heyday of railway construction, especially high-speed railway construction, gradually forming a railway network pattern with wide coverage, rich layers, and a complex structure. Complex networks can describe the structure and evolution characteristics of railway networks from both global and local perspectives. This paper conducted complex network modeling and analysis of railway network characteristics in four years, and found that both the network size and structural characteristics showed an upward trend, basically reflecting the development of China's railways from 2008 to 2019. The conclusions of this study are consistent with the findings of other papers [19,20]. Nonetheless, this paper also has limitations. Due to the difficulty in obtaining

train operation data, only the data from 2008, 2010, 2015, and 2019 were selected to analyze the spatiotemporal evolution of the railway network structure. Future studies can consider incorporating more years of train operation data and conducting a time-series railway network analysis on its spatiotemporal evolution.

In this paper, the SOM algorithm was used to integrate the railway network feature indices of 4 years into a unified framework, and we analyzed the evolution patterns of the entire railway network. Due to the use of the dimension reduction and visualization functions of the SOM algorithm, the number of neurons in the output panel was much larger than the input data, and the k-means method was used to further cluster the output panel of the SOM algorithm. However, current studies have not provided a definitive conclusion on how much larger the size of the output panel lane should be, so the output panel size used in this study was mainly based on examples found in the literature [30,36]. Although the evolution pattern of China's railway network is well recognized, there is still much work to be conducted. In the future, we will explore the driving factors for the evolution of the spatial pattern of railway networks and will analyze the interactive relationship between railway networks and socio-economic systems. Moreover, different study objects have different time and space scales [37]. In this paper, the evolution pattern of the railway network was analyzed at the city scale. Whether the evolution pattern at the county level or the station level is different from that at city level still requires further study.

6. Conclusions

In view of the existing problems in previous studies of modeling railway data in different periods separately, only comparing the network feature indices of two or more adjacent periods for the evolution analysis, and failing to integrate the railway network from different periods into a unified framework, this study integrated railway networks in different periods based on complex network theory and self-organizing maps method, and analyzed the evolution pattern of the railway network. Firstly, based on the complex network theory, this study constructed the geographical railway network in 2008, 2010, 2015, and 2019, and analyzed the changes in network feature indices. From 2008 to 2019, with the increasing scale of the railway network, the network feature indices showed an upward trend, indicating that the connections between cities become closer and closer, and that travel and transfer became more and more convenient. Secondly, the multi-time series data of railway network features were simultaneously input into the SOM network for training and output, so as to probe into the spatiotemporal evolution pattern of the railway network structure. The expansion pattern of the railway network from 2008 to 2019 was divided into the core-peripheral pattern, belt expansion pattern, strings of beads pattern, and multi-center network pattern. Finally, the change trajectory of each city node was compared using cluster analysis. It was found that the new nodes were basically peripheral nodes. As cities connected more closely, the structure of the railway network changed accordingly. The evolution of the change trajectory of the nodes was highly related to the scale and hierarchical structure of the urban system, and the position of a city in the railway network reflected its level in the national urban system to a certain extent.

China has built a railway network with wide coverage, rich levels, and a complex structure, achieving good connectivity between various levels, and serving the rapid passenger and freight flows between regions. At a stage when the layout of the railway network tends to be stable, improving operational services will be a necessity and prerequisite for the effectiveness of China's railway network in the future. Moreover, in March 2021, China, for the first time, issued a policy setting restrictive requirements for the construction of high-speed railways, banning the construction of subways and light railways in disguised forms, and clarified the construction threshold for high-speed rail projects with a speed of 350 km per hour. It is expected that the operating mileage of high-speed rail in China will maintain a growth trend, but the mileage of new developments will be reduced. Therefore, subsequent development should focus on building a railway network with a

reasonable layout, wide coverage, high efficiency, and convenience based on the law of railway development, considering economic and social benefits, and efficiency and fairness.

Author Contributions: Lingzhi Yin performed the experiments and wrote this manuscript; Yafei Wang provided important suggestions. All authors have read and agreed to the published version of the manuscript.

Funding: This paper was supported by Zhejiang Provincial Natural Science Foundation of China under Grant No. LTGG23D010002; the National Natural Science Foundation of China under Grant No. 41801297, and the Science Foundation of Zhejiang Sci-Tech University (ZSTU) under Grant No. 20032312-Y.

Data Availability Statement: Not applicable.

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

References

- Macchi, M.; Garetti, M.; Centrone, D.; Fumagalli, L.; Pavirani, G.P. Maintenance management of railway infrastructures based on reliability analysis. *Reliab. Eng. Syst. Safe* **2012**, *104*, 71–83. [\[CrossRef\]](#)
- Hong, L.; Ouyang, M.; Peeta, S.; He, X.; Yan, Y. Vulnerability assessment and mitigation for the Chinese railway system under floods. *Reliab. Eng. Syst. Safe* **2015**, *137*, 58–68. [\[CrossRef\]](#)
- Lam, C.Y.; Tai, K. Network topological approach to modeling accident causations and characteristics: Analysis of railway incidents in Japan. *Reliab. Eng. Syst. Safe* **2020**, *193*, 106626. [\[CrossRef\]](#)
- Yin, L.; Zhu, J.; Li, W.; Wang, J. Vulnerability Analysis of Geographical Railway Network under Geological Hazard in China. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 342. [\[CrossRef\]](#)
- Ouyang, M.; Zhao, L.; Hong, L.; Pan, Z. Comparisons of complex network based models and real train flow model to analyze Chinese railway vulnerability. *Reliab. Eng. Syst. Safe* **2014**, *123*, 38–46. [\[CrossRef\]](#)
- Szymula, C.; Bešinović, N. Passenger-centered vulnerability assessment of railway networks. *Transp. Res. B-Meth.* **2020**, *136*, 30–61. [\[CrossRef\]](#)
- Yin, L.Z.; Wang, Y.F. Network characteristics and vulnerability analysis of Chinese railway network under earthquake disasters. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 697. [\[CrossRef\]](#)
- Zhang, J.; Hu, F.; Wang, S.; Dai, Y.; Wang, Y. Structural vulnerability and intervention of high-speed railway networks. *Physica A* **2016**, *462*, 743–751. [\[CrossRef\]](#)
- Wang, W.; Cai, K.; Du, W.; Wu, X.; Tong, L.C.; Zhu, X.; Cao, X. Analysis of the Chinese railway system as a complex network. *Chaos Soliton Fract.* **2020**, *130*, 109408. [\[CrossRef\]](#)
- Strogatz, S.H. Exploring complex networks. *Nature* **2001**, *410*, 268–276. [\[CrossRef\]](#)
- Kreutzberger, E.D. Distance and time in intermodal goods transport networks in Europe: A generic approach. *Transp. Res. A-Pol.* **2008**, *42*, 973–993. [\[CrossRef\]](#)
- Ghosh, S.; Banerjee, A.; Sharma, N.; Agarwal, S.; Ganguly, N.; Bhattacharya, S.; Mukherjee, A. Statistical analysis of the Indian railway network: A complex network approach. *Acta Phys. Pol. B Proc. Suppl.* **2011**, *4*, 123–138. [\[CrossRef\]](#)
- Calzada-Infante, L.; Adenso-Díaz, B.; Carbajal, S.G. Analysis of the European international railway network and passenger transfers. *Chaos Soliton Fract.* **2020**, *141*, 110357. [\[CrossRef\]](#)
- Sen, P.; Dasgupta, S.; Chatterjee, A.; Sreeram, P.A.; Mukherjee, G.; Manna, S.S. Small-world properties of the Indian railway network. *Phys. Rev. E* **2003**, *67*, 036106. [\[CrossRef\]](#) [\[PubMed\]](#)
- Albert, R.; Barabási, A.L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **2002**, *74*, 47. [\[CrossRef\]](#)
- Caset, F.; Vale, D.S.; Viana, C.M. Measuring the accessibility of railway stations in the Brussels Regional Express Network: A node-place modeling approach. *Netw. Spat. Econ.* **2018**, *18*, 495–530. [\[CrossRef\]](#)
- Cao, W.W.; Feng, X.N.; Zhang, H. The structural and spatial properties of the high-speed railway network in China: A complex network perspective. *J. Rail Transp. Plan. Manag.* **2019**, *9*, 46–56. [\[CrossRef\]](#)
- Wang, J.; Jin, F.; Mo, H.; Wang, F. Spatiotemporal evolution of China's railway network in the 20th century: An accessibility approach. *Acta Geogr. Sin.* **2009**, *43*, 765–778. [\[CrossRef\]](#)
- Huang, Y.; Lu, S.; Yang, X.; Zhao, Z. Exploring Railway Network Dynamics in China from 2008 to 2017. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 320. [\[CrossRef\]](#)
- Lu, S.; Huang, Y.; Zhao, Z.; Yang, X. Exploring the Hierarchical Structure of China's Railway Network from 2008 to 2017. *Sustainability* **2018**, *10*, 3173. [\[CrossRef\]](#)
- Wei, S.; Jiao, J.; Wang, L.; Xu, J. Evolving characteristics of high-speed railway network structure in Yangtze river delta, China: The perspective of passenger flows. *Appl. Spat. Anal. Policy* **2020**, *13*, 925–943. [\[CrossRef\]](#)
- He, D.; Chen, Z.; Pei, T.; Zhou, J. The regional and local scale evolution of the spatial structure of high-speed railway networks—A case study focused on Beijing-Tianjin-Hebei Urban Agglomeration. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 543. [\[CrossRef\]](#)

23. Ghaseminezhad, M.H.; Karami, A. A novel self-organizing map (SOM) neural network for discrete groups of data clustering. *Appl. Soft Comput.* **2011**, *11*, 3771–3778. [[CrossRef](#)]
24. Ren, J.H.; Chen, J.C.; Wang, N. Visual analysis of SOM network in fault diagnosis. *Phys. Procedia* **2011**, *22*, 333–338.
25. Lee, C.D.; Rinner, C. Visualizing urban social change with Self-Organizing Maps: Toronto neighbourhoods, 1996–2006. *Habitat. Int.* **2015**, *45*, 92–98. [[CrossRef](#)]
26. Augustijn, E.W.; Zurita-Milla, R. Self-organizing maps as an approach to exploring spatiotemporal diffusion patterns. *Int. J. Health Geogr.* **2013**, *12*, 60. [[CrossRef](#)] [[PubMed](#)]
27. Andrienko, G.; Andrienko, N.; Bremm, S.; Schreck, T.; Von Landesberger, T.; Bak, P.; Keim, D. Space-in-Time and Time-in-Space Self-Organizing Maps for Exploring Spatiotemporal Patterns. *Comput. Graph Forum.* **2010**, *29*, 913–922. [[CrossRef](#)]
28. Qi, J.; Liu, H.; Liu, X.; Zhang, Y. Spatiotemporal evolution analysis of time-series land use change using self-organizing map to examine the zoning and scale effects. *Comput. Environ. Urban* **2019**, *76*, 11–23. [[CrossRef](#)]
29. Jiao, J.; Wang, J.; Jin, F. Impacts of high-speed rail lines on the city network in China. *J. Transp. Geogr.* **2017**, *60*, 257–266. [[CrossRef](#)]
30. Soh, H.; Lim, S.; Zhang, T.; Fu, X.; Lee, G.K.K.; Hung, T.G.G.; Di, P.; Prakasam, S.; Wong, L. Weighted complex network analysis of travel routes on the Singapore public transportation system. *Physica A* **2010**, *389*, 5852–5863. [[CrossRef](#)]
31. Lin, Y.J.; Ban, Y.F. Complex network topology of transportation systems. *Transp. Rev.* **2013**, *33*, 658–685. [[CrossRef](#)]
32. Xu, W.T.; Zhou, J.P.; Qiu, G. China's high-speed rail network construction and planning over time: A network analysis. *J. Transp. Geogr.* **2018**, *70*, 40–54. [[CrossRef](#)]
33. Chen, N.; Ribeiro, B.; Vieira, A.; Chen, A. Clustering and visualization of bankruptcy trajectory using self-organizing map. *Expert Syst. Appl.* **2013**, *40*, 385–393. [[CrossRef](#)]
34. Vesanto, J.; Alhoniemi, E. Clustering of the self-organizing map. *IEEE Trans. Neural Netw.* **2000**, *11*, 586–600. [[CrossRef](#)] [[PubMed](#)]
35. Kohonen, T. Self-organized formation of topologically correct feature maps. *Biol. Cybern.* **1982**, *43*, 59–69. [[CrossRef](#)]
36. Skupin, A.; Hagelman, R. Visualizing Demographic Trajectories with Self-Organizing Maps. *Geoinformatica* **2005**, *9*, 159–179. [[CrossRef](#)]
37. Li, Z.; Wang, J.; Tan, S.; Xu, Z. Scale in geo-information science: An overview of thirty-year development. *Geomat. Inf. Sci. Wuhan Univ.* **2018**, *43*, 2233–2242.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.