A study of region division based on spatial units fusion of clustering algorithm

Yu Wang*

School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 610031, Sichuan, China
School of Traffic and Transportation Engineering, Dalian Jiaotong University, Dalian 116028, Liaoning, China

Received 1 March 2014, www.cmnt.lv

Abstract

Region division is the foundation of regional socio-economic development planning. Traditional regional division is only built upon the information of attributes in each spatial unit. However, the spatial relationships and their spatial interaction between units are ignored. In the study, based on the spatial linkage theory, the principle and steps of spatial units fusion of clustering algorithm is proposed. According the spatial range of Chinese railway as a case, determine the spatial units and the linkage model was built based on the proximity and attribute characters, the experiments of regional division for Chinese railway are achieved. The experiment shows the results are highly accordant with the real situations and have proved the feasibility of the algorithm.

Keywords: spatial linkage theory, region division, spatial units fusion of clustering algorithm, Chinese railway

1 Introduction

Regional planning is an important part of regional socio-economic development planning. Region division is the foundation of regional planning, and the division foundation in current research is classified into two types. One take the economic attribute as the main basis [1,2], as economic belt [3], economic region [4] and metropolitan area [5]. But the academe has not come to an agreement with it, the reason is the principles and methods are different and can also affect the regional planning establishment and implementation. Another take the spatial units attribute as the main basis [6]. But this method only consider the individual attribute data and the relationship are ignored in practice.

Based on the spatial linkage theory [7], starting with the relationship with the contiguous spatial units, the fusion of clustering algorithm [8,9] were studied and so as to realize the region division.

2 Spatial units fusion of clustering algorithm

2.1 ALGORITHM PRINCIPLE

Within the selected spatial range, the spatial set can be formed under the proper spatial units are selected, by computing the units correlation data, the correlation coefficient set is determined, fused with the two adjacent spatial units which has maximum correlation coefficient, connect the units in one direction according to the units attribute until there are not any external connection by the fusion units, so the units which have fused cluster a class. According to this algorithm, all the spatial units are clustered several classes in the spatial range.

2.2 ALGORITHM STEPS

Step 1: Determination of the spatial range s.
Step 2: Determination of the spatial units s_j:
\[ S = \{ s_1, s_2, ..., s_i \} \]

Step 3: Determination of the model of the relation coefficient f.

Step 4: Take a spatial unit as the starting point S_j, computed the relation coefficient f_{ij} with the neighboring spatial units, fused with the two neighboring spatial units which has maximum relation coefficient max f_{ij}.

Step 5: Determination of the neighboring spatial units in one direction connection T_{ij}. The units attribute data is used to represent the (Q_{1}, Q_{2}), determine the connect direction T_{ij} by comparing Q_{1} and Q_{2}. if \( Q_{1} / Q_{2} > 1 \), \( T_{ij} = 1 \), otherwise \( T_{ij} = 0 \).

Thus, the two contiguous units have only one direction connection.

Step 6: In the same way, by computing as step4 and step5 until there are not any external connection by the fusion units, so the units which have fused cluster the first class \( C_1 \).

Step 7: According to this algorithm, all the spatial units are clustered several classes, \( S = ( C_1, C_2, ..., C_k ) \).

* Corresponding author’s e-mail: wangyudjtlu163.com
3.2 DETERMINE THE SPATIAL UNITS

Chinese railway transport mainly presented as inter-provincial transport relationship, and the volume of freight flow mostly completed between 1,000~2,000 km [11]. So the spatial units selection are based on the provinces, municipalities and autonomous regions. Considering the uniqueness of the adjacent space, fused the only adjacent units to each other at first and the number of spatial units are 28. As shown in Table 1.

### 3.2 SPATIAL UNITS CORRELATION MODEL

Spatial units correlation model was built from the method of spatial transport linkage theory [7], according to the data of \( O - D \) matrix and considering the bidirectional transport data transfer. The model can be used to evaluate the correlation strength with the adjacent spatial units, as shown by Equation (1):

\[
f_{ij} = a_i \frac{O_{ij}}{O_i} + a_j \frac{D_{ij}}{D_j} + a_u \frac{O_u}{O_u} + a_v \frac{D_v}{D_v}, (i, j = 1, 2, ..., n),
\]

where \( f_{ij} \) is the correlation coefficient, \( O_{ij} \) is the freight flow from unit \( i \) to \( j \), \( D_{ij} \) is the freight flow from unit \( j \) to \( i \), \( O_i \) and \( O_j \) is the total output freight flow with unit \( i \) and \( j \), \( D_i \) and \( D_j \) is the total input freight flow with unit \( i \) and \( j \), \( a_i \) and \( a_j \) is the standardization coefficient, \( n \) is the number of spatial units.

Regional division of Chinese railway has based on the analysis on the freight flow in china. \( a_i \) and \( a_j \) were obtained from “China Statistical Yearbook for Regional Economy-2013” [12]. \( O_{ij}, D_{ij}, O_i, D_i \) were obtained from “China Transportation Yearbook 2013”[13].

### 3.3 CALCULATION PROCESS

Take \( S_1 \) as the starting point, search for the adjacent spatial units as \( S_2, S_3, S_4, S_10, S_{11}, S_{12}, S_{13}, S_{14}, S_{15}, S_{16}, S_{17}, S_{18} \) and calculate the \( f_{ij} \). As shown in Table 2.

<table>
<thead>
<tr>
<th>Adjacent units</th>
<th>a1</th>
<th>O1/Oi</th>
<th>D1/Dj</th>
<th>a2</th>
<th>O2/Oj</th>
<th>D2/Dj</th>
<th>fij</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1 - S2</td>
<td>0.96</td>
<td>0.03</td>
<td>0.53</td>
<td>0.88</td>
<td>0.49</td>
<td>0.05</td>
<td>1.00</td>
</tr>
<tr>
<td>S1 - S3</td>
<td>0.96</td>
<td>0.01</td>
<td>0.19</td>
<td>0.58</td>
<td>0.49</td>
<td>0.02</td>
<td>0.68</td>
</tr>
<tr>
<td>S1 - S4</td>
<td>0.96</td>
<td>0.01</td>
<td>0.01</td>
<td>0.88</td>
<td>0.49</td>
<td>0.02</td>
<td>0.68</td>
</tr>
<tr>
<td>S1 - S14</td>
<td>0.96</td>
<td>0.07</td>
<td>0.04</td>
<td>1.37</td>
<td>0.49</td>
<td>0.02</td>
<td>0.68</td>
</tr>
<tr>
<td>S1 - S15</td>
<td>0.96</td>
<td>0.02</td>
<td>0.01</td>
<td>0.81</td>
<td>0.49</td>
<td>0.02</td>
<td>0.68</td>
</tr>
</tbody>
</table>

From Table 2, we note that \( S_1 \) and \( S_2 \) have the maximum correlation coefficient, \( \max f_{ij} = 1.00 \). Determine \( S_1 \) and \( S_2 \) in one direction connection, the exchanges of freight volume can be regarded as \( (Q_1, Q_2) \), as \( Q_i = O_{ij} \), \( Q_j = D_{ij} \), \( O_{12} / D_{12} = 0.02 < 1 \), so the only direction connection between \( S_1 \) and \( S_2 \) was \( S_1 \rightarrow S_2 \).

Take \( S_2 \) as the starting point, search for the adjacent spatial units as \( S_3, S_4, S_{10}, S_{11}, S_{12}, S_{13}, S_{14}, S_{15}, S_{16}, S_{17}, S_{18} \) and calculate the \( f_{ij} \). As shown in Table 3.
From Table 3, we note that $S_2$ and $S_1$ have the maximum correlation coefficient, $\max f_{31} = 1.00$. Determine $S_1$ and $S_2$ in one direction connection, as $O_{31}/D_{31} = 64.3 > 1$, so the only direction connection between $S_2$ and $S_1$ was $S_2 \rightarrow S_1$.

Take $S_3$ as the starting point, search for the adjacent spatial units as $S_1$, $S_2$, $S_{24}$, $S_{25}$ and calculate the $f_{3j}$. As shown in Table 4. From Table 4, we note that $S_3$ and $S_1$ have the maximum correlation coefficient, $\max f_{31} = 0.68$.

Determine $S_1$ and $S_3$ in one direction connection, as $O_{13}/D_{13} = 57.6 > 1$, so the only direction connection between $S_3$ and $S_1$ was $S_1 \rightarrow S_3$.

Take $S_{14}$ as the starting point, search for the adjacent spatial units as $S_1$, $S_9$, $S_4$ and calculate the $f_{14j}$. As shown in Table 5.

From Table 5, we note that $S_{14}$ and $S_1$ have the maximum correlation coefficient, $\max f_{141} = 0.33$. Determine $S_{14}$ and $S_1$ in one direction connection, as $O_{14}/D_{14} = 1.72 > 1$, so the one direction connection between $S_1$ and $S_{14}$ was $S_{14} \rightarrow S_1$.

There are not any external connection by the unit $S_1$, so the units $S_1$, $S_2$, $S_3$, $S_{14}$ which have fused cluster the first class $C_1 = \{S_1, S_2, S_3, S_{14}\}$, the fused path was $S_2 \rightarrow S_1; S_3 \rightarrow S_1; S_{14} \rightarrow S_1$.

As the calculation process is same, the other class calculation is not repeated, the fused cluster steps as shown in Table 6.
The transport relation with spatial units has its own characteristics and the characteristics distributed on the geospatial form the different transportation regions. Shown in Table 7.

<table>
<thead>
<tr>
<th>No.</th>
<th>Code</th>
<th>Name of region</th>
<th>Composition of region</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>C₁</td>
<td>Northern</td>
<td>Hebei, Beijing, Tianjin, Shanxi, Mongolia West, Shandong</td>
</tr>
<tr>
<td>2</td>
<td>C₂</td>
<td>Northeast</td>
<td>Heilongjiang, Jilin, Liaoning, Mongolia East</td>
</tr>
<tr>
<td>3</td>
<td>C₃</td>
<td>East</td>
<td>Shanghai, Jiangsu, Anhui, Zhejiang</td>
</tr>
<tr>
<td>4</td>
<td>C₄</td>
<td>Southeast</td>
<td>Fujian, Jiangxi</td>
</tr>
<tr>
<td>5</td>
<td>C₅</td>
<td>Central</td>
<td>Shaanxi, Henan, Hubei</td>
</tr>
<tr>
<td>6</td>
<td>C₆</td>
<td>Southern</td>
<td>Guangdong, Hainan, Hunan</td>
</tr>
<tr>
<td>7</td>
<td>C₇</td>
<td>Southwest</td>
<td>Guangxi, Guizhou, Sichuan, Chongqing, Yunnan</td>
</tr>
<tr>
<td>8</td>
<td>C₈</td>
<td>Northwest</td>
<td>Gansu, Ingrid, Xingjian, Qinghai, Tibet</td>
</tr>
</tbody>
</table>

4 Conclusion

Traditional regional division model for spatial units is only built upon the information of attributes in every unit. However, the spatial relationships and their spatial interaction between units are not considered sufficiently. In this study, based on scale-space theory, the spatial units cluster fusion algorithm is proposed so that the elements of spatial relationship between units could be integrated besides considering the information of attributes. By this approach, considered both of the spatial proximity and attribute similarity at certain spatial range, the spatial units could be clustered into one class if their connective direction is the same and achieve the regional division.

This paper uses the spatial transport linkage theory to construct the spatial units correlation model for railway freight transportation regional division. According to the data of O – D matrix with the adjacent spatial units and considering economic parameters as limitation, the regional division is divided 8 parts of railway freight transportation in China. As Northern, Northeast, East, Southeast, Central, Southern, Southwest and Northwest. As shown in Figure 1. The results are highly accordant with the real situations and can offer spatial background for regional railway freight transportation planning and development.

Further research is also proposed to improve spatial units correlation model. The data selection should not only reflect the present but also reflect the future trends of the regional development. Economic indicators should be further improved when considering the link between transportation and economic and also considering the impact of macroeconomic policies and geographic and other factors and make the model more comprehensive.

FIGURE 1 The spatial layout of railway freight transportation regional division

Acknowledgments

The project has been supported by National Natural Science Foundation, China (Grant No. 71173177) and International Exchange Programs, Muroran Institute of Technology, Japan.

References


**Author**

Yu Wang, born in 1979, Dalian, Liaoning, China.

Current position, grades: lecturer at the School of Traffic and Transportation Engineering, Dalian Jiaotong University, China.

University studies: B.E. at Southwest Jiaotong University in China. ME at Southwest Jiaotong University in China.

Scientific interests: traffic and transportation planning and management.

Publications: more than 10 papers.

Experience: teaching experience of 5 years, presided 1 scientific research project.