

## Research Article

# Quantitative Analysis of Urban Regional Traffic Status

**Qing-fang Yang, Ru-ru Xing, Li-li Zheng, and Shu-xing Wang**

*College of Transportation, Jilin University, Changchun 130022, China*

Correspondence should be addressed to Li-li Zheng; [zlldtq1024@163.com](mailto:zlldtq1024@163.com)

Received 12 January 2017; Revised 2 April 2017; Accepted 4 April 2017; Published 15 May 2017

Academic Editor: Qin Yuming

Copyright © 2017 Qing-fang Yang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to monitor the real-time operation condition of urban region traffic flow, and to quickly identify regional traffic status, this paper adopts CNM (Clauaset-Newman-Moore) Community Division Method of Complex Network to analyze traffic status information deeply implied from the regional road network traffic flow data, which aims to objectively develop the reasonable classification of regional traffic state with no classification criteria of determining regional traffic state. Combined with the regional road network traffic data from a certain city, the example analysis shows that this proposed method can easily provide the reasonable division of regional traffic condition and verifies the feasibility of the regional traffic state classification method. Besides, the example analysis gives the rough regional traffic status determination standard, laying theoretical basis for accurately judging the regional traffic state.

## 1. Introduction

Recently, the rapid development of urban economy brings along the improvement of resident consumption levels. With an increasing amount of private car ownership, there is an increasing trend of serious traffic problems. Urban traffic congestion has gradually spread from a single point line to a connected piece of area and even the whole road network. The evolution of traffic status follows certain patterns, especially the pattern of regional traffic status, which is the prominent jam phenomenon shown by literature [1]. And it is the premise to solve the problem of regional traffic jams in order to quickly and accurately determine the regional traffic congestion and block the traffic congestion to spread further. Therefore, it is particularly vital to study regional traffic state and to quantitatively analyze the regional traffic state discriminant standard which is the crucial problem to be solved.

Regional traffic status, like the road traffic status, belongs to a kind of traffic participants' awareness information, which means it is a subjective concept. Therefore, many studies on regional traffic status demonstrate different understandings. However, there is few research for regional traffic status at this stage. The related research is divided into two categories: qualitative research and quantitative research. The studies of regional traffic status mostly stay on the qualitative research, especially most of the ones studied by Chinese scholars.

Guo et al. [2] and Ji et al. [3] mostly adopted the way of introducing specific evaluation parameters to qualitatively analyze the regional traffic status. The evaluation parameters are derived from crossing traffic parameters and road traffic parameters. However, it is not a reliable method that only uses the intersection or road traffic flow parameters in the road network to directly reflect the running state of the road network traffic flow. It needs a macro analysis of the regional road network traffic state.

Therefore, Leng [4] used three parameters, road network efficiency, road network accessibility, and road network connectivity, to evaluate network running state. But the research cannot clearly show how the three evaluation parameters reflect the state of the network. Also from the macroscopic angle, He et al. [5] introduced the road network congestion intensity parameter and combined Gaussian Mixture Model to cluster the sample data and then analyzed and evaluated the regional traffic congestion scope.

In the absence of considering the real regional traffic state classification level, Xiao-feng Ji directly referred to the standard of the evaluation index system of the city traffic management to divide the regional traffic state into four grades. Then, he extracted the traffic flow characteristics to qualitatively analyze the regional traffic state. This part of the research is not rigorous.

Considering the necessity for the principle of clustering algorithms to set the number of categories and, also, considering the principle of the need to set the number of categories, Guo directly divided the regional traffic state into three levels according to the “flow-density” diagram.

To deal with this problem, He introduced Silhouette average width to determine the optimal clustering number of the sample. However, the research aims at the spatial distribution of regional traffic congestion problem.

Zhang et al. [6] introduced two parameters—road accessibility and road connectivity, and fixed a time period for the quantitative research of the regional traffic state. Although, the fluctuation change of the regional traffic status in the research period can be reflected by evaluating the changes of the parameter values. He did not quantitatively research the fluctuation change.

Daganzo [7], Gonzales et al. [8], and Geroliminis and Daganzo [9] researched regional traffic status from the perspective of Macroscopic Fundamental Diagrams and explored the evolution role of regional traffic status through building the models of the Macroscopic Fundamental Diagrams, which is able to roughly reflect the regional traffic status. But, the divided regional traffic state is obtained by the Macroscopic Fundamental Diagrams, which is rough.

Hence, this paper, on the basis of the existing research, from a macro point of view, introduces the global efficiency as the evaluation parameter. Then, it uses the complex network clustering algorithm for quantitative analysis of the reasonable regional traffic condition level. Finally, it puts forward a quantitative method to solve the problem for the regional traffic state discrimination.

The main contribution of this paper is to propose a method for confirming the level of dividing regional traffic state, as the basis to identify regional traffic state. So, this paper uses Complex Network Theory to research regional traffic state from the perspective of quantitative analysis. A rough discriminant standard for regional traffic status will finally be obtained, which is realized to generally determine the regional traffic state.

Due to the complex structure in regional road network, the research should take the connection relationships between the roads into consideration. Hence, abstract the urban regional transportation network into a complex network. Then, use CNM (Clauset-Newman-Moore) Community Division Method to deeply explore regional traffic status information from a data set composed of traffic parameters named global efficiency, and dynamically divide the regional traffic status into several grades. On this basis of the above, build the corresponding relation between the global efficiency and the parameters of determining current urban trunk state discriminant standard. Finally, according to the function transformation, this paper roughly makes sure of the state discriminant threshold of regional road network under various traffic states and obtains a general standard which will quantitatively reflect the regional network traffic state situation.

After this introduction, the remainder of the paper is organized as follows: Section 2 presents the theoretical basis to the studied method and the research background

environment; Section 3 describes the proposed method; Section 4 uses the regional traffic data to analyze and discuss the method; Section 5 summarizes the main research conclusions; Acknowledgments simply introduces the sources of funding of this paper.

## 2. Urban Regional Traffic Network Construction

*2.1. The Basis of Using Complex Network to Research Regional Traffic.* The research of this article mostly uses the Complex Network Theory. The reasons why this article adopts the Complex Network Method lie in three aspects:

- (1) From the topology structure level of traffic network, complex network abstracted by urban traffic network conforms with scale-free network but also shows the features of small-world complex network [10].
- (2) From the dynamics behavior level of traffic network, it is found that the cause of the wide range of traffic congestion is a threshold in either point in scale-free networks, and edge congestion or network cascading failure. If it exceeds the threshold, the traffic network will face larger areas of congestion.
- (3) From evaluation characteristics of traffic network, urban traffic network has self-similarity structure. Its fractal dimension is not invariable and has a reciprocal relationship with the time value.

Based on the above, this article constructs a complex network to analyze the complex regional traffic status and applies the Complex Network Theory to divide the regional traffic status.

*2.2. Urban Regional Traffic Network Construction.* For the study of complex urban regional traffic status, this research will not only consider the static network topology character of the urban area network but also analyze the features of dynamic time-varying embodied traffic flow operating condition from the network. Therefore, this article will firstly abstract urban regional traffic network into a complex network. Then, it will research the essence of the regional road network traffic flow operating state from the angle of network topology.

This article uses the Primal Approach Method [11] to regard intersections from the regional traffic network as complex network nodes, each section as edge, and vehicles driving on the roads as information transmission packets of complex network. Considering the complex urban regional transportation network structure and nonuniformity running direction of traffic flow, and the different levels of the maximum capacity under traffic flow, this paper builds a directed weighted complex network  $N = (V(N), E(N), w(N))$  on the basis of regional road network traffic, which is vertex set;  $E(N)$  is frontier set;  $w(N)$  is weighted set. This abstraction not only reflects imbalance road network but also incarnates the influence of regional network from different level roads.

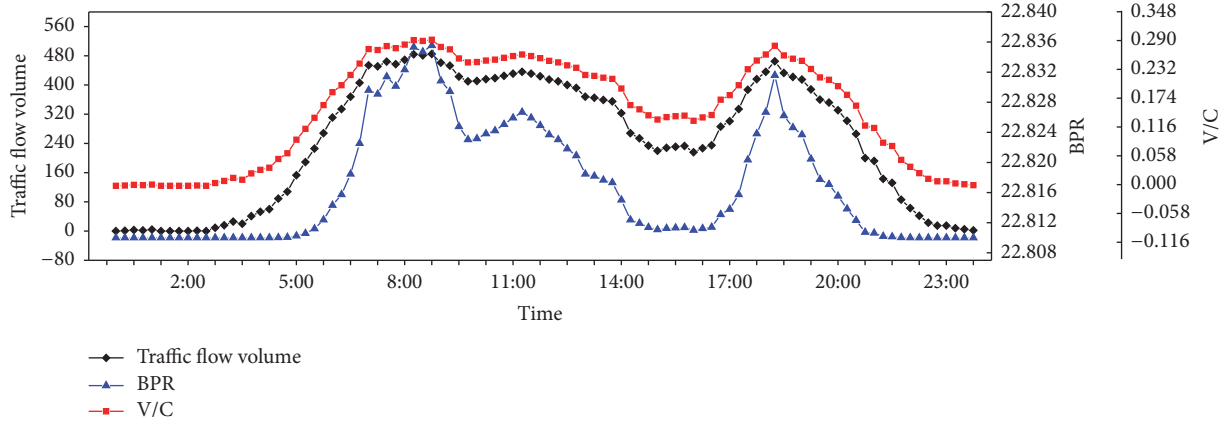


FIGURE 1: Contrastive analysis between BPR and V/C.

2.3. *Determining Method of Network Weighted Set  $w(N)$ .* The research of the complex network edge-weight roughly has been divided into two categories: fixed edge-weight model and dynamic-weight model [12]. Urban traffic network is a complex dynamic network. And the operation efficiency of most traffic flow in the network depends on the roads and intersections. Namely, in complex network, the transmission efficiency of information packet is affected by the edge-weight. Thus, this article builds a complex network. And the edge-weight from this network is recognized as a dynamic parameter changing over time. Literature [13] introduced a calculation method of road impedance (BPR), as follows:

$$w_a = w_0 \left[ 1 + \alpha \left( \frac{q_a}{c_a} \right)^\beta \right], \quad (1)$$

where  $w_a$  is the impedance of the road  $a$  and  $w_0$  is free-flow impedance, which means  $w_0$  is the free-flow travel time.  $q_a$  is traffic flow volume on the road  $a$ ;  $c_a$  is the actual capacity of road  $a$ , which means  $c_a$  is the actual vehicles number per unit time.  $\alpha$  and  $\beta$  are the retardation factors, the parameter values are generally  $\alpha = 0.15$  and  $\beta = 4$  under highway traffic flow distribution program in the United States.

In addition, some scholars [14] use V/C as the source of complex network edge-weight. Namely,

$$w_{ij} = \frac{q_{ij}}{c_{ij}}, \quad (2)$$

where  $w_{ij}$  is the edge-weight between node  $v_i$  and node  $v_j$  on complex network;  $q_{ij}$  is the actual traffic flow volume between node  $v_i$  and node  $v_j$  under urban regional traffic network;  $c_{ij}$  is the actual capacity between node  $v_i$  and node  $v_j$  under urban regional traffic network, which means it stands for the maximum flow volume passing under per hour on the road in a certain period of time.

In view of the two edge-weights calculation forms, put traffic flow data of an arterial road in Changchun to analyze the sensitivity shown by two edge-weights calculation equations. So, the contrastive analysis between BPR and V/C is seen in Figure 1.

In Figure 1, it is shown that V/C can accurately depict the changes in road traffic flow, which is able to reflect the traffic

state actual information. BPR function, on the other hand, due to the equation containing power exponent function, changes the traffic flow operating tendency during part time.

However, the road traffic flow volume  $q_{ij}$  is related to the road length. The following is the modified equation (2).

$$w_{ij}^* = \frac{1}{l_{ij}} \times \frac{q_{ij}}{c_{ij}} = \frac{w_{ij}}{l_{ij}}, \quad (3)$$

where  $l_{ij}$  is the length of the edge between node  $v_i$  and node  $v_j$ . The modified edge-weight  $w_{ij}^*$  represents V/C within the unit length.

Therefore,  $w_{ij}^*$  is chosen as a way to calculate the value of a network edge-weight.

### 3. The Quantitative Analysis Method of Regional Traffic State

After solving the problems of determining edge-weight value in complex network, the network information transmission efficiency needs a parameter to be quantitatively evaluated. In directed and weighted network, global efficiency is an evaluation parameter for the network information transmission efficiency, and its numerical size reflects the complex network connecting degree, which means depicting the specific situation of the regional traffic state. Therefore, this article introduces the global efficiency to evaluate the regional road network traffic flow operating state. And literature [15] considered the fact that under the same traffic condition traffic flow data has same properties. So, this paper focuses on how to quantitatively analyze the similarity between traffic flow data under all kinds of traffic conditions.

3.1. *The Definition and Calculation of Global Efficiency.* The global efficiency is a measure for reflecting network operation ability, and it can judge the efficiency of the whole network information transmission, namely, the operating efficiency of traffic flow within regional road network [16]. To define the global efficiency, give some concepts, like shortest path, dissimilar-weight, and similar-weight first.

In undirected complex network, the distance  $d_{ij}$  (the shortest path length) between node  $v_i$  and node  $v_j$  refers to the minimum number of edges connected with node  $v_i$  and node  $v_j$ . In addition, the inverse of this distance refers to the efficiency between node  $v_i$  and node  $v_j$ .

For the weighted network, there are two kinds of edge-weights—dissimilar-weight and similar-weight. Normally, the most popular aspect is the dissimilar-weight relating to distance: the larger weight value means the further distance between two nodes. Nevertheless, it is significantly different from the similar-weight: the larger weight value indicates the closer distance between nodes (higher density among nodes) [17].

The distance  $\hat{d}_{ij}$  between node  $v_i$  and node  $v_j$  needs to consider each edge-weight in a network. If weight  $w_{ij}$  of edge  $e_{ij}$  expresses dissimilar-weight, the relationship between the distance  $\hat{d}_{ij}$  and its weight  $w_{ij}$  is  $\hat{d}_{ij} = w_{ij}$ . In the case of dissimilar-weight, if node  $v_i$  and node  $v_j$  connect through two edges ( $e_{ik}$  and  $e_{kj}$ ) with weight  $w_{ik}$  and weight  $w_{kj}$ , respectively, the relationship is  $\hat{d}_{ij} = w_{ik} + w_{kj}$ . If weight  $w_{ij}$  of edge  $e_{ij}$  expresses similar-weight, the relationship is  $\hat{d}_{ij} = 1/w_{ij}$ . In the case of similar-weight, the relationship is  $\hat{d}_{ij} = 1/w_{ik} + 1/w_{kj}$ .

Therefore, based on the complex network environment built in this paper, the equation for calculating the shortest path is as follows:

$$\hat{d}_{ij} = w_{ik} + w_{kj}. \quad (4)$$

At this time, the efficiency  $\hat{\varepsilon}_{ij}$  between node  $v_i$  and node  $v_j$  is as follows:

$$\hat{\varepsilon}_{ij} = \frac{1}{\hat{d}_{ij}} = \frac{1}{w_{ik} + w_{kj}}, \quad (5)$$

where  $w_{ik}$  and  $w_{kj}$  express the edge-weights of the shortest path connected with node  $v_i$  and node  $v_j$ , respectively.

In the weighted network, the global efficiency  $E_{\text{global}}$  can be defined as the average efficiency between all the nodes in the network.

$$E_{\text{global}} = \frac{1}{N(N-1)} \sum_{i,j \in G, i \neq j} \hat{\varepsilon}_{ij} = \frac{1}{N(N-1)} \sum_{i,j \in G, i \neq j} \frac{1}{\hat{d}_{ij}}, \quad (6)$$

where  $N$  is the number of nodes in the complex network,  $\hat{\varepsilon}_{ij}$  is the efficiency between node  $v_i$  and node  $v_j$ , and  $\hat{d}_{ij}$  is the shortest distance between node  $v_i$  and node  $v_j$ .

**3.2. The Traffic Significance of the Global Efficiency.** From (6), the network global efficiency  $E_{\text{global}}$  can be obtained by calculating the shortest path length of each node in the network. In addition,  $E_{\text{global}} \propto 1/w_{ij}$ . Due to the fact that edge-weight  $w_{ij}$  is determined by  $V/C$  within the unit length, the network global efficiency  $E_{\text{global}}$  is inversely proportional to traffic flow density  $K_{ij}$  ( $K_{ij} = q_{ij}/l_{ij}$ ), namely,  $E_{\text{global}} \propto 1/K_{ij}$ , which shows that global efficiency decreases with the increase of road traffic flow density. In other words, the higher

the values of the global efficiency are, the less congested the roads are and the better the traffic state is. As a result, the global efficiency can depict cities regional road network traffic state from the perspective of complex network.

**3.3. The Regional Traffic State Classification Method Based on CNM.** Based on the recent research, in order to quantitatively analyze the regional traffic state, firstly study the regional traffic state classification method for determining a rough regional traffic state identification standard.

Literature [18] used the Clustering Method to determine the optimal class number of road traffic status with the principle that the difference of data between any two classes is as large as possible and the difference of data within the same class is as small as possible. This paper refers to this theoretical ideology, based on the complex network environment, considering the phenomenon that traffic data has low difference in similar traffic status and significant difference under heterogeneous traffic condition, applying the CNM algorithm to divide regional traffic status, which intends to scientifically require the optimal classification number of regional traffic status.

At present, Community Partition Algorithm is roughly divided into two categories: graphics segmentation algorithms in computer science and hierarchical clustering algorithm in sociology [19]. The latter is divided into Division Method and clustering algorithm. The CNM algorithm can directly build a modularity incremental matrix in order to save the storage space. Considering the scale of the complex network constructed from urban region road network and data amount involved in network, this paper adopts the CNM algorithm for mining community.

**(1) CNM Algorithm Principle.** Due to the network module increment as the research object of the CNM algorithm, define the modularity in complex network firstly.

The calculation of modularity in undirected network module is as follows:

$$Q = \sum_i (e_{ii} - a_i)^2, \quad (7)$$

where  $e_{ii}$  denotes the ratio between all the edges within community  $i$  and all the edges in the network and  $a_i = \sum_j e_{ij}$  denotes the ratio between the edges of nodes connected with nodes in community  $i$  and all the edges of the network.  $e_{ij}$  denotes the ratio between all the edges within community  $i$  and community  $j$ .

The CNM algorithm directly constructs a modularity incremental matrix  $\Delta Q$ . Similar to the adjacency matrix, the sum of each element in each row of this modularity incremental matrix  $\Delta Q$  is  $a_i$ . It is easily understood that when two communities are not connected to each other, the value of the modularity incremental matrix is invariable. The modularity incremental matrix  $\Delta Q$  only stores the connected communities, which saves the storage space. In these matrix mining communities, the largest elements of each row in the modularity incremental matrix  $\Delta Q$  and its corresponding community numbers are saved into the maximum heap  $H$ , which means the maximum heap  $H$  is similar to the data

storage space of the CNM algorithm. Use the maximum heap  $H$  to find the element and its corresponding community numbers in the shortest time. Then, merge the corresponding communities and achieve the ultimate community mining.

In the face of massive traffic information from the regional road network, the quantitative method of analyzing regional traffic state should have a high operation rate and large storage space. The CNM algorithm constructs a modularity incremental matrix, with saving the computing storage space and also improving the operation rate. Considering the advantage of the CNM algorithm, the method this paper proposed is able to improve the operating speed of regional traffic state identification.

(2) *The Regional Traffic State Classification Method Based on CNM*. Based on (6) in Section 3.1, combined with road traffic flow information under regional network states, global efficiency data set  $G$  of the regional road network can be got (i.e.,  $G = \{E_{\text{global } i} \mid i = \text{collect data interval}\}$ ). Regard each data element in the data set as nodes in the network, and build a complex network  $O$  based on the global efficiency data set  $G$ . The complex network  $O = (V(O), E(O))$  is an undirected network, where  $V(O)$  is vertex set and  $E(O)$  is frontier set. The vertex set  $V(O)$  is derived from each element of the global efficiency data set  $G$ . And each element of the frontier set  $E(O)$  is decided by the similarity of each element of the global efficiency data set  $G$ . Use the CNM algorithm to divide communities from this undirected network, and then objectively get the classification of regional traffic state. According to the CNM algorithm, the specific implementation process of regional traffic state classification method is shown in Figure 2.

*Step 1* (build a complex network  $O$  based on the global efficiency sample set  $S$ ). Through the collected information of regional traffic flow parameters in multiple time scales, calculate the global efficiency of regional road network under the whole time and construct the initial sample set of global efficiency  $S = \{s_j\}$ . Then abstract each sample of the set as a node  $v_i$  in the complex network  $O$ , and abstract the connected relationship based on the numerical similarity of every sample as edges. According to the characteristics of this network being abstracted by sample set  $S$ , it is an undirected network.

*Step 2* (establish Euclidean distance matrix  $\mathbf{M}$ ). According to the following equations, calculate initial samples of Euclidean distance matrix  $\mathbf{M}$ :

$$m_{ij} = \sqrt{(s_i - s_j)^2}$$

$$\mathbf{M} = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{nn} \end{bmatrix}, \quad (8)$$

where  $s_i$  and  $s_j$  are the elements of global efficiency sample set and  $\mathbf{M}$  is a  $(n \times n)$  matrix.

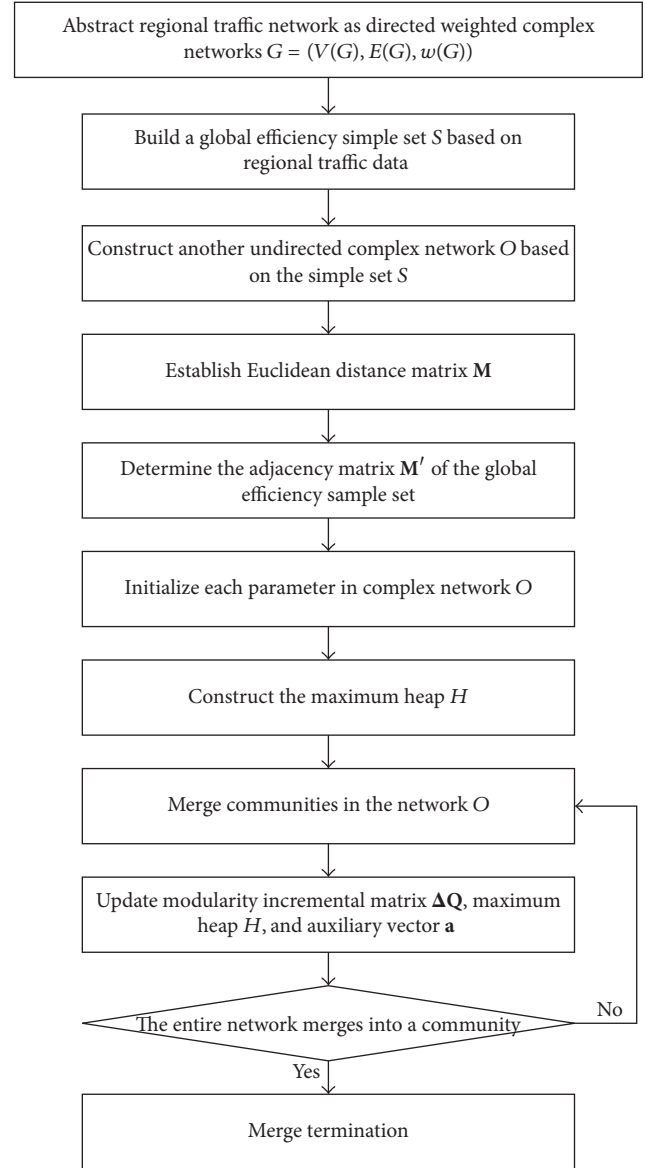


FIGURE 2: Flow chart of quantitative method.

Due to the built complex network as an undirected network, it can be ratiocinated that  $m_{ij} = m_{ji}$  and  $m_{ii} = 0$ , which means the Euclidean distance matrix  $\mathbf{M}$  is a symmetric matrix.

*Step 3* (determine the adjacency matrix  $\mathbf{M}'$  of the global efficiency sample set  $S$ ). Based on the constructed Euclidean distance matrix  $\mathbf{M}$ , compare the truncated threshold  $R$  with each element of matrix  $\mathbf{M}$ . If the numerical value of element  $M_{ij}$  in the matrix  $\mathbf{M}$  is less than the truncated threshold  $R$ , define that the two nodes  $v_i$  and  $v_j$  are connected to each other. Using the truncated threshold  $R$  to rearrange the Euclidean distance matrix  $\mathbf{M}$ , the global efficiency adjacency matrix  $\mathbf{M}'$  of the global efficiency sample set  $S$  will be obtained.

$$\mathbf{M}' = \begin{bmatrix} m'_{11} & m'_{12} & \cdots & m'_{1n} \\ m'_{21} & m'_{22} & \cdots & m'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m'_{n1} & m'_{n2} & \cdots & m'_{nm} \end{bmatrix} \quad (9)$$

$$\text{among, } m'_{ij} = \begin{cases} 1, & m'_{ij} \leq R \\ 0, & m'_{ij} > R. \end{cases}$$

*Step 4* (initialize each parameter in complex network  $O$ ). Assuming that the complex network  $O$  contains  $N$  nodes (that is, the global efficiency sample set  $S$  contains  $N$  samples), initialize its network parameters. Regard the network as  $N$  communities, which means each node is seen as a community, while the initial modularity of every community equals zero ( $Q_i = 0$ ). The calculation methods of the initial values of

$$\Delta q_{ij} = \begin{cases} \frac{1}{2M} - \frac{k_i k_j}{(2M)^2}, & \text{node } v_i \text{ and node } v_j \text{ are connected by edges} \\ 0, & \text{else.} \end{cases} \quad (11)$$

*Step 5* (construct the maximum heap  $H$ ). The maximum heap contains the largest element  $\hat{v}_{ij}$  of each row from the modularity incremental matrix  $\Delta \mathbf{Q}$  in the network. And it also includes the community number  $i$  and number  $j$  corresponding to this largest element  $\hat{v}_{ij}$ .

After Step 4, the initial modularity incremental matrix  $\Delta \mathbf{Q}$  corresponding to the maximum heap  $H$  can be got.

*Step 6* (merge communities in the network  $O$ ). According to the value of the maximum heap  $H$ , firstly, pick out the

$$\Delta q_{jk} \leftarrow \begin{cases} \Delta q_{ik} + \Delta q_{ik}, & G_k \text{ is connected with } G_i \text{ and } G_j \\ \Delta q_{ik} - 2a_j a_k, & G_k \text{ is connected with } G_i, \text{ not connected with } G_j \\ \Delta q_{jk} - 2a_i a_k, & G_k \text{ is connected with } G_j, \text{ not connected with } G_i. \end{cases} \quad (12)$$

(2) The updating rules of  $H$ : after updating  $\Delta \mathbf{Q}$ , update the maximum elements of the corresponding row and column in the maximum heap  $H$ .

(3) The updating rules of  $\mathbf{a}$ :

$$a_j \leftarrow a_i + a_j, \quad \text{then } a_i \leftarrow 0. \quad (13)$$

And record the merged modularity value:

$$Q \leftarrow Q + \max \{ \Delta q_{ij} \}. \quad (14)$$

*Step 8* (merge termination). In the process of merging communities, if the updated largest element of network modularity incremental matrix changes from positive to

each element  $e_{ij}$  and element  $a_i$  in the matrix  $\mathbf{E}$  and auxiliary vector  $\mathbf{a}$  are shown below.

$$e_{ij} = \begin{cases} \frac{1}{2M}, & \text{node } v_i \text{ and node } v_j \text{ are connected by edges} \\ 0, & \text{else} \end{cases} \quad (10)$$

$$a_i = \frac{k_i}{2M},$$

where  $k_i$  represents the degree of node  $v_i$  (the number of adjacent edges of node  $v_i$  is its degree  $k_i$ ) and  $M$  represents the total number of edges in the network.

In that way, the initial elements in modularity incremental matrix  $\Delta \mathbf{Q}$  will satisfy the following conditions:

maximum value  $\Delta q_{ij}$ , and merge the community  $G_i$  and community  $G_j$  corresponding to the element  $\Delta q_{ij}$ . Then mark the merged community as  $G_j$ .

*Step 7* (update the parameters). After the communities merging, update modularity incremental matrix  $\Delta \mathbf{Q}$ , maximum heap  $H$ , and auxiliary vector  $\mathbf{a}$ . The updating rules of three kinds of data are as follows.

(1) The updating rules of  $\Delta \mathbf{Q}$ : delete the elements of row  $i$  and column  $i$ , and update the elements of row  $j$  and column  $j$ . Then, we can obtain

negative, stop this merger. The division of the community structure is optimal.

The differences between the divided communities demonstrate the differences among regional traffic status. If the number of divided communities is more, regional traffic status changes more; when the number of communities is small, the state of regional traffic is stabilized in this case. But when the number of the communities is only one, it is shown that there are two kinds of traffic condition in this regional network: firstly, the overall regional network is unblocked, namely, any roads in a state of free-flow area. Secondly, the overall regional network is blocked, namely, all roads in the crowded state.

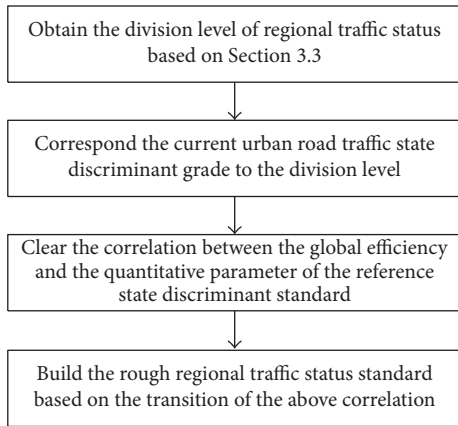


FIGURE 3: Process of determination method of the rough regional traffic congestion standard.

Therefore, in determining the quantitative standard of regional traffic status, all-time traffic data will be chosen as this research object.

3.4. *The Method Application: The Quantitative Method of Obtaining Regional Traffic State Standard.* According to the division principle in Section 3.3, the classification level of regional traffic status will be objectively obtained. In order to discriminate the actual regional traffic status, this part is based on the quantitative method of regional traffic state from Section 3.3. It proposes a method that a rough standard of regional traffic state could be obtained. The proposed classification method provides theoretical basis for the further study of regional traffic state discrimination. The specific process of determining the rough standard is shown in Figure 3.

### 4. Instance Analysis

4.1. *Example Verification.* This article utilizes a regional network in Changchun as an example for confirming the proposed method feasibility, where this network has obvious regional traffic congestion characteristic. There are 14 crossings and 21 links in this traffic network, as shown in Figure 4, where line width of roads stands for road grades in the area. Combining the collected artificial-investigation road traffic flow information of the regional network and the data collected from sensors laid in the network to build the data set, the data set reports every 15 min with all the time interval 0:00–24:00 from 23 to 27 December in 2013, including the traffic flow volume, the vehicle speed, and the network geometric attributes (traffic signals, facility type, etc.). Then calculate the quantitative parameter of regional traffic status division—the global efficiency  $E_{\text{global}}$ ; finally, build the global efficiency data set  $E = \{e_i\}$ .

(1) *Overall Analysis of Network.* Convert the relationship data of each element in the global efficiency set  $E = \{e_i\}$  into the adjacent matrix on the basis of the calculation method of network adjacent matrix in Steps 1 and 2 from Section 3.3,

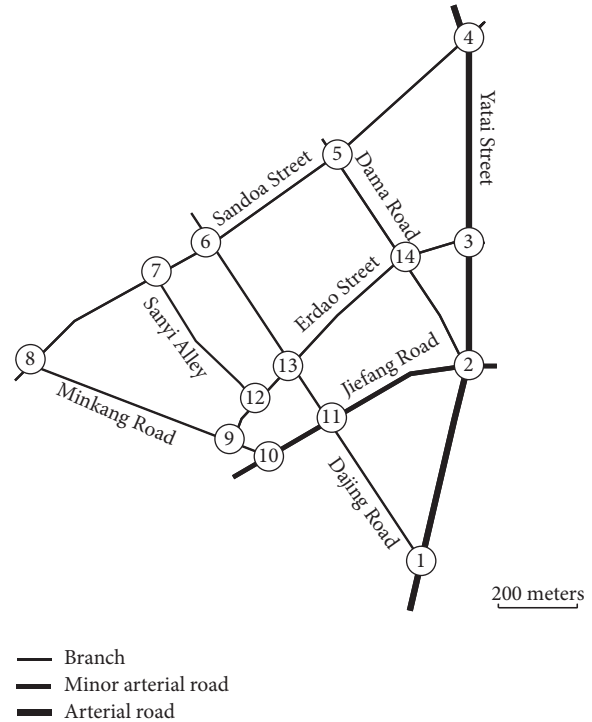


FIGURE 4: Schematic diagram of example area.

where “1” denotes that the two elements in global efficiency set are related and “0” denotes that there is no relation between these elements. Utilize UCINET software to directly describe the relation of elements in the global efficiency set by network diagram, as is shown in Figure 5. And utilize R software to make the correlogram of nodes in complex network in Figure 6. Due to the bigger scale of regional traffic data and the higher number of abstracted nodes in complex network, the whole complex network is not easy to display in image. Therefore, one complex network constructed with one-day global efficiency data is indicated in Figure 5, which reflects the connection character of regional traffic network.

In Figure 5, “circle” stands for the nodes in a network, namely, the corresponding elements in the data set of the global efficiency  $E = \{e_i\}$ . The “double-side arrow” stands for the fact that nodes are linked by the sides, which means there are certain relationships between the corresponding elements in the data set of the global efficiency  $E = \{e_i\}$ , but it does not reflect that the network is a directed complex network. In Figure 6, the higher the color is, the greater the relevance of the variable is.

In addition, through the analysis of Figure 5, it is found that these nodes are divided into three types based on node degree: one is that there are node degrees of zero in these nodes, which means such nodes do not connect to other nodes. The second is the one that has one node degree, which means they only connect to one node. The third is the one where those node degrees are greater than one, which means these nodes connect to multiple nodes. The changes of regional traffic status are just reflected by the changes of

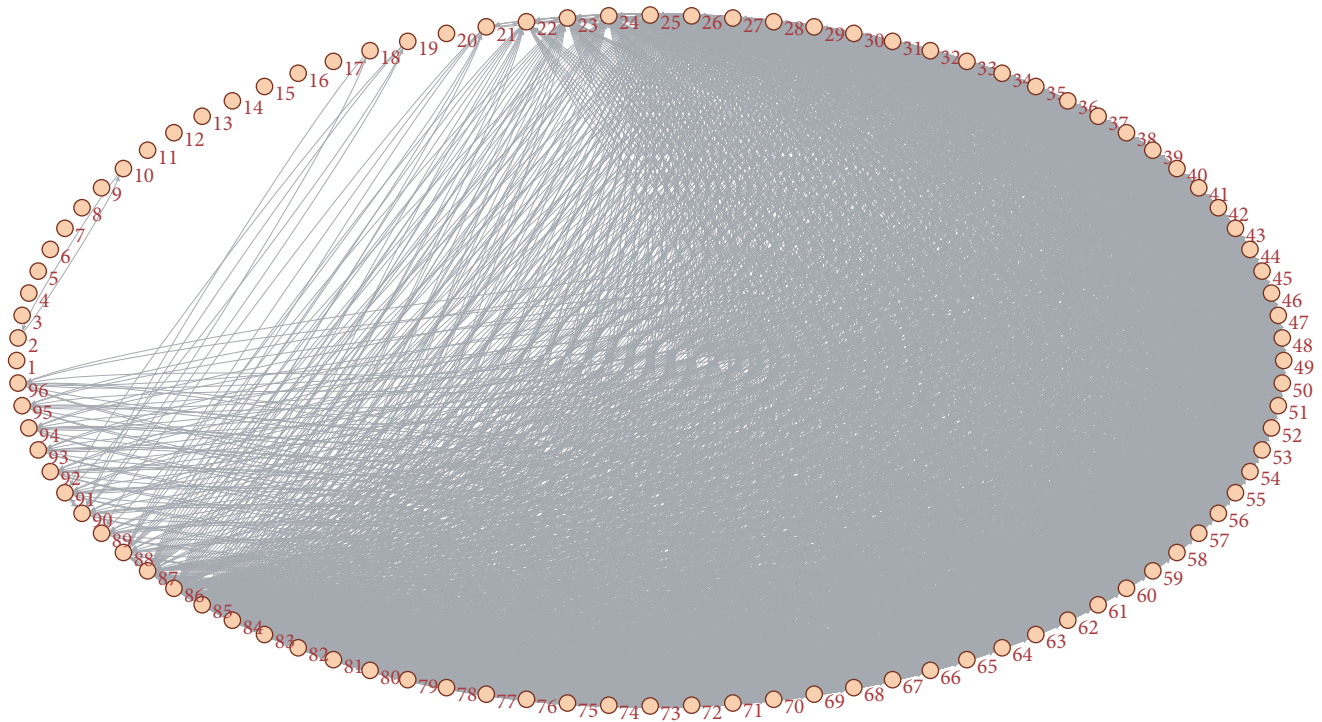


FIGURE 5: Complex network based on global efficiency data set.

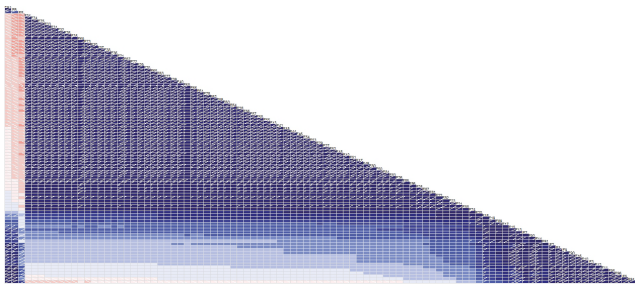


FIGURE 6: Correlogram of nodes in complex network.

node degree, which confirms that the global efficiency  $E_{\text{global}}$  can reflect the regional traffic state information.

(2) *Community Division.* In dividing communities from the complex network in Figure 5, use MATLAB to convert the algorithm from Section 3.3 into computer language in order to get the corresponding community division result. The specific results of node dividing are shown in Figure 7. Due to the large amount of data, its dividing result cannot be fully displayed in Figure 7, which reflects the result with one-day global efficiency data.

In Figure 7, it can be seen that the red dotted line on the basis of the degree of module maximum value divides complex network into four communities. It means that, according to data characteristics of actual network quantitative parameters, this complex network based on the global efficiency set ( $E = \{e_i\}$ ) is divided into 4 communities. Thus, it reflects that this urban regional traffic status must

be divided into 4 levels by the objective law of traffic data information.

(3) *Determine the Regional Traffic State Discrimination Standard.* The Public Security Traffic Management Department published “Urban Traffic Management Evaluation Index System” in 2012. The system divides road traffic status into four categories to describe regional congestion, by using the vehicle average travel speed of urban arterial road to evaluate road traffic states. Use the built data set to fit the relationship between the global efficiency  $E_{\text{global}}$  of regional traffic network and the vehicle average travel speed  $\bar{v}$  of urban arterial road in the regional traffic network, as shown in Figure 8 and Table 1.

In Figure 8, 10 kinds of common function are, respectively, used in curve fitting to analyze the relationship between the global efficiency and the average travel speed. According to the values of all kinds of curve estimation parameters in Table 1, it is found that the values of sig. of all the curves are less than 0.05, which means the relationships between each curve and the two parameters are very significant. Therefore, this article chooses the estimation function of the highest correlation curve as a function of the relations between the two parameters, namely, quadratic function. Then, the particular form of this function is shown below:

$$E_{\text{global}} = 0.0559 - 0.0042\bar{v} + 0.0001\bar{v}^2. \quad (15)$$

As seen in Figure 8, the global efficiency and the average travel speed are in positive relationship, which reflects that the global efficiency value will increase with the growth of the average travel speed in urban arterial road. At the same time,



TABLE 1: Model summary and parameter estimates.

Equation	Model summary					Parameter estimates			
	R-squared	F	df1	df2	Sig.	Constant	b1	b2	b3
Linear	0.9139	5064.8053	1.0000	477.0000	0.0000	-0.0452	0.0029		
Logarithmic	0.8494	2689.9098	1.0000	477.0000	0.0000	-0.2259	0.0793		
Quadratic	0.9722	8324.9882	2.0000	476.0000	0.0000	0.0559	-0.0042	0.0001	
Cubic	0.9705	7827.1259	2.0000	476.0000	0.0000	0.0249	-0.0008	0.0000	1.449E - 006
Compound	0.9461	8369.1022	1.0000	477.0000	0.0000	0.0046	1.0751		
Power	0.9078	4697.8380	1.0000	477.0000	0.0000	0.0000	1.9898		
S	0.8394	2494.0357	1.0000	477.0000	0.0000	-1.4498	-51.3362		
Growth	0.9461	8369.1022	1.0000	477.0000	0.0000	-5.3797	0.0725		
Exponential	0.9461	8369.1022	1.0000	477.0000	0.0000	0.0046	0.0725		
Logistic	0.9461	8369.1022	1.0000	477.0000	0.0000	216.9557	0.9301		

The independent variable is 39.480000.

TABLE 2: Status identification standard.

	First level	Second level	Third level	Forth level
Average travel speed	$v \geq 40$ km/h	$v < 40$ km/h, $v \geq 30$ km/h	$v < 30$ km/h, $v \geq 15$ km/h	$v < 15$ km/h
Global efficiency	$E_{\text{global}} \geq 0.0479$	$E_{\text{global}} < 0.0479$ , $E_{\text{global}} \geq 0.0199$	$E_{\text{global}} < 0.0199$ , $E_{\text{global}} \geq 0.0154$	$E_{\text{global}} < 0.0154$

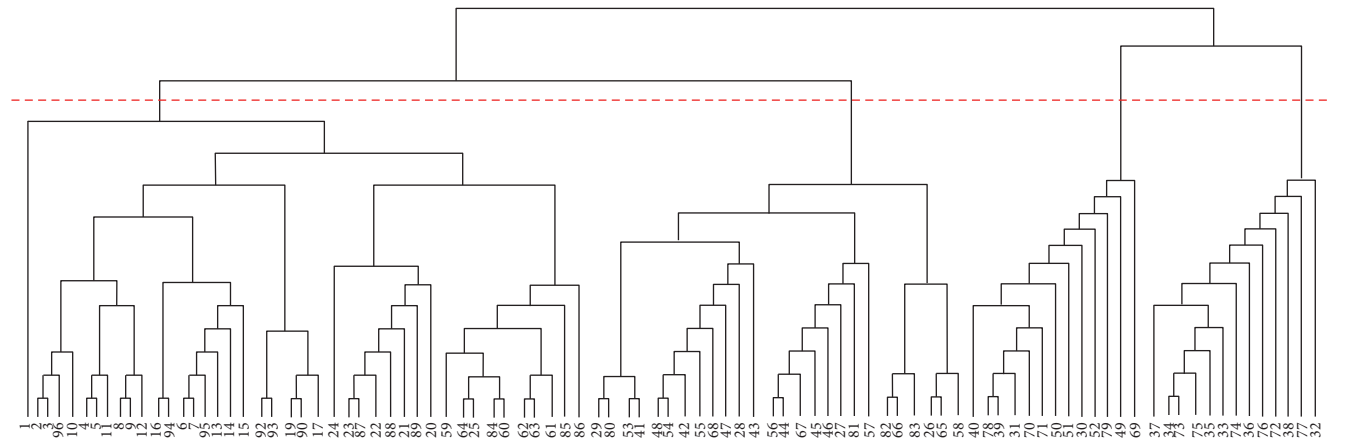


FIGURE 7: Dividing result with one-day data in regional traffic network.

the relationship between these two parameters in Figure 8 confirms the above expounded traffic significance of the global efficiency.

According to the relationship of quantitative parameters, convert the level of the evaluation index system of the city traffic management into regional transportation discriminant criteria, as shown in Table 2.

Therefore, the traffic state of this example area is divided into four levels. Get real-time traffic parameters to determine the global efficiency and then obtain the corresponding state level of regional traffic congestion, which will provide strong information support for traveler to travel and traffic managers to make the congestion control scheme.

4.2. The Feasibility Analysis of the Proposed Method. According to the example in Section 4.1, the proposed method can

reasonably quantify regional traffic status, from a theoretical basis of obtaining the regional traffic state information. The method, from the perspective of data mining, based on mass data of regional traffic flow parameters, analyzes the regional traffic state information. On the other hand, compared to the widely used clustering algorithm, Neural Network Algorithm [20–24] uses the K-means Method [25] with better clustering effect to feasibly analyze the regional traffic state’s quantitative methods.

Based on the global efficiency data set, this regional traffic status is also divided into 4 levels by using K-means Method. It means this regional traffic status is better to be researched as 4 levels. And the proposed method is feasible to quantitatively analyze regional traffic status and provides the best grade of regional traffic state. Figure 9 shows the effect of cluster index of K-Means clustering results, where the weighted average

TABLE 3: The comparison results of three methods.

	Proposed method	SOM Algorithm	K-means Method
Calculation time (s)	8.510	18.193	10.830

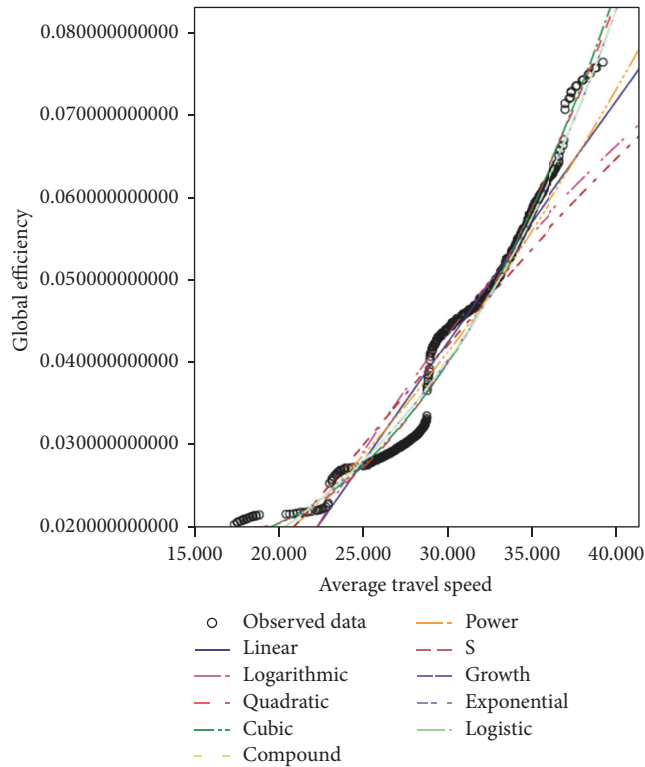


FIGURE 8: Regression analysis of evaluation parameters.

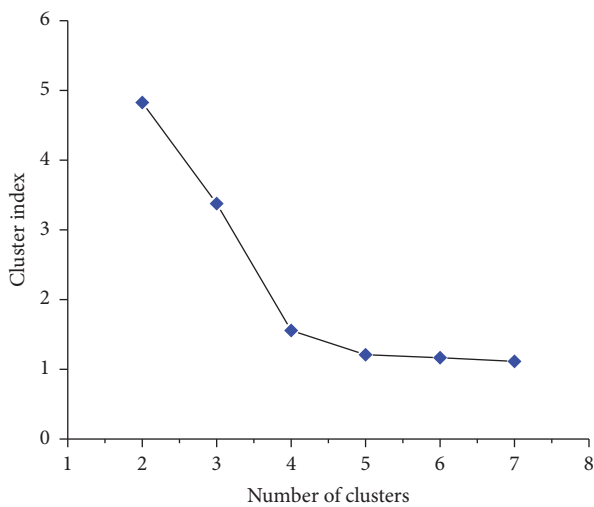


FIGURE 9: Function relationship of evaluation parameters.

of the mean centroid distances of  $K$  clusters is chosen as clustering index. It can be easily seen that when the number of clusters is 4, cluster index declines rapidly, which confirms that the correct number of clusters is 4 and the result is the same as the proposed method.

Use the SOM Algorithm, the  $K$ -means Method, and the proposed method to cluster the everyday collected data. Then, the averaged calculation times of every three methods are shown in Table 3. It can be seen that the proposed method is faster than the SOM Algorithm and the  $K$ -means Method in Table 3.

Because of the traffic flow data related to the area scope, the Neural Network Algorithm spends too much time in building a network. So, it has the longest time. The clustering algorithms needs to firstly calculate the clustering centers. But, the proposed algorithm directly picks up the parameters used in the operation and stores them in the form of the maximum heap. Therefore, it takes the shortest time to operate. Table 3 confirms that the proposed algorithm can quickly provide the regional traffic state classification results.

### 5. Conclusion

Based on the lack of the present quantitative research of urban regional traffic status and the demand of the urban transportation system, this paper studies the intrinsic characteristics of regional traffic flow parameters in the network environment. From the perspective of complex network topology, this article abstracts the urban regional network into a directed and weighted complex network and regards it as study foundation. It introduces the global efficiency as quantitative parameter of regional road network traffic state. With complex network community division theory, traffic state level in the region has been obtained. Then, build the corresponding relationship between the global efficiency and the average travel speed of arterial road in regional network, which further determines the boundary value of each class of discriminant standard. Thus, this article can provide the final regional traffic state discriminant criteria.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

This research is funded by the National Natural Science Foundation of China (no. 51308249) and Shandong Province Tube Enterprises Technological Innovation Projects (no. 20122150251-1).

### References

[1] Y. Chen, Y. Zhang, J. Hu, and D. Yao, "Pattern discovering of regional traffic status with self-organizing maps," in *Proceedings of the IEEE Intelligent Transportation Systems Conference (ITSC '06)*, pp. 647–652, September 2006.

- [2] W. Guo, Y. D. Yao, Y. Fu et al., "Study on the method for regional traffic flow feature extraction and traffic status evaluation," *Journal of Highway and Transportation Research and Development*, vol. 22, no. 7, pp. 101–104, 2005.
- [3] F. X. Ji, L. Liu, and Q. G. Wu, "Extraction method for traffic information of regional road network," *Journal of Southwest Jiaotong University*, vol. 43, no. 3, pp. 22–26, 2008.
- [4] H. P. Leng, *Research on Urban Traffic Network State Evaluation*, Jilin University, Changchun, China, 2008.
- [5] C. Z. He, Q. Y. Zhou, and Z. Yu, "Regional traffic state evaluation method based on data visualization," *Journal of Traffic and Transportation Engineering*, vol. 16, no. 1, pp. 133–140, 2016.
- [6] H. S. Zhang, Y. Zhang, and D. C. Hu, "Quantitative approach for regional traffic state analysis," *Journal of Jilin University Engineering and Technology Edition*, vol. 39, no. 2, pp. 336–342, 2009.
- [7] C. F. Daganzo, "Urban gridlock: macroscopic modeling and mitigation approaches," *Transportation Research Part B: Methodological*, vol. 41, no. 1, pp. 49–62, 2007.
- [8] J. E. Gonzales, C. Chavis, Y. Li et al., "Multimodal transport modeling for Nairobi, Kenya: insights and recommendations with an evidence-based model," in *Proceeding of the Transportation Research Board 90th Annual Meeting*, D.C, Washington, DC, USA, 2011.
- [9] N. Geroliminis and C. F. Daganzo, "Existence of urban-scale macroscopic fundamental diagrams: some experimental findings," *Transportation Research Part B: Methodological*, vol. 42, no. 9, pp. 759–770, 2008.
- [10] J. J. Wu, *Studies on the Complexity of Topology Structure in the Urban Traffic Network*, Beijing Jiaotong University, Beijing, China, 2008.
- [11] S. Porta, P. Crucitti, and V. Latora, "The network analysis of urban streets: a primal approach," *Environment and Planning B: Planning and Design*, vol. 33, no. 5, pp. 705–725, 2006.
- [12] P. Y. Gu, J. H. Chen, D. W. Gong, and M. Zhou, "A generalized form of weighted network evolving model," *Mathematics in Practice and Theory*, vol. 43, no. 22, pp. 253–259, 2013.
- [13] C. F. Shao, *Traffic Planning*, China Railway Publishing House, Beijing, China, 2014.
- [14] B. F. Liu and Y. Gao, "The research of clock traffic congestion model and synchronization technology," *Applied Mechanics and Materials*, vol. 346, no. 346, pp. 145–148, 2013.
- [15] F. Q. Zhang, *Study on Forecasting Method of Urban Arterial Road Traffic Congestion*, Chang'an University, Xi'an, China, 2015.
- [16] Z. S. Guo and Z. M. Lu, *Complex Network Basic Theory*, 284, China Science Publishing & Media Lid. (CSPM), Beijing, China, 2012.
- [17] L. Tian, R. Z. Di, and H. Yao, "Effect of distribution of weight on the efficiency of weighted networks," *Acta Physica Sinica*, vol. 60, no. 2, pp. 307–351, 2011.
- [18] P. Zhang, J. F. Guo, Y. X. Quan et al., Traffic congestion evaluation method based on congestion index, CN103366557A, 2013.
- [19] F. X. Wang, X. Li, and R. G. Chen, *Complex Network Theory And Application*, Tsinghua University Press, Beijing, China, 2006.
- [20] X. Wu, V. Kumar, Q. J. Ross et al., "Top 10 algorithms in data mining[]," *Knowledge and Information Systems*, vol. 14, no. 1, pp. 1–37, 2008.
- [21] Z. Huang, "A fast clustering algorithm to cluster very large categorical data sets in data mining," *Research Issues on Data Mining & Knowledge Discovery*, pp. 1–8, 1998.
- [22] J. Vesanto, J. Himberg, E. Alhoniemi et al., "Self-organizing map in Matlab: the SOM toolbox," in *Proceedings of the Matlab DSP Conference*, pp. 35–40, 2000.
- [23] Z. Anke, Q. Xinjian, and C. Guojian, "Clustering analysis of gene data based on PCA and SOM neural networks," in *5th International Conference on Intelligent Systems Design and Engineering Applications, ISDEA 2014*, pp. 284–287, 2014.
- [24] S. Zhou, L. Fu, and B. Liang, "Clustering analysis of ancient celadon based on SOM neural network," *Science in China, Series E: Technological Sciences*, vol. 51, no. 7, pp. 999–1007, 2008.
- [25] S. Bandyopadhyay and U. Maulik, "An evolutionary technique based on K-Means algorithm for optimal clustering in, ja: math," *Information Sciences*, vol. 146, no. 1–4, pp. 221–237, 2002.



# Hindawi

Submit your manuscripts at  
<https://www.hindawi.com>

