Journal of Environmental Informatics 15(2) 111-119 (2010)

Journal of Environmental Informatics

www.iseis.org/jei

Urban Environment-Oriented Traffic Zoning Based on Spatial Cluster Analysis

N. N. Wu[1], X. P. Yan[1,*], G. H. Huang[2], C. Z. Wu[1], and J. Gong[1]

[1]*Engineering Research Center for Transportation Safety (Ministry of Education), Intelligent Transport Systems Research Center, Wuhan University of Technology，Wuhan, 430063, China*

[2]*Faculty of Engineering, University of Regina, Regina, Saskatchewan S4S 0A2, Canada*

Received 11 Januray 2010; revised 11 May 2010; accepted 17 June 2010; published online 20 June 2010

ABSTRACT. Environment-oriented traffic management zoning (TMZ) provides an important way to study spatial variation of urban traffic environment. The TMZ can be used to manage the urban traffic environment and establish regional improvement strategies. In this study, a spatial immune cluster method based on aiNet (Artificial Immune Network) is designed to facilitate the traffic environment zoning. Spatial coordinates are employed to measure the spatial correlation of the traffic system, and the weightings of different indices are calculated with the use of PCA (principle component analysis). For the case of Beijing, the whole city area is divided into evenly distributed grid cells. An urban traffic environment index system is developed to measure the traffic environment condition of different grids based on the concept of traffic sustainability. The GIS (Geographical information System) software of ArcGIS and database programming are used to calculate the indexical values of each grid. In order to classify the grid data set, we use two analysis methods, i.e. immune cluster algorithm and K-mean algorithm, and it is found that the immune cluster algorithm has better performance for the classification. With the use of this method, the study area is divided into 4 zones.

Keywords: traffic environment, management zones, spatial clustering, geographical information system, principal component analysis, K-mean, immune clustering

1. Introduction

In the process of urbanization, modern transportation provides a convenient way for traveling; however, it has negative impacts on urban sustainable development. For example, the problems of vehicle emission, noise hazards, traffic congestion and petroleum resource shortage, have becoming more and more serious in many countries. Sustainable traffic environment is developed not only for meeting the city's growing traffic demand, but also for coordinating the contradiction between the urban transportation needs, socio-economic development and the shortage of environmental resources.

In the last few years, many studies have been conducted for the promotion of the urban traffic environment, especially in the relationship between transportation activity and the environment (Wu et al., 2010). The results could be used to predict the pollution caused by transportation (Perez and Trier, 2001; Cai et al., 2008; Cristion et al., 2008; Xu et al., 2008).

Recently, GIS tool and a number of analysis techniques have been used to study the formation of the traffic environment and social problems. For example, Oyana (2009) employed SOM algorithm and GIS to study the relationship between environmental factors and the disease of obstructive sleep apnea; some studies have been developed to explore the spatial distribution of vehicle emission concentration and voice in a microscopic environment based on the GIS and vehicle emission model (Tan, 2003; Chen et al., 2004; Zhang et al., 2008).

In addition, some assessment and planning models have been proposed for the transport and environment system based on a certain index system. Typical examples of evaluation methods are fuzzy evaluation (Yang and Wang, 1999), gray theory (Huang et al., 2005), Analytic Hierarchy Process (AHP) (Zhang, 2009), and inexact algorithm (He et al., 2006; Wu et al., 2009).

Urban traffic environment system is a complex unity composed of urban socio-economic subsystem, traffic subsystem, resource and environment subsystem. The whole traffic environment system shows a high degree of spatial differences and correlation. However, most of the studies aimed at a single aspect (e.g., traffic emission and voice), or on the whole city level (comprehensive assessment of whole city) (Jiang et al., 2007).

Urban traffic environment zoning considers the variation and relationship of different parts of the urban transportation environment, and divides the whole city into regions of interzone homogeneity. It is the first step in regional control and planning of the urban traffic environment. Therefore, it is impor-

* Corresponding author. Tel.: +86 27 87651456; fax: +86 27 86582280.
E-mail address: xpyan@whut.edu.cn (X. P. Yan).

ISSN: 1726-2135 print/1684-8799 online
 doi:10.3808/jei.201000171

tant to find an approach to realize the urban traffic environment division to bring insight into more scientific, accurate and reasonable management and adjustment measure for the urban traffic environment system.

In order to achieve accurate partition of urban traffic environment zones, three problems should be solved. Firstly, it is significant to understand the meaning and function of the traffic environment, and establish the index system to quantize and record the traffic environment condition at different times and locations. Secondly, it is necessary to develop a research framework and object model. The traffic network and spatial grid model can be used, both of which can solve spatial problems in transportation. Traffic network model is common in the traffic analysis zone of traffic prediction, traffic regional signal control and navigation (Li et al., 2009; Xu, 2009). The spatial grid system is usually used in the domain of traffic impact analysis, urban land resources management, and so on (Wang and Luan, 2000). However, traffic network is incapable of dealing with area elements. The spatial grid model is adopted in this study, with the aid of GIS software. Finally, it is important to figure out an effective zoning algorithm. Currently, K-mean and fuzzy K-mean are the most extensively used methods (Ortega et al., 2002; Rodrigo and Oscar, 2007).

This study starts by building the zoning index system of the urban traffic environment and discussing the raw data collection and data processing in Beijing, China. Then an immune cluster algorithm is designed to divide all urban grids into different groups, and the zoning scheme of the study area is formed based on the clustering result and special considerations of traffic zone definition. Finally the effectiveness of immune cluster in the application of TMZ and the zoning result of Beijing are discussed.

2. Materials

2.1. The Study Area

The study was conducted in the urban and suburban area of Beijing, China (see Figure 1). The study area represents an important function as a major political, educational and cultural center in China. It is also the focus of economic development and international communication.

According to the third comprehensive survey of Beijing transportation in 2005, the total travel amount and average travel distance have increased by 26.9 and 16.25% respectively, compared to 2000 (Beijing Transportation Committee 2005). Beijing is facing a sharp increase in both the traffic demand and environment pressure.

To measure the spatial variation of urban traffic environment system, the study area is divided into evenly distributed grids. Each grid is set to 1 km × 1 km, so as to make sure that at least one road is contained in each grid.

2.2. Zoning Index System

Urban traffic system is a complex system composed by socio-economic, traffic, resource and environment subsystems. According to the views on traffic sustainability, urban traffic environment should not only meet the city's growing traffic demand, but also limit the negative impact on urban resources and the environment.

In order to carry out cities' transportation sustainability assessment, many researchers have studied the index structure of Chinese cities' transportation environment system. Guo and Guo (2010) gave a detailed list of quantitative indices in terms of environmental, resource, and social subsystems, which almost covers all factors related to city transportation environment system. Other researchers built a city ecological transportation assessment index system to measure city-level transportation (Dai et al., 2005; Wang, 2009).

Figure 1. The study area.

In this study, zoning method is designed on the basis of spatial clustering according to the distribution and spatial variation of grid properties. The selection of indices to calibrate each grid is a very important step. According to the structure of the urban traffic environment system, a zoning index system should consider the factors of transportation conditions socio-economy, and environment.

The basic responsibility of the urban traffic system is serving the movement of people and goods from one location to another. In order to achieve this goal, the urban road network of enough density and sufficient transportation infrastructure are needed (Jiang et al., 2007). Besides, high efficiency of traffic is required. Public transportation is an important measure to promote efficiency, and alleviate the pressure of the growing transportation demand. Since bus system is still the most important way of public transportation for most Chinese cities, bus organization and planning factors, such as bus network, routes， stations and the proportion of bus travel, are selected to measure the service level of public transportation (Dai et al., 2005). Additionally, traffic volume, speed, and passenger flow are most commonly used to measure the output and dynamic properties of transportation. They are important properties of road network. In grid level, traffic volume and passenger capacity is replaced by vehiclemiles of travel (VMT), and passenger-miles of travel. And average congestion level is used, for congestion increase the traffic emission and voice.

On the other hand, urban transportation is affected by the condition of socio-economy, especial the distribution of land use. They are the source of traffic demand, and affect the distribution of traffic (Wang et al., 2009). Different kind of land use raise different requirement on transportation. Densely populated areas need convenient traffic and rigid control on vehicle emission and traffic voice.

In order to evaluate the negative impact of transportation on environment, indices of traffic voice level and vehicle emission are used in this paper. And urban greening can lessen the effect of traffic voice and emission, which is also regarded as a significant index of urban traffic environment.

From the practical management perspective, spatial structure of the zones is also an important issue. In city road network model, network topology (connectivity of road section) is often used to measure the spatial correlation and continuity (Xu, 2009) of traffic. In this grid model of urban transportation, we use the coordinates of grid center to measure the spatial autocorrelation (Lieske and Bender, 2009) of traffic system. According to the above considerations, the urban traffic environment zoning index system is established (Table 1).

2.3. Data Collection

The required data include graphic data, and traffic parameters on various road sections. To effectively manage those data, ArcGIS 9.2 (by ESRI) is used to facilitate the data storing and the latter stage of data processing. Different kinds of graphic data are stored in various map layers, representing the residential districts, areas of road land use, road network of different grades, bus routs, bus stations, road intersections, and green areas. The daily average traffic parameters are collected and stored in the relational database, and the traffic parameters of each road are related with the corresponding graphic data. Average traffic volume, vehicle type proportions, speed, congestion level, road-side decibel (voice) on various roads, are determined from the statistical data during the year of 2004 ~ 2005. The average emissions per km of HC, CO and NOx on each road are calculated on the basis of traffic volume, vehicle type information, and average speed, with the aid of microscopic vehicle emission model of CEME (Wu et al., 2008; Xu et al., 2008;).

2.4. Grid Indexical Value Calculation

In this section, the raw data is translated into grid properties. The indexical values of various grids are calculated with the help of ArcGIS tools and database programming.

The Analysis Tools of Intersect is used to cut a large graph into separate grids, because a graph is possibly shared by several adjacent grids. And the geometric property of each graph, including the X and Y coordinates of centroid, length, and area, are calculated by the Field Calculator (Tang and Yang, 2006). Indexical values of $n_1 \sim n_8$ and n_{15} in Table 1 are calculated by the aforementioned tools. The value of n_9 (density of service coverage of bus stations) is obtained by first building a buffering zone (a distance of 300 m is adopted (GB 50220-95)) around each bus station, and then summing up the service area in each grid.

Table 1. Urban Traffic Environment Zoning Index System

Category	n_i	Index
Geographical position of grid center	n_1	X coordinate (m)
	n_2	Y coordinate (m)
Social and economical indices	n_3	Density of residential districts (%)
	n_4	Density of traffic land use (%)
Transportation facility indices	n_5	Density of road length (km.km^{-2})
	n_6	Main road length percentage (%)
	n_7	Density of bus route length (km.km^{-2})
	n_8	Density of intersections (n.km^{-2})*
	n_9	Density of service coverage of bus stops (%)
Dynamic traffic parameters	n_{10}	Vehicle-miles of travel (VMT) (km.d^{-1})
	n_{11}	Bus VMT rate (%)
	n_{12}	Passenger-miles of travel (km.d^{-1})
	n_{13}	Average speed (km.h^{-1})
	n_{14}	Congestion level
Environmental indices	n_{15}	Density of Greening area (%)
	n_{16}	Roadside voice level
	n_{17}	NOx emission (kg.km^{-2})
	n_{18}	HC emission (kg.km^{-2})
	n_{19}	CO emission (kg.km^{-2})

* n = the number of interactions.

Other indexical values are calculated on the basis of the road network data and traffic parameters on various road sections. The database programming of Visual C# is used into this step. Equations (1) to (7) are programmed into the programmer code, to calculate the grid properties of index $n_{10} \sim n_{14}$, and $n_{16} \sim n_{19}$.

Vehicle-miles of travel (VMT):

$$VM_i = \sum_j l_{ij} \cdot f_{ij} \tag{1}$$

where i stand for each grid, VM_i represents the value of vehicle-miles of travel (VMT), j stands for the road segment contained in the grid i, f_{ij} is the average traffic volume, and l_{ij} is the length of the road section.

Bus VMT rate:

$$\mathrm{BR}_i = \sum_j f_{ij} . \mathrm{b}_{ij} . l_{ij} \div \sum_j f_{ij} . l_{ij} \tag{2}$$

where BR_i represents the value of bus VMT rate of grid i; b_{ij} is the vehicle type proportion of bus on each road section.

Passenger-miles of travel:

$$PM_i = \sum_j \sum_k f_{ij}.r_{ijk}.p_k.l_{ij} \quad (3)$$

where PM_i represents the value of passenger-miles of travel in grid i, k stands for vehicle type, p_k is the average passenger loading quantity of vehicle type k (averagely, 2 for car and taxi, 60 for bus), r_{ijk} is the average vehicle type proportion of vehicle type k on each road section.

Average speed:

$$S_i = \sum_j f_{ij}.l_{ij} \div \sum_j f_{ij}.\frac{l_{ij}}{v_{ij}} \quad (4)$$

where S_i represents the value of average speed of grid i, v_{ij} is average speed on each road section.

Congestion level:

$$CL_i = \sum_j l_{ij}.c_{ij} \div \sum_j l_{ij} \quad (5)$$

where CL_i represents the value of congestion level of grid i, c_{ij} is the average congestion level on each road section (typically, 1 represent smoothing traffic, 2 represent slow traffic, 3 represent seriously congested traffic).

Roadside voice level:

$$VOI_i = \sum_j voi_{ij}.l_{ij} \div \sum_j l_{ij} \quad (6)$$

where VOI_i represents the value of roadside voice level of grid i, voi_{ij} is the average roadside voice level on each road section (typically, 1 means $Leq \leq 68$, 2 means $68 < Leq \leq 70$, 3 means $70 < Leq \leq 72$, 4 means $72 < Leq \leq 74$, 5 means $Leq > 74$).

NOx/HC/CO emissions:

$$E_i = \sum_j e_{ij}.l_{ij} \quad (7)$$

where E_i represents the value of NOx/HC/CO emissions of grid i, e_{ij} is the average vehicle emission amount of NOx/HC/CO per km on the road section (depending on factors such as vehicular speed and volume).

3. Methodology

The objective of using cluster algorithm is to find various groups, in the way that the objects belonging to the same group resemble each other, whereas the objects in different groups are dissimilar. In this study, the grids calibrated by the zoning indexical values are divided into several groups through spatial clustering. Management zones are formed on the basis of clustering results. Cluster algorithms are differentiated from each other in the following aspects: (1) definition of the distance function between two objects; (2) method to find the groups. In this study, the immune clustering enlightened by artificial immune system was proposed to classify the grids into different groups. And it is compared to K-mean clustering.

3.1. K-mean Algorithm

The K-mean analysis is performed by using the statistical analysis software of SPSS 13.0 for Windows (SPSS Inc., 2004; Yang, 2005; Xu, 2007). The Euclidean distance is used to measure the distance between grid sets. Before the cluster analysis, all indexical values for each grid are standardized according to Equation (8):

$$SI_{ik} = (I_{ik} - \bar{I}_k)/S_k \quad (8)$$

where SI_{ik} is the standardized value of the k^{th} index for the grid i; I_{ik} is the original indexical value; S_k is the standard deviation of the original indexical value k.

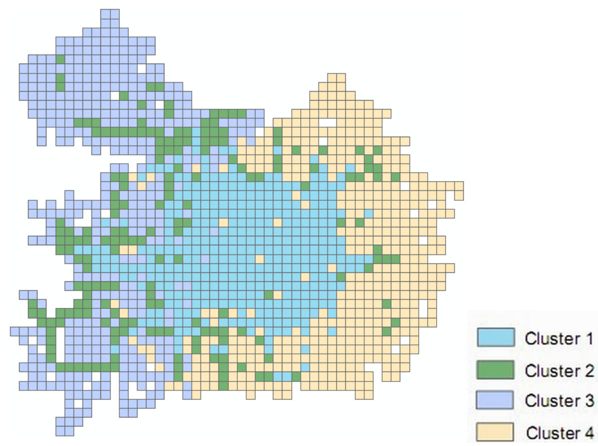

Figure 2. K-mean clustering result.

In SPSS, settings for the K-mean analysis are as follows: maximum iteration = 100, convergence criterion = 0.00001. The pre-defined cluster number is demanded. Determining an exact cluster number is a complex and difficult problem. Current solutions in theoretical studies are often dependent on the characteristic of clustering algorithm adopted (Cheng et al., 2007). In practical application, the background knowledge and expert opinion are usually incorporated to figure out a reasonable cluster number. In urban traffic environment zoning, it is very important to consider the convenience of regulation of urban traffic system. Too many zones will increase the difficulty of regional management. Referring to the practice of Shanghai transportation planning (Yi, 2006) and the suggestion of Beijing Traffic Management Bureau, we divide Beijing into 4 zones. The K-mean analysis result is shown in Figure 2.

3.2. Immune Clustering

During the last decade, artificial immune systems have attracted a lot of attention and are extensively used to solve various engineering problems. It is enlightened by the mechanism of our bodies against the constant attack from external antigens, such as strains and pathogens. It functions by a set of cells, named antibodies, which continuously interact with each other and the antigens (Yang 2005; Merve and Ahmet, 2009). aiNet (Artificial Immune Network) was proposed by de Castro and Von Zuben (2001). aiNet algorithm is a kind of immune network learning algorithm. Though several iterations of clonal selection and antibody variation, the final memory antibodies can memorize the mode of antigens (Xu, 2007).The aiNet algorithm has the characteristics of fast convergence and global optimization. But the data of memory antibodies are not the clustering centers of the original grid data. The graph or traditional cluster algorithm is needed to obtain the ultimate cluster centers. In this section, several definitions are explained and a cluster algorithm based on aiNet system is built.

3.2.1. Index Weighting

To overcome the subjectivity of expert determining weighting values of different indices, an objective comprehensive method to determine index weighting is used, which is based on the principal component analysis (Pang et al., 2001).

In this study, principal components analysis (PCA) is carried out based on the grid sample. The statistical analysis software of SPSS is used in the former processing. Pre-processing in SPSS includes target standardization, calculation of correlation matrix, eigenvalues and proportions of variance, and initial loading of variables for PCs. Next the principle components are selected, in the way that the cumulative proportion is higher than 80%. The principal components expression is computed as follows:

$$\begin{cases} F_1 = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p \\ F_2 = a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p \\ \dots\dots\dots\dots \\ F_k = a_{k1}X_1 + a_{k2}X_2 + \dots + a_{kp}X_p \end{cases} \quad (9)$$

$$a_{ij} = \mu_{ij} / \sqrt{\lambda_i} \quad (10)$$

where μ_{ij} is the initial loading of the variance *j* for the PC$_i$; λ_i is eigenvalue of PC$_i$.

The comprehensive score model based on PCA is calculated, in Equations (11) and (12):

$$Y = b_1X_1 + b_2X_2 + \dots + b_pX_p \quad (11)$$

$$b_i = \sum_{j=1}^{k} a_{ji}\theta_j \div \sum_{j=1}^{k}\theta_j \quad (12)$$

where θ_j is the explaining proportion of the j^{th} PC; b_i represents the initial weighting of the i^{th} variance. After the normalization of the initial weighting, the final weightings of all variable are acquired (see Table 2).

Table 2. Weighting for Grid Indices

Indices	n_i	Weight
X coordinate	n_1	0.116
Y coordinate	n_2	0.078
Density of residential district	n_3	0.047
Density of traffic land use	n_4	0.044
Density of road length	n_5	0.045
Main road length percentage	n_6	0.07
Density of bus route length	n_7	0.046
Density of intersects	n_8	0.044
Density of service coverage of bus stops	n_9	0.032
vehicle-miles of travel (VMT)	n_{10}	0.057
Bus VMT rate	n_{11}	0.039
Passenger-miles of travel	n_{12}	0.052
Average speed	n_{13}	0.026
Congestion level	n_{14}	0.058
greening rate	n_{15}	0.01
Roadside voice level	n_{16}	0.067
NOx emission	n_{17}	0.056
HC emission	n_{18}	0.056
CO emission	n_{19}	0.056

3.2.2. Affinity

In this study, the standardized grid data set (in Equation (7)) are recognized as the population of intrusive antigens. Suppose that *Ag* is the population of antigens, and that *Ab* is the network antigens:

$$Ag = (Ag_1, Ag_2, ..., Ag_Q),\ i = 0, 1, ..., Q. \quad (13)$$

$$Ab = (Ab_1, Ab_2, ..., Ab_Q),\ i = 0, 1, ..., P \quad (14)$$

where Ag_i is the i^{th} antigen, Ab_i the i^{th} antibody in the immune network, *Q/P* is the total amount of grids/antigens.

The affinity is the quantitative interaction between the antibodies and antigens. A larger value means a bigger opportunity for the antibody to be selected and cloned. This is defined in Equation (15):

$$Fin_{ij} = \sqrt{\sum_{k=1}^{m} W_k (Ag_{ik} - Ab_{jk})^2} \quad (15)$$

where Fin_{ij} is finity between an antigen and an antibody, Ag_{ik} is m^{th} property of the antigen Ag_i, Ab_{ik} is m^{th} property of the antigen Ab_i, and W_k is the weighting of the k^{th} property for the grid dataset.

3.2.3 Similarity

Similarity is an important quantitative index for the interaction between different antibodies. Antibodies with a large similarity will be suppressed to keep the diversity of immune system. This mechanism prevents the clustering algorithm from prematurity:

$$Sim_{ij} = 1 \Big/ (1 + \sqrt{\sum_{k=1}^{m} W_k (Ab_{ik} - Ab_{jk})^2}) \quad (16)$$

where Sim_{ij} is the similarity between Ab_i and Ab_j.

3.2.4 Algorithm

Procedures in designed immune clustering algorithm are:

(1) Initialization: Create an initial random population of antigens.

(2) Immune mechanism: For each antigen Ag_i, do (i) to (iv).

(i) Clonal selection and expansion -- For each network anti- body, determine its affinity with the antigen Ag_i (Equation (15)). Select a certain number of antibodies of high affinity, and clone them to get new population of antibodies $A_b(K_1)$. The cloning number for each selected antibody is defined in the Equation (17).

$$q_j = \text{int}[N_c \cdot Fin_{ij} \Big/ \sum_{i=1}^{e} Fin_{ij}] \quad (17)$$

where N_c is the total cloning number; q_j is the cloning number of j^{th} antibody.

(ii) Antibody variation -- Mutate each antibody according to Equations (18), (19), and select a fixed percentage of antibodies to form a new population of antibodies $A_b(K_2)$:

$$Ab_j = Ab_j - \mu\left(Ab_j - Ag_i\right) \quad (18)$$

$$\mu = \alpha \times \exp(-Fin_{ij} / \beta) \quad (19)$$

where μ is a variation indicator for the antibody; fin_{ij} is the affinity between Ag_i and Ab_j, $\alpha, \beta \in (0,1)$

(iii) Antibody suppression -- Determine the similarity between the antibody set $A_b(K_2)$ and eliminate those antibodies whose affinity with other antibody are smaller than σ_s.

(iv) Memory cells collection -- Add the remaining antibodies in to the memory set $M(k)$.

(3) Memory set suppression: Eliminate those memory antibodies whose affinities are smaller than σ_d.

(4) Cycle: Repeat steps (2) to (3), until a pre-specified number of iterations is reached.

(5) Clustering center extraction: Use the AHP (Analysis Hierarchy Process) clustering analysis on the final memory set and calculate the average values in each cluster as the clustering centers (Liu, 2008).

(6) Classification: All grids are classified into different groups in a way that each grid belongs to the cluster, center of which are nearest to it.

The result of immune clustering is displayed in Figure 3.

4. Results and Discussion

4.1. Comparison between Two Clustering Results

Although the two methods show similarities on cluster 3 and cluster 4, they are different from each other on the location and distribution of clusters 1 and 2. In the K-mean analysis, the distribution of grids in the cluster 2 is so irregular and dispersive that visually no zone can be extracted from this cluster.

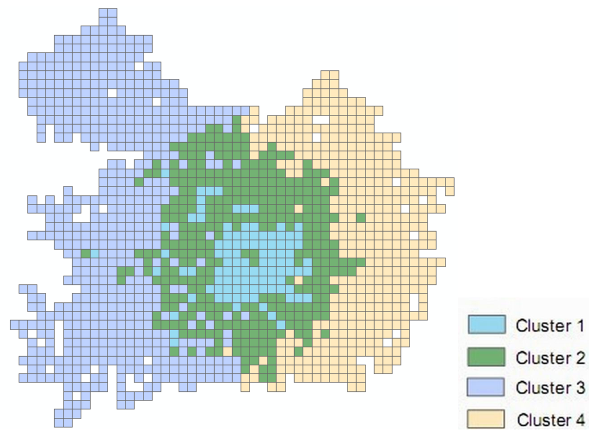

Figure 3. Immune clustering result.

The difference is produced by at least two reasons: 1) complexity and peculiarity of traffic environment zoning, and 2) the distinction of the two algorithms.

The performance of urban transportation is affected by many complex factors. In this study, 19 indices are selected to measure the variation of transportation environmental condition, and some of those indices indicate great variability over the space and time. In this sense, the irregular distribution of grid properties may exist. Those abnormal data present great challenges for the cluster analysis. Competent cluster analysis should satisfy the following requirements: 1) effectiveness in handle with high-dimensional data, and 2) stability when abnormal data is present.

The two clustering algorithms are different from each other. The immune clustering mentioned is composed of two steps, the mode diagnosis depending on the artificial immune network (aiNet), and an additional cluster analysis to classify the input data. Through the first step, data of smaller scale are obtained to represent the original mass data, and the effect of abnormal data is reduced (Xu, 2007). K-mean analysis is tested by many researches to be capable of deal with large samples. Whereas it works under objective function concerning the original samples, in order to achieve the smallest variation in clusters. Therefore, the K-mean analysis is sensitive to abnormal data. In this case, the cluster 2 of K-mean algorithm is possibly resulted from those abnormal data. A qualified cluster for traffic zoning should make sure that the grids in one cluster are not only similar in properties but also distributed intensively. Only in this way, a zone can be got from each cluster.

In conclusion, although K-mean algorithm is very com-

Table 3. Average Indexical Values for Each Zone

Zoning Indices		1	2	3	4
X coordinate	n_1	0.56	0.51	0.27	0.79
Y coordinate	n_2	0.42	0.44	0.50	0.45
Density of residential district	n_3	0.52	0.34	0.09	0.12
Density of traffic land use	n_4	0.71	0.40	0.13	0.08
Density of road length	n_5	0.49	0.27	0.11	0.11
Main road length percentage	n_6	0.14	0.12	0.01	0.04
Density of bus route length	n_7	0.45	0.21	0.02	0.02
Density of intersects	n_8	0.47	0.31	0.05	0.04
Density of service coverage of bus stops	n_9	0.35	0.16	0.01	0.01
vehicle-miles of travel (VMT)	n_{10}	0.63	0.32	0.11	0.11
Bus VMT rate	n_{11}	0.62	0.60	0.38	0.42
Passenger-miles of travel	n_{12}	0.56	0.25	0.03	0.03
Average speed	n_{13}	0.73	0.81	0.85	0.86
Congestion level	n_{14}	0.13	0.07	0.00	0.00
greening rate	n_{15}	0.07	0.08	0.06	0.05
Roadside voice level	n_{16}	0.05	0.06	0.00	0.02
NOx emission	n_{17}	0.61	0.30	0.13	0.10
HC emission	n_{18}	0.60	0.29	0.10	0.10
CO emission	n_{19}	0.59	0.29	0.10	0.10

monly used in the management zones, the immune cluster analysis is more effective in the urban environment-oriented traffic zoning. It is extrapolated that K-mean may not be competent in such situations that the considered properties are complex and diverse.

4.2. Management Zone Formation

During years of study of management zone design, many scientific literatures have established some guidelines to the definition of traffic zones. For the management of urban traffic environment zoning it requires: 1) homogeneity of property, 2) compactness of zone shapes, 3) adjustment of zone boundaries to political and administrative boundaries, and 4) respect of physical separators (Ding et al., 1993; Ortuzar and Willumsen, 2001). The first two requirements are satisfied through the spatial clustering.

As shown in Figure 3, the boundaries between zones are very illegible. After considering the current administrative boundary and rivers, a final zoning scheme is achieved (see Figure 4). The average index values are display in Table 3.

Zone 1 consists of Xicheng District, Dongcheng District, Chongwen District, and Xuanwu District. The population density is very large. Up to 2006, total household population of this area was 399.8 million, accounting for 49.4% of Beijing's eight central administrative districts. The number of private car ownership is very large, and traffic speed is very small. The living environment is very bad. The traffic emission is very serious and this is caused by traffic congestion and instability. Zones 3 and 4 are located in the outside part. In those zones, the transport infrastructure input is much smaller. And there are service blind spots of bus transportation. The vehicle emissions are less, and the living environment is very suitable in those zones.

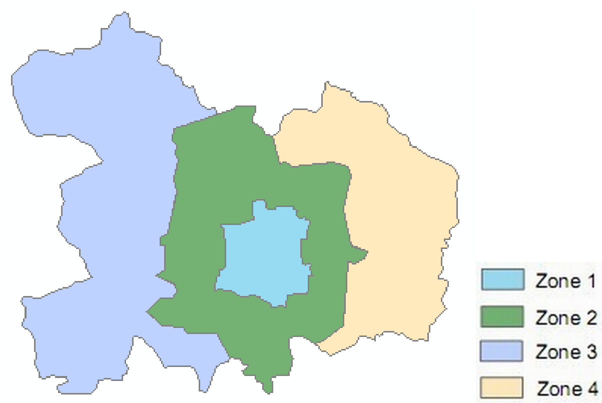

Figure 4. Zoning scheme.

5. Conclusions

For the control and management of regional traffic environment, urban traffic environment zoning is very important. How to classify traffic management zones has been attracting more and more concerns, it is known that the common and effective way is using spatial cluster and GIS. On the basis of the framework of spatial grid model, spatial clustering, which integrates spatial coordinates, socio-economic, traffic, and environmental properties of grid cells, could effectively accomplish the traffic environment zoning. As such, both of similarity and spatial closeness of the grids within each zone are well kept.

The accuracy of the definition of management zone (MZ) depends on the representation of index system, and the effectiveness of cluster algorithm. In the study, an index system consisting of 19 indices is designed for urban grids, and an immune cluster algorithm is applied in the case of Beijing. In comparison to the K-mean analysis, the immune algorithm presents better performance in dealing with high-dimensional data, and robustness against abnormal samples. It is concluded that the immune cluster algorithm is more effective in the application of urban traffic environmental zoning. Additionally, the special considerations of traffic zone definition can increase the feasibility of traffic environment zones in practice.

Acknowledgments. The authors acknowledge the support of the National Basic Research Program of China (No. 2005CB724205). The authors also appreciate Dr. Nengchao Lv's cooperation and also special thanks are given to the editor and other two anonymous reviewers for their helpful comments on an earlier manuscript.

References

Cai, M., Yin, Y.F., and Xie M. (2008). Prediction of hourly air pollutant concentration near urban arterials using artificial neural network approach. *Transp. Res. D: Transp. Environ.,* 14 (1), 32-41, doi:10.1016/j.trd.2008.10.004

Chang, K.T., Khatib, Z., and Ou, Y.M. (2002). Effects of zoning structure and network detail on traffic demand modeling. *Environ. Plann. B,* 29(1), 37-52, doi:10.1068/b2742.

Chen, H.M., Wang, X.Q., and Chen, C.C. (2005). GIS-based urban atmospheric dispersion model: a case study of Fuzhou. *Geo-information Science,* 7(4), 101-106.

Cheng, L.F., Jiang, Q.S., and Wang S.R. (2007). A hierarchical method for determining the number of clusters. *Journal of Software,* 19(1), 62-72.

Cristion, E., Luis, S.V., and Roberto, M.C. (2008). A simulation platform for computing energy consumption and emissions in transportation networks. *Transp. Res. D: Transp. Environ.,* 13 (7), 412 -427, doi:10.1016/j.trd.2008.07.006.

Dai, Y., Chen, C.H., and Jing, Q.G. (2005). Indicators establishment and evaluation approach for sustainable transportation. *World Sci-Tech R & D,* 27(5), 94-99.

de Castro, L. N., and Von Zuben, F. J. (2001). *aiNet: An artificial immune network for data analysis*, Idea Group Publishing, USA.

Ding, C., Choi, K., and Kim, T.J. (1993). GIS-based traffic analysis zone design. *Proc. of the 3rd International Conference on Computers in Urban Planning and Urban Management,* Atlanta, Georgia, 1993.

Guo, X.B., and Guo, J.R. (2010). Discussion of the sustainable development of city transportation. *Shanxi Arichitecture,* 36(4), 30-31, doi:100926825(2010)1020030202.

He, L., Chan, C.W., Huang, G.H. and Zeng, G.H. (2006). A probabilistic reasoning-based decision support system for selection of remediation technologies for petroleum-contaminated sites. *Expert Syst. Appl.,* 30, 783-795, doi:10.10 16/j.eswa.2005.07.038.

Huang, H.L., Zeng, G.M., Huang, G.H., Yu, H.Y., Zhang, B.B. and Li, J.B. (2005). Application of grey clustering method on evaluating compost maturity. *J. Safety Environ.,* 6(5), 87-90, doi:1009-6094 (2005) 06-0087-04.

Jiang, Y.M., Guo, H.C., Huang, K., Yu, Y.J., and Liu, Y. (2007). Comprehensive assessment model and its application for urban eco-traffic system. *Research of Environmental Science,* 20(6), 158-163, doi: 1001 -6929(2007)06-0158-06.

Li, X.D., Chu, H., and Yang X.G. (2009). The concept of city traffic network zone. *Joural of Wuhan University of Technology,* 33(5), 972-975, doi: 10.3963/j.issn.100622823.2009.05.040.

Lieske, D.J., and Bender, D.J. (2009). Accounting for the influence of geographic location and spatial autocorrelation in environmental models: a comparative analysis using north American songbirds. *Environ. Inf.,* 1391), 12-32, doi:10.3808/jei.2009001 37.

Liu, Q. (2008). The application of artificial immune clustering algorithm in automatic division of Traffic Intervals. *Application,* 2, 64-67, doi: 1003-0492(2008)01-0064-04.

Merve, A., and Ahmet, A. (2009). A collaborative filtering method based on artificial immune network. *Expert Syst. Appl.,* 36, 8324-8332, doi:10.1016/j.eswa.2008.10.029.

Ortega, J., Foster, W., Ortega, R., (2002). Defini´on de sub-rodales para una silvicultura de precisi´on: Una aplicaci´on del m´etodo Fuzzy K-means. *Ciencia e Investigaci´on Agraria*, 29 (1), 35 -44.

Ortuzar, J., and Willumsen, L. (2001). *Modeling transport,* 3rd ed., Wiley, New York.

Oyana, T.J. (2009). Visualization of High-Dimensional Clinically Acquired Geographic Data Using the Self-Organizing Maps. J. *Environ. Inf.,* 13(1), 33-44, doi: 10.3808/jei.200900138.

Pang, Y.J., Liu, K.D., and Zhang, B.W. (2001). The method of determining the objectivity index weight in the synthetic evaluation system. The *Concept and Application of System Engineering*, 8, 37-42, doi: 100026788(2001)0820037206.

Perez, P., and Trier, A., (2001). Prediction of NO and NO_2 concentrations near a street with heavy traffic in Santiago, Chile. *Atmos. Environ.,* 35, 1783-1789, doi:10.1016/S1352-2310(00) 00288-0.

Rodrigo, A.O., and Oscar, A.S, (2007). Determination of management zones in corn (Zea mays L.) based on soil fertility. *Comput. Electron. Agric.,* 58, 49-59, doi:10.1016/j.compag.2006.12.011.

Tan, Z.Y. (2003). *GIS-based automobile Emission's pollution information system of Shantou,* M.A. Dissertation, college of Engineering, Shantou University, Shantou, Guangdong, China.

Tang, G.A., and Yang, X. 2006. *Spatial analysis experiment tutorial for geographic information system of ArcGIS*, Science Press, Beijing.

Wang, M.J., and Luan, W.X. (2000). Regional adjustment reseach for the Yellow Sea coastal area in china. *Marine Science Bulletin,* 19(6), 50-56.

Wang, Z. (2009). *Study on comprehensive assessment method for urban eco-traffic system,* M.A. Dissertation, School of Transportation, Hefei University of Technology, Hefei, Anhui, China.

Wang, Z., Guo, H.C., Yu, Y.J., Xu, Z.X., and Zhan, X.H. (2009). Research progress in the interrelationship between urban land use and transport. *Human Geography*, 180(40), 91-97.

Wu, C.Z, Huang, G.H., and Yan, X.P. (2008). An intelligent agent mobile emissions model for urban environmental management. *Int. J. Software Eng. Knowl. Eng.*, 18(4), 485-502, doi:10.1142/S0218 194008003751.

Wu, C.Z, Huang, G.H., Yan, X.P., Cai, Y.P., and Li, Y.P. (2009). An inexact optimization model for evacuation planning. *Kybernetes,* 38(10), 1676-1683, doi: 10.1108/03684920910994 033.

Wu, C.Z., Huang, G.H., Yan, X.P., Cai, Y.P., and Li, Y.P. (2010). An interval-parameter mixed integer multi-objective programming for environment-oriented evacuation management. *Int. J. Syst. Sci.,* 41(5), 547-560, doi: 10.1080/00207720903072332.

Xu, C.G.. (2007). *The research of artificial immune system and application in the clustering,* M.A. Dissertation, School of Computer Science, South China Normal University, Guangzhou, Guangdong, China.

Xu, C.W., Wu, C.Z, and Chu, X.M. (2008). The research of average emission factor of light car in Wuhan based on CEME model. *Computer and Communications,* 26(4), 185-188.

Xu, L.Q. (2009). Tree growing algorithm for road network zoning. Application Research of Computers, 26(10), 3663-3665, doi: 10. 3969/ jissn.100123695.2009.10.017.

Yang, X.B. (2005). *Research of key technologies in clustering* *analysis,* Ph.D. Dissertation, School of Computer Science, Zhejiang University, Hangzhou, Zhejiang, China.

Yang, Z.S., and Wang, M. (1999). The grey statistical evaluation for city traffic systems. *Journal of Highway and Transportation Research and Development*, 16(2), 49-52.

Yi, J. (2006). *The research of the reasonable design of traffic analysis zone*, M.A. Dissertation, School of Architecture and Urban Planning, Tongji university, Tongji, Shanghai, China.

Zhang, C., Gong, J., and Yan, X.P. (2008). The dispersion of traffic emission based on GIS. *Transport Information and Safety,* 146 (27), 113-116.

Zhang, W.H. (2009). *Study on comprehensive assessment method for urban eco-traffic system,* M.A. Dissertation, School of Transportation, Hefei University of Technology, Hefei, Anhui, China.