I.J. Engineering and Manufacturing, **2012,3, 21-29**
Published Online June 2012 in MECS (http://www.mecs-press.net)
DOI: 10.5815/ijem.2012.03.04

Available online at http://www.mecs-press.net/ijem

Road Network Pattern Classification Using GEV Distribution Parameters

Chao Yang, Qi Liu

Key Laboratory of Road and Traffic Engineering of the Ministry of Education, School of Transportation Engineering Tongji University 4800 Cao'an Road, Shanghai, 201804, P. R. of China

Abstract

Network pattern mentioned in this paper is referred to the geographical layout and structure of a network, which is related to the connection, direction, and combination features of roads in a road network. A quantitative method is proposed in this paper to classify patterns of networks, through which a network could be identified quantitatively to be one of the three standard patterns, i.e., grid network, circle+radial network and tree-patterned network. Metric distances of shortest paths are taken as the main features of the networks and are described by parameters through Generalized Extreme Value (GEV) fitting. The criteria for pattern classification were established according to the cluster analysis of the parameters calculated from a set of trial networks. Six real networks were calculated using the method and their patterns are identified according to the proposed criteria. It turned out that the method could capture the features of the network patterns well. This method may set a threshold of the more general and deep studies of network pattern classification, which may offer help to road network planning and assessment.

Index Terms: Road network pattern; network topology; classification; GEV distribution

1. Introduction

One may tell the name of a city immediately only by taking a look at its road network because the network of a city has its own pattern, which could also be named as morphology. As is concerned here, network pattern is referred to the geographical layout and structure of a network, which is related to the connection, direction, and combination of roads in a road network.

Road networks of cities get their own patterns through years of evolution. Looking into the history of urban evolution, it is obvious of the interaction between urban land expanding and road network growing, which is also mentioned as a chicken and egg problem, i.e., the growth of traffic network depends on urban land use pattern, and nevertheless, has obvious influences on urban land use in return. The circular influences

* Corresponding author.
E-mail address: tongjiyc@tongji.edu.cn , deriekliu@163.com

can be shown more clearly in Fig. 1.

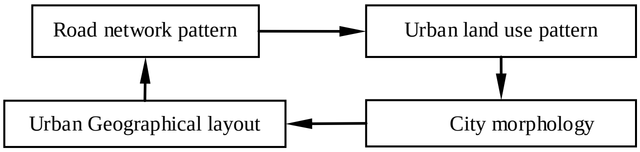

Fig. 1 Interactions between road network pattern and urban land use.

Different patterns of road networks may have different influences on both urban land use evolution and network performances. In the early years of the 20th century, planners and architects put forward a method by which strictly hierarchical roads are organized into a tree patterned network because they thought a network in that pattern could be more efficient and safe for the vehicles [1]. According to this theory, ITE has published Recommended Practice for Subdivision Streets [2], in order to improve the traffic safety of community streets through hierarchical and unconnected network layout.

Recently, as accessibility of multi-model traffic become more and more concerned, traffic planners start to consider increasing the capacity of a whole network by strengthen its density and connectivity [2], and highly connected pattern becomes a favorable one in network planning [3].

In spite of the development that has been gained in the network pattern study, the insight into network patterns is still not clear and only a few researchers have worked on it. Sun and Lovegrove have studied four patterns of networks and compared their levels of traffic safety [4]. Sun measured urban rail transit network pattern using fractal approach and described the interactions and morphological similarities between urban land use morphology and urban transport network pattern [5].

One of the very important problems during the study of network pattern is that there is not yet a standard for pattern classification which is agreed by most researchers. Furthermore, there is not even an effective method to distinguish between types of patterns. Take Sun's research [4] for instance, the four types of network were just put out by the author, thus may be difficult for other researches or cities to follow.

To solve this kind of problems, this paper proposes a quantitative method to identify different types of network patterns. This method could capture the main features of a network's pattern and it proved out to be able to distinguish different types of network patterns.

The following section will talk about models used to describe networks. In the third section, the network pattern classification method and related criteria will be proposed using several trial network samples. In the forth section, actual road networks will be calculated in order to validate the method. Conclusions will be stated in the final section.

2. Network Description Using Gev Model

In this paper, network is abstracted as a planar graph: G= {V, E}, where V is a collection of nodes connected by links (edges) E. Nodes represent intersections while links represent road segments between intersections. Adjacent nodes are defined as the nodes that connect with each other without passing through other nodes. The adjacent relations can be represented by the topological adjacency matrix {aij}:

$$a_{ij} = \begin{cases} 1, & \text{if there is an edge joining node i to node j;} \\ \infty, & \text{otherwise;} \end{cases} \quad (1)$$

and can also be represented by the metric adjacency matrix $\{b_{ij}\}$:

$$b_{ij} = \begin{cases} d_{ij}, & \text{if there is an edge joining node i to node j;} \\ \infty, & \text{otherwise;} \end{cases} \quad (2)$$

where dij=metric distance of the edge between node i and j.

Shortest paths between each pairs of nodes are figured out using matrix {aij} and {bij}. By using {aij}, the shortest path is the one that contains the least number of edges and the distance is the number of edges within the path. By using {bij}, the shortest path is the one that has the shortest metric distance, i.e., the sum of metric distances of the edges in the path. In this paper, shortest paths between nodes in a network are studied to capture the features of networks having different types of patterns.

Yang and Wang have studied morphological properties and functions of urban transit networks and proved that GEV distribution fits the network shortest paths very well [6]. Generalized extreme value (GEV) distribution is used to describe networks and capture their features. The GEV distribution function is:

$$y = f\left(x \middle| k, \mu, \sigma\right) = \left(\frac{1}{\sigma}\right) \exp\left(-\left(1 + k\frac{(x-\mu)}{\sigma}\right)^{-\frac{1}{k}}\right)\left(1 + k\frac{(x-\mu)}{\sigma}\right) - 1 - \frac{1}{k} \tag{3}$$

for $1 + k\frac{(x-\mu)}{\sigma} > 0$. As is shown in Fig. 2, μ∈R is the location parameter and the larger μ is for a distribution, the righter the distribution is positioned. σ > 0 is the scale parameter and the larger σ is, the gently the curve is, which indicates a smaller variation. ∈R is the shape parameter, which governs the tail characteristics of a distribution. =0 leads to the Type I distributions whose tails decrease exponentially, such as the normal distribution; >0 leads to the Type II distributions whose tails decrease as a polynomial, such as Student's t distribution; <0 leads to the Type III distributions whose tails are finite, such as the beta distribution.

As an example of description, road networks from four cities were calculated, and the probability distribution of their shortest path distances were fitted using GEV distribution. In spite of different network patterns, GEV distribution fits the networks very well, as shown in Fig. 3.

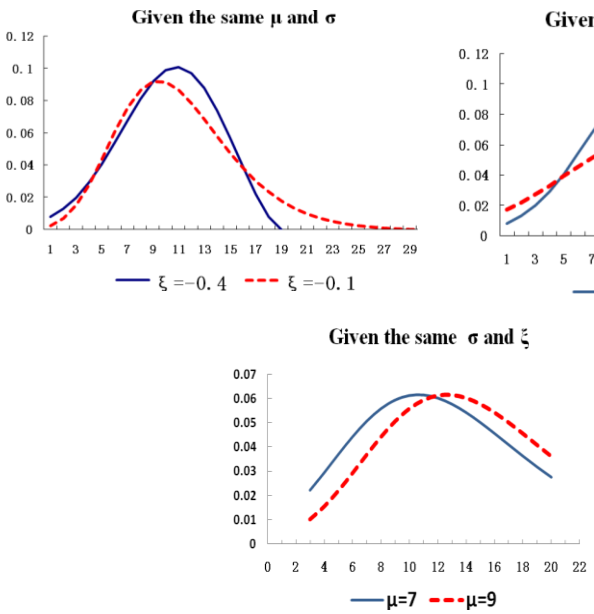

Fig. 2 Functions of three parameters in the generalized extreme value distribution.

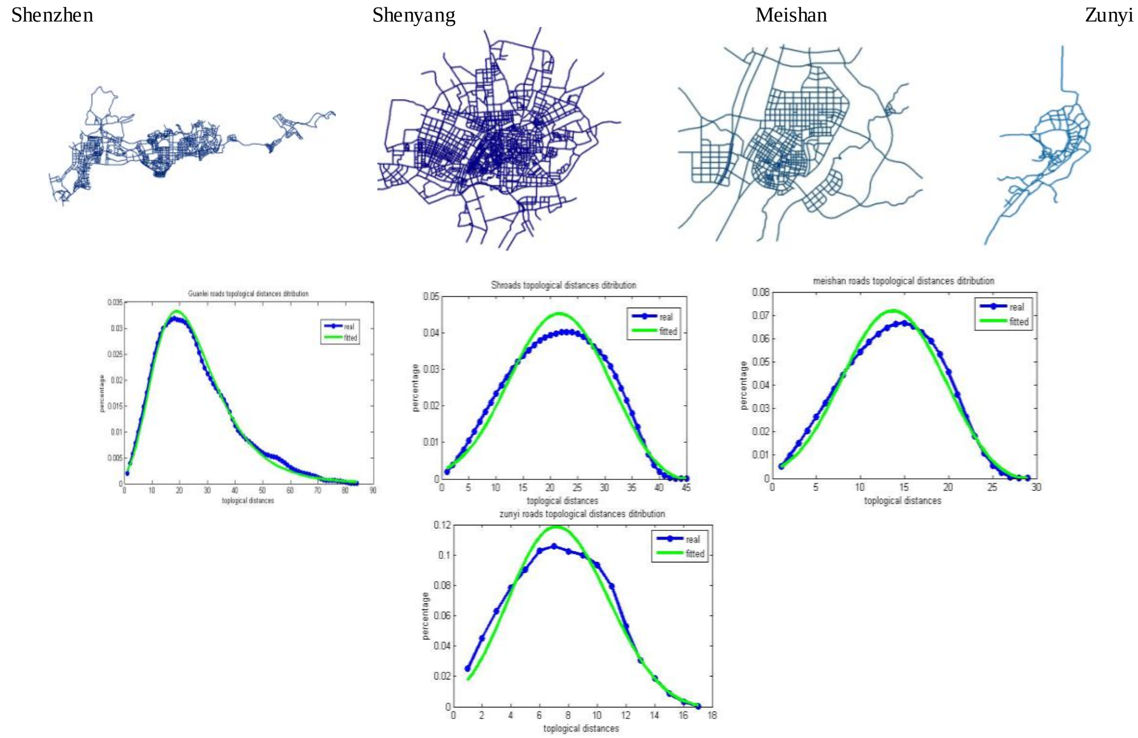

Fig. 3 Maps and GEV fittings of four cities' road network.

Furthermore, network features can be observed from the probability distribution curves fitted for different city networks, which could also be caught by the parameters of the GEV distribution (TABLE I). Take Shenzhen for instance, the parameter k for its fitting is positive, indicating its tail decreases as a polynomial, while the parameter σ is the largest of the four, which means its variance being the largest. Thus it could be primarily concluded that of all the path distances comparing with the median distance in Shenzhen, the longer ones take a lower part than the shorter ones. This is consistent with the pattern of its network, which is mainly positioned in the east and west direction. For the city Zunyi, although not as obviously as Shenzhen, the layout of the network is also in such pattern, which may probably be one of the reasons why the parameter k for Zunyi is larger than that for Shenyang and Meishan.

TABLE I CONCLUSIONS OF GEV FITTINGS FOR THE FOUR CITIES

city	ξ (shape)	σ(scale)	μ(location)	mean	log-likelihood	R-square
Shenzhen	0.022	11.076	19.159	25.795	539.977	0.993
Shenyang	-0.326	8.655	18.448	21.266	274.770	0.960
Meishan	-0.314	5.411	11.757	13.560	167.538	0.975
Zunyi	-0.259	3.219	6.236	7.421	80.944	0.955

From the above example, it could be noticed that the parameter is sensitive to the distances distributions of a network and σ to the variance of the distribution. Also, the parameter μ could reflect the whole levels of distances in a network. In the next section, a method in order to identify different network patterns through these parameters will be proposed.

3. Network Pattern Classification Method

As illustrated by the above example, the parameters gained from GEV fitting could distinguish network patterns. To see how well it fits, three types of trial networks were generated and calculated using GEV distribution.

Networks generated here include grid network, circle+radial network and tree-patterned network, as in Fig. 4. Only four samples out of ten for each type of networks are shown here since the patterns are very easy to catch. For grid networks, the label 3m3 means there are 3 roads longitudinal and the other 3 latitudinal. For circle+radial networks, the label 1O4I means that in the network, there is 1 circle road and 4 radial roads. In order to make each intersection connects no more than four roads, the center of a circle+radial network is actually a much smaller circle which joins all the radial roads. For tree-patterned networks, roads in the networks are hierarchical and are classified into three grades. Again, considering the numbers of roads connected to one intersection, each network was generated in four symmetric quadrants. The label a3c2 means each arterial road in one quadrant is divided into 3 sections and each collector (or subsidiary) road is divided into 2 sections.

In order to make their sizes comparable, all the networks are created within an area of approximate 10km2, and will be calculated using the matrix {bij}, as was introduced in section 2. As a result, the longer the roads are in the network, the denser the road network is, but the distances of the shortest path in the network will not change much. In this way, network patterns could be comparable.

Shortest paths of thirty trial networks of three types of patterns were fitted using GEV distribution. The sizes of the trial networks calculated here are listed in Table II Only 28 networks rather than 30 were analyzed because several tree-patterned networks of small sizes don't consistent with the cluster result. That may still make sense since sometimes a network with its size too small doesn't show its pattern features clearly.

According to these parameters gained from the GEV fitting, points were plotted in a three dimensional space with each dimension representing one parameter as illustrated in Fig. 5. Points in this space were further clustered into three groups through the method of cluster analysis using Matlab as a software tool. The three clusters were such set to keep the points within each cluster as close to each other as possible, and as far from points in other clusters as possible. Each cluster is characterized by its centroid, or center point. Noted that the distance used in this clustering, which is squared Euclidean distance specifically, does not represent spatial distance but only distance in the space.

Grid network Circle+radial network Tree-patterned network

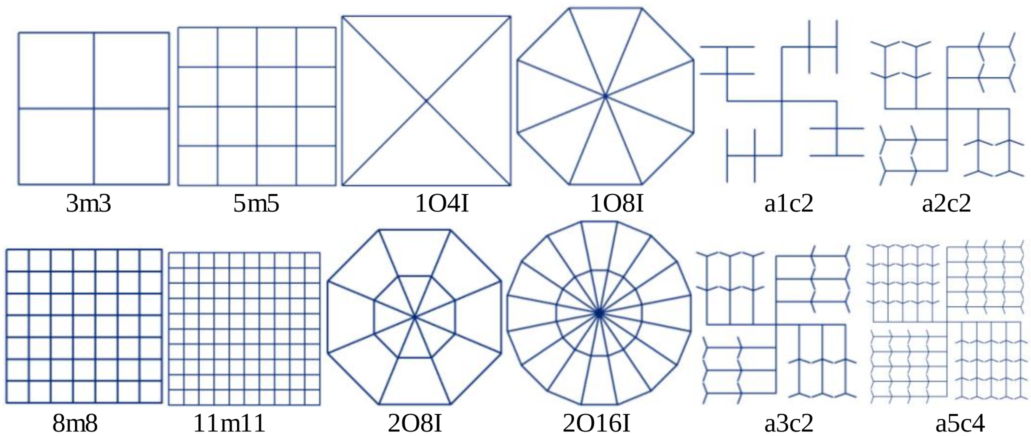

Fig. 4 Sa mples of the trial networks

TABLE II SIZES OF THE TRIAL NETWORKS CALCULATED

grid	3m3	4m4	5m5	6m6	7m7	8m8	9m9	10m10	11m11	16m16
circle	1O4I	1O6I	1O8I	2O4I	2O6I	2O8I	2O10I	2O16I	2O20I	2O30I
tree	-	-	a3c2	a3c3	a3c4	a4c4	a5c4	a5c5	a5c6	a8c8

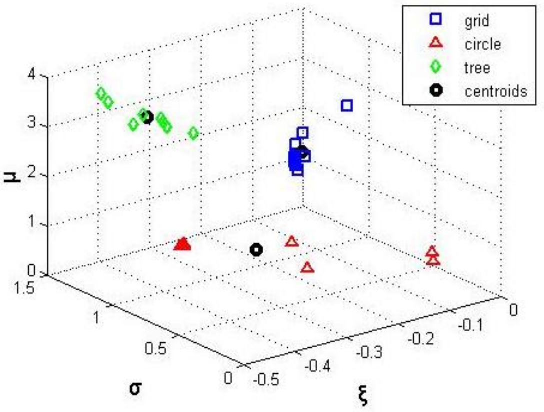

Fig. 5 GEV parameters of the trial networks and the cluster conclusion.

According to Fig. 5, the cluster results fitted the original parameter data well, which means the cluster result is consistent with the three original network pattern settings. In order to further show the quality of the cluster fitting the original data, a hierarchical tree of the cluster was plotted in Fig. 6 using Matlab again as a software tool. In a hierarchical tree, clusters at one level are joined as clusters at the next level. There are many U-shaped lines connecting objects in a hierarchical tree. The height of each U represents the distance between the two objects being connected.

Three clusters are presented using different colors. Labels of nodes represent types of patterns, which are grid (numbered 1-10), circle+radial (numbered11-20), and tree-patterned (numbered 21-28) networks, separately. The correlation coefficient between the hierarchical graph (or the cluster result) and the original data is calculated to be 0.7801, which indicates that they are highly correlated linearly.

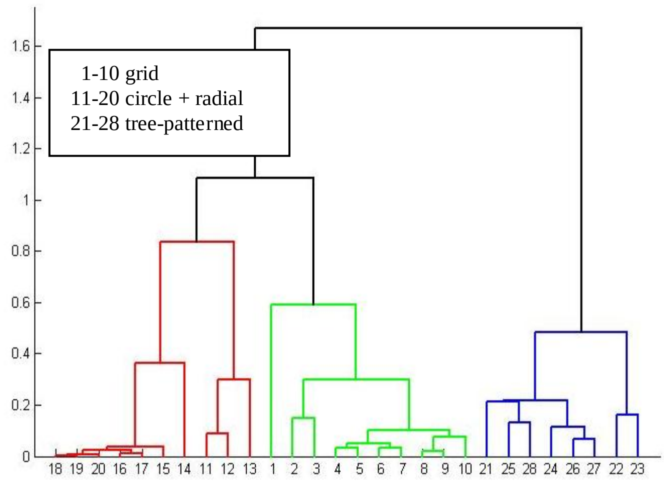

Fig. 6 Hierarchical tree of the cluster.

The criteria of pattern classification have been established according to the cluster analysis result. When classifying a network, three GEV parameters should be estimated first, which would generate a point in the three dimensional space mentioned above, then distances between the new point and three cluster centroids would be calculated, according to which the pattern of the new network would be identified. An example will be shown in the next section.

4. Validation of The Classification Method

In order to validate the pattern classification method and criteria proposed above, six road networks (in Fig. 7) are tested. They were cut from three cities in China. The area of each network is about 10 km2, which is consistent with the identification criteria.

GEV fitting parameters and the distances in the space were calculated and listed in TABLE III, from which the pattern class of each network could be identified quantitatively. Take Chengdu1 as an example, its distance to the cluster centroid of grid is the smallest comparing with that to others, which means its pattern is most likely to be grid.

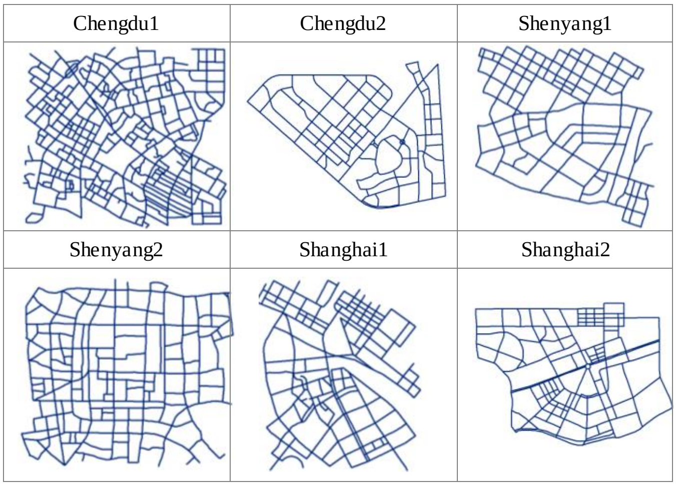

Fig. 7 Network samples for validation

TABLE III GEV PARAMETERS AND DISTANCES BETWEEN THE NETWORK AND THE CLUSTER CENTROIDS OF STANDARD PATTERN CLUSTERS

	GEV parameters			Distance to the centroids		
	ξ	σ	μ	Grid	Cricle+radial	Tree-patterned
Chengdu1	-0.222	0.886	1.651	0.140	0.516	2.214
Chengdu2	-0.133	0.936	1.631	0.128	0.535	2.266
Shenyang1	-0.169	1.077	1.836	0.021	0.931	1.621
Shenyang2	-0.158	0.974	1.737	0.061	0.700	1.935
Shanghai1	-0.246	1.121	1.954	0.021	1.172	1.302
Shanghai2	-0.190	0.838	1.469	0.301	0.289	2.817

In order to make the pattern classification result more clear, the new samples shown as points plotted in the space and a transferred graph is shown in Fig. 8 and Fig. 9. In the transferred graph, a triangle is used representing the whole pattern space based on the centroids mentioned above. Each node of the triangle represents one type of pattern with one specific color. The position of each sample point stands for its closeness to each of the three standard patterns.

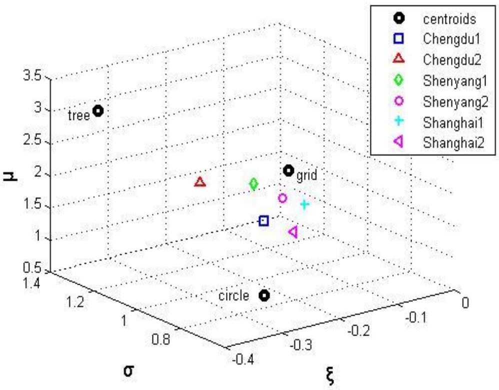

Fig. 8 The new sample points.

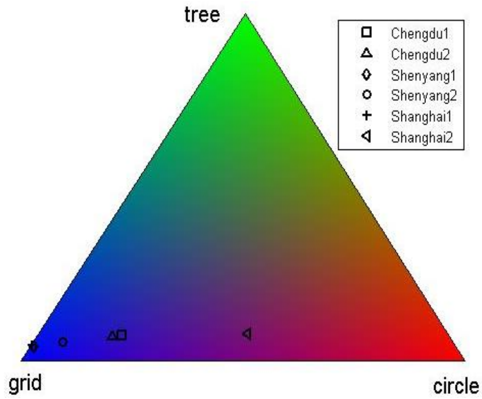

Fig. 9 A transferred graph.

It can be observed from the transferred graph that five out of the six network samples is positioned within the grid area while Shanghai2 sample is positioned closer to the circle node comparing with grid node, which is consistent with TABLE III. Referring to Fig. 7, it could be observed that it indeed is circle-and-radial network, which indicates that the method proposed in this paper has captured the pattern features of the networks.

All the trial and sample networks calculated in this paper are within an area of 10km2, which is based on the purpose of covering the main features in urban road network patterns. For a smaller area, the structure of a network may not be completed and for a larger area, the structure may be too complex to be classified into one type of pattern. Nevertheless, this method could still be used for any size of networks and capture their pattern features from the perspective of shortest paths distributions. Noted if a different size of network would be calculated using this method, the criteria should be calculated and validated before classification of sample networks because the three parameters, especially the location parameter μ, will be influenced by the size of the networks.

5. Conclusions

A quantitative method is proposed in this paper to classify patterns of road networks. Metric distances of shortest paths were taken as the main features of the networks and the features were transferred to three parameters through fitting to GEV distribution. The parameters of several trial networks have been studied and according to which a set of pattern criteria were established using cluster analysis. Six real networks were calculated using the method and their patterns were identified according to the proposed criteria. It turned out that the method could capture the main pattern features of the networks well.

This method may set a threshold of a more general method through which networks of cities could be calculated and identified quantitatively. Then other researches on networks based on their patterns may be carried on, such as efficiency comparison of networks of different patterns.

Nevertheless, there are several limitations in the proposed method that deserves further researches. The criteria proposed in this paper did identify the network features, but during the setting up procedure of the criteria, the three sets of trial networks were created artificially, which may be subjective comparing with the networks in the real world. Furthermore, only distances of shortest paths were considered as features of network pattern, which may not be enough to capture all the features. That is to say, several other important properties related to network patterns may be missed in this method.

Overall, this paper provides a new perspective and method to urban road network analysis, and also has its valuable meanings to the planning and assessment of urban road networks.

Acknowledgement

This research is sponsored by Shanghai Rising-Star Program (08QA14066) and Program for New Century Excellent Talents in University (NCET-08-0406).

References

[1] M.Southworth, E. Ben-Joseph, “Streets and the shaping of towns and cities,”.McGraw-Hill, 1996.

[2] ITE, “Recommended Practice for Context Sensitive Solutions in Designing Major Urban Thoroughfares for Walkable Communities,” Washington D.C., 2006.

[3] ITE， “Smart Growth Task Force. Smart Growth Transportation Guidelines: An ITE Proposed Recommended Practice,” *Institute of Transportation Engineers*, Washington D.C., 2003.

[4] J. Sun, G. R. Lovegrove, “Using Community-Based Macro-Level Collision Prediction Models to Evaluate the Safety Level of Neighborhood Road Network Patterns,” *Transportation Research Board Annual Meeting* 2010, 10-0535.

[5] Z. Sun, “The Study of Fractal Approach to Measure Urban rail Transit Network Morphology,” *Journal of Transportation Systems Engineering and Information Technology*, vol. 7, pp. 29-38, 2007.

[6] C. Yang, Y. Wang, “Morphological and Functional Analysis of Urban Rail Transit Networks,” *Transportation Research Board Annual Meeting* 2010, 10-3018.