2010 13th International IEEE
Annual Conference on Intelligent Transportation Systems
Madeira Island, Portugal, September 19-22, 2010

A Spatial Analysis Approach for Describing Spatial Pattern of Urban Traffic State

Haixiang Zou, Yang Yue, Qingquan Li and Yonghui Shi

Abstract—**To urban transportation management, it is valuable to obtain the whole road network traffic state, while most existing traffic analysis methods pay little attention to this problem. This paper proposes a method to analyze spatial pattern of road network traffic state, which describes traffic characteristic from network level. Above all, we treat traffic data as road link's attribution and establish a spatial model to describe road network traffic data from spatial perspective. Then using real travel speed data extracted from long-term floating car data, we analyze spatial dependency of traffic state based on spatial autocorrelation method, which shows that urban road network traffic state exist association in space, and this method can quantify the influence degree of road link, which is useful to micro-view traffic management; using trend surface analysis method, we analyze spatial heterogeneity of urban traffic state, which qualitatively shows that urban traffic state is unstable in space and is related to land use. Further, divergence tendency analysis shows that urban traffic state is influenced by road network structure, and it can also quantitatively analyze the influence range, which is useful to detailedly discover hot spot in urban traffic, such as traffic congestion.This study shows that it is a feasible approach to use spatial analysis method to study the overall characteristics of urban traffic system, which is of benefit to both dynamic traffic control and long-term traffic planning.**

I. INTRODUCTION

Traffic state is an important indicator to measure road network performance. Transport managers are often more concerned with road network overall traffic state, instead of that of some individual road links only. However, existing studies often focus on isolated freeways or urban primary arteries, with links enclosed. Limited deployment of traffic sensors is one of the reasons; the other reason is that most studies pay more attention to the temporal dimension of road link traffic following the time series representation of traffic data [1]. Therefore, the analysis of urban road network traffic state from spatial perspective is generally lacked.

This work is supported by the National Natural Science Foundation of China (No. 60872132, No. 40830530), the National Basic Research Program of China (973 program, No. 2006CB705500), Hong Kong RGC Grant (No. 754109) and State Key Lab of Resources and Environmental Information System open fund.

Haixiang Zou is with School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, Hubei, China (E-mail: zou.mono@gmail.com).

Yang Yue and Qingquan Li is with State Key Laboratory for Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, Hubei, China (E-mail: {yueyang, qqli}@whu.edu.cn)

Yonghui Shi is with research department of Wuhan traffic management bureau, Wuhan, Hubei, China

At present, with the availability of a large amount of floating cars, it has become convenient to collect traffic data over the whole road network; and some traffic data analysis methods have taken road network correlation into account. Sun et al. [2] proposed a spatiotemporal Bayesian network predictor, which incorporates all the spatial and temporal information available in a transport network to forecast traffic flow; Tomoaki et al. [3] proposed a floating car data imputation method based on feature space projection, which takes link correlation into account. However, these works are constrained to microscopic traffic analysis. Some methods have also been developed to analyze traffic data from spatial, i.e., macroscopic view. Yuan and Mills [4] analyzed spatial–temporal characteristics of traffic in large-scale networks on the influences of local behavior over the whole network; Yue and Yeh [5] choose a small region road network as study object, use cross-correlation analysis to investigate spatiotemporal dependency of traffic flow. But these methods still failed to describe the inner spatial characteristics of traffic state from road network level.

A priori, road network traffic states (such as volume, speed) are spatially correlated because vehicles are constrained on road network; and many traffic phenomena have explicit spatial aggregation, such as traffic congestion. Therefore, it is feasible to analyze urban traffic state using spatial analysis methods, which have relatively mature theory and have been widely used in many fields. This study proposes a spatial analysis model to analysis urban traffic state.

The rest of this paper is organized as follows: part II establishes an road network traffic data spatial model; part III introduces spatial analysis techniques and related definitions used in this study; part IV illustrates some real-data experiments; and part V gives some conclusions and puts forward some future research directions.

II. SPATIAL MODEL FOR ROAD NETWORK TRAFFIC DATA

To utilize spatial analysis methods and analyze road network traffic state, it is necessary to establish a spatial model to describe road network traffic data from spatial perspective. Following definitions are given to describe some of the modeling assumptions.

Firstly, since urban traffic network is composed of many road links and intersections in terms of spatial structure, it can be abstracted as line pattern objects and point pattern objects.

978-1-4244-7659-6/10/$26.00 ©2010 IEEE

Definition1: A road intersection is treated as a node which has more than two adjacent links.

Definition2: A road link is a segment between two consecutive intersections.

Since it is possible to define a time interval *Δt* in which traffic state of a road link is relatively homogeneous, a road link can be abstracted as a point pattern defined by its traffic state:

Definition3: A road link can be abstracted into a *Link Point* which is the mid-point of the road link, and the attribute of *Link Point* is the traffic state during time interval *Δt*.

Therefore, a road network is transformed into sparse points in space, and traffic data upon it (*St*) is split into cross-sectional data p_n by *Δt* (see Fig.1).

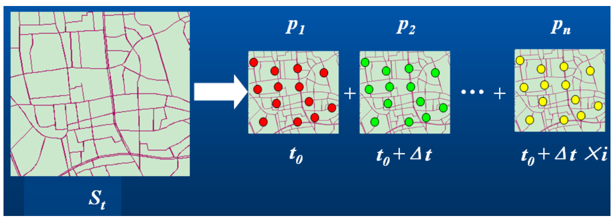

Fig.1. Split traffic data into cross-sectional data as fixed time granularity

Distance is the most important indictor to reflect the extent of spatial correlation between spatial objects. Because intersection is the major issue affecting urban traffic state, our model uses the number intersections between two road links as the distance of them in network space instead of their real distance in free space.

Definition4: The *network distance* between two road links is the least accessible number of intersections between two link points.

III. SPATIAL ANALYSIS TECHNIQUES

Traffic data is usually not randomly distributed over road network but takes on association and aggregation, therefore, traffic state spatial dependency and heterogeneity are two important aspects in understanding road network performance. *The First Law of Geography* [6] point out that, adjacent geographical units exists spatial dependency, and it decays with increase of distance. Goodchild et al. [7] also pointed out spatial data tends to local aggregated and also shows unstationary, therefore, spatial heterogeneity coexist with spatial dependency. In this section, we introduce two spatial analysis techniques to be used in this study to analyze traffic spatial dependency and spatial heterogeneity.

A. *Spatial autocorrelation analysis*

Spatial autocorrelation analysis can be used to measure spatial dependency, which assesses the extent to which the value of a variable at a given location influences values of that variable at contiguous locations [8]. Many global or local statistics are proposed to assess spatial autocorrelation from Global Moran I [9] and have been applied in many space-related fields, such as econometrics, geography, ecology, and immunology. Global Moran I's definition is shown as follow [10]:

$$I(d) = \frac{\sum_{i}^{n}\sum_{j\neq i}^{n} w_{ij}(x_i - \overline{x})(x_j - \overline{x})}{S^2 \sum_{i}^{n}\sum_{j\neq i}^{n} w_{ij}} \qquad (2)$$

Where, S^2, x_i and x_j denote the attribute X's variance, observed value at location *i* and *j* respectively, $\overline{x}$ is the average of the $\{x_i\}$ over the n locations, and w_{ij} is the spatial weights Matrix, which can be defined by different application [11].

The Moran I value is usually between -1 and 1, and it is statistically significant and positive when the observed value of locations within a certain distance (d) tend to be similar, negative when they tend to be dissimilar, and approximately zero when the observed values are arranged randomly and independently over space [12].

Under the assumption that Moran I is a normal distribution, Cliff and Ord [10] calculate the expected value of Moran I, set as *E(I)*; and the variance of Moran I, set as *VAR (I)*. The test on the null hypothesis that there is no spatial autocorrelation between observed values over the *n* locations can be conducted based on the standardized statistic as formula (3):

$$Z(d) = \frac{I(d) - E(I)}{\sqrt{VAR(I)}} \qquad (3)$$

According to normal distribution test value, null hypothesis can be rejected when $|Z| > 1.96$, which means the attribute values of spatial object do not fit normal distribution. In other words, probability of the attribute existing spatial autocorrelation is more than 95%.

B. *Trend surface analysis*

Trend surface analysis uses continuous function to fit the observed data and is commonly used to analyze data in map form [13]. Kernel density estimation (KDE) is a trend surface analysis technique based on non-parameter statistic. It uses form of density graph to visualize space pattern, which can be applied to analyze distribution characteristic of spatial data. The core estimator is defined as formula (4) [14]:

$$f(s) = \sum_{i=1}^{n} \frac{1}{\pi r^2} k\left(\frac{d_{is}}{r}\right) \qquad (4)$$

Where, *f(s)* is the density at location s, *r* is the search radius (bandwidth) of the KDE (only points or lines within *r* are used to estimate *f(s)*), *k* is the weight of a point *i* at distance d_{is} to location *s*. *k* is usually modeled as a function (called kernel function) of the ratio between d_{is} and *r*. Gaussian function is a commonly function form in KDE, which is shown in formula (5) [14]:

$$k\left(\frac{d_{is}}{r}\right) = \begin{cases} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{d_{is}^2}{2r^2}\right) & 0 < d_{is} \leq r \\ 0 & d_{is} > r \end{cases} \qquad (5)$$

However, the choice of bandwidth is more important than choosing the weighting function. The shorter bandwidth can obtain more detail, but also costs longer calculate time.

To demonstrate the usage of the two methods in analyzing road network traffic state, in the next section we apply them to the real urban road network traffic data, and analyze the spatial characteristics of urban traffic state through some experiments.

IV. EXPERIMENTS

A. *Experiment data*

In our study, real floating car data collected from approximately 1,200 taxis with three-week temporal extent (seven weekends included) in the downtown of Nanchang (a medium-sized city in China, see Fig.2) is used as experimental data, and the average distance between two intersections is about 370m. GPS signals are sent every 100 meter by unloaded taxis or every 400 meter by loaded taxis and cover about 90% roads of the whole road network. In the data pre-processing step, we extract travel speed per 5 minute as experiment sample, which is a representative traffic parameter and can be conveniently extracted from floating car, according to a method proposed in our previous study [15].

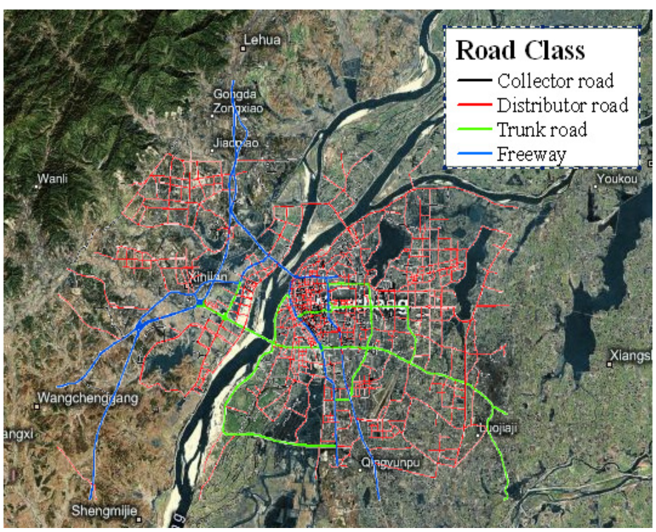

Fig.2. Nanchang's road network map

Furthermore, to analyze traffic state in different time periods, we divide one day into 4 time period according to classical urban traffic pattern: workday's morning peak (8am-9am), workday's non-peak (3pm-4pm), workday's evening peak (6pm -7pm) and weekend (8am-10pm; traffic state of weekend is steady and has no distinct peak period, so the time interval is longer than others).

B. *Spatial dependency analysis*

We use the average of travel speed samples in each time period as a road link's attribute in that time period. And based on spatial weight matrix: $W = \sum_{d=1}^{D} W_d$ (d is order number) proposed by Anselin and Smirnov [16], which is commonly used to build high-order spatial weight matrix, we build seven spatial weight matrixes from 1st order to 7th order. For example, according to Definition4, when the 5th order spatial weights matrix is defined, the nearest neighbor links of link i are the set of all links that can pass through 4 or less than 4 intersections to reach link i, and if one link has existed in low order neighbors, it cannot exist in high order neighbor again. Fig.3 show one link's 1st order to 7th order neighbor links.

According to formula (2) and formula (3), we calculate global Moran I and corresponding Z score using different spatial weigh matrix in different time period. Shown in table 1, it can be observed that: (1) all the Moran I values are positive and all the z scores are bigger than 1.96; (2) Moran I decrease with increase of order number and tend to be zero from special order.

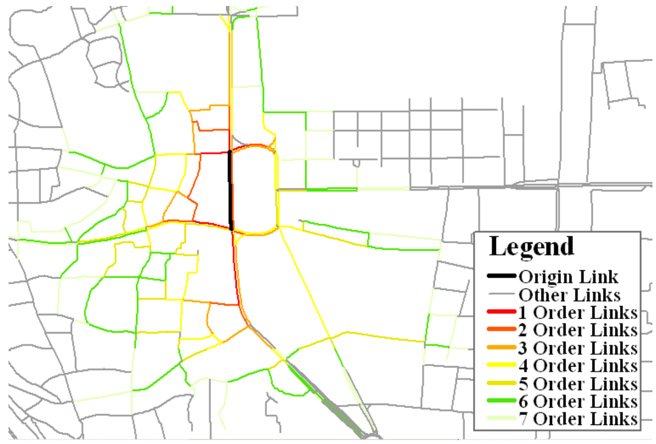

Fig.3. One link's hierarchically adjacent links (From 1st order to 7th order)

According to the meaning of global Moran I and Z score described above, it can come to the conclusions that:

- Road network traffic state exist significant spatial autocorrelation in every time period, which means that traffic state exist association in urban road network, and it is the same as experiential knowledge that two contiguous roads have similar;
- The extent of spatial autocorrelation decreases with increase of road link distance, which coincides with real-world situation that traffic state is usually dissimilar between two far apart roads. In this case, it shows that Moran I trends to 0 (<0.1) from 5 order in morning peak, which means that the average influence degree of one road link is five intersections in road network; and weekend's Moran I is bigger than other periods, which means that influence degree of road link in weekend is stronger than other periods, in other words, the distance between two road link have similar traffic state is longer than other periods.

By spatial dependency analysis, on the one hand, it can qualitatively prove that urban road network traffic state exist significant similarity in space, and on the other hand, it can also be used to quantify the influence degree of road link in urban traffic system, which is useful to micro-view traffic management, for example, it can offer accurate sample range to short-term forecasting.

TABLE I. THE Z SCORE TEST RESULTS OF TRAVEL SPEED'S GLOBAL MORAN I IN DIFFERENT TIME PERIOD

Order	Morning Peak(8:00-9:00)		Non-peak(15:00-16:00)		Evening Peak(18:00-19:00)		Weekend(8.am-10.pm)	
	Moran I	*Z Score*	*Moran I*	*Z Score*	*Moran I*	*Z Score*	*Moran I*	*Z Score*
1	0.2148	5.4087	0.2848	6.1831	0.2358	5.2770	0.3006	7.3655
2	0.1651	6.3077	0.2127	7.0030	0.2227	7.6301	0.2333	8.7308
3	0.1332	6.6787	0.1788	7.7005	0.1781	7.9844	0.1941	9.5592
4	0.1111	6.9198	0.1660	8.8484	0.1532	8.4722	0.1624	9.9139
5	0.0909	6.8115	0.1424	9.0619	0.1327	8.7090	0.1426	10.400
6	0.0790	6.9370	0.1243	9.1813	0.1130	8.5747	0.1257	10.671
7	0.0681	6.8628	0.1065	8.9792	0.0966	8.3317	0.1112	10.784

C. *Spatial heterogeneity analysis*

To analyze spatial heterogeneity of traffic state, we build travel speed trend surface of the whole road network based on KDE method introduced in part III. To obtaining more details, the bandwidth is set to 500 meter.

First, to examine traffic state in the whole space, Euclidean distance in free space is used. Fig. 4 shows travel speed density graph in workday's morning peak, which explicitly demonstrate that, the spatial distribution of traffic state is unstable, and shows local aggregation. In Fig. 4, the deeper color means lower travel speed, and the shallower color means higher travel speed.

By analyzing the distribution of different aggregation densities, we can find out different traffic state areas. It can be applied to detect rough influence areas of traffic congestion.

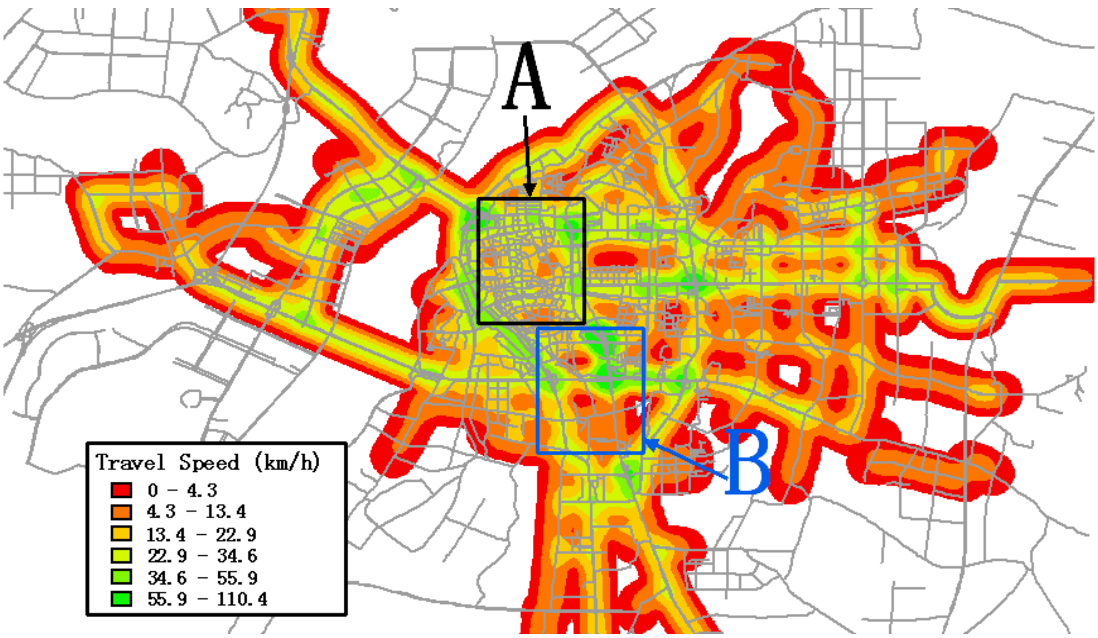

Fig.4. The travel speed density graph of road network in workday's morning peak

Moreover, our experiment shows that the existing of traffic spatial heterogeneity also related to land use. Fig. 5 compares the travel speed's distribution in commercial zone (Region A in Fig.4) and residential zone (Region B in Fig.4), it shows that the different functional zone's traffic state significantly different from each other even in the same time period.

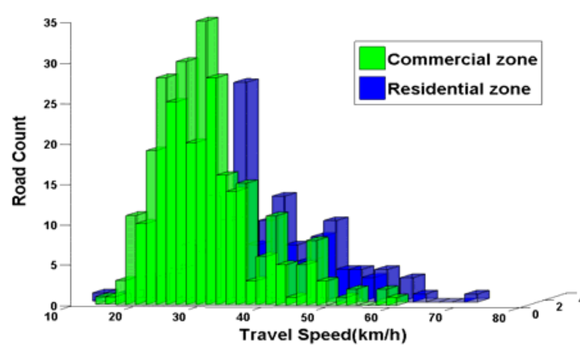

Fig.5. The comparison of travel speed distribution between different functional zones in workday's morning peak

D. *Divergence tendency analysis*

To further investigate the traffic spatial heterogeneity over road network, spatial divergence tendency is conducted which takes road links' travel speed and its spatial location into account. To facilitate the analysis, a mid-point of a road link is selected as the benchmark in the experiment, which is located almost at the center of study area (see fig.6).

Then, each road link's network distance with the benchmark point is calculated based on Definition 4, which is defined as *network divergence distance* (NDD). For example, in fig.6, the red line is the route which contains the least number of intersections from road link A to network's center and the sum is 37, therefore the NDD is equal to 37.

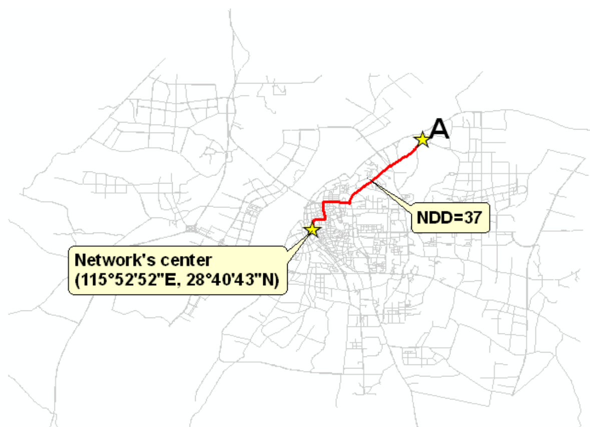

Fig.6. The center of study area and network divergence distance (NDD)

Fig.7 depicts the relation of the examined road links' divergence distance with their associated travel speed in different time periods. The figure shows that:

- The shape of the sectors in each time periods all illustrates that, the relationship between road link's spatial location and travel speed is nonlinear. In general, the further network distance from the benchmark point, the higher travel speed. This also proves traffic spatial heterogeneity.
- The included angles of the sectors are different in each time period. This reflects that traffic state spatial heterogeneity and divergence are temporal-varying. For example, workday's morning peak has the smallest included angle among them, which means that the whole traffic state of workday's morning peak is the most homo-geneous.

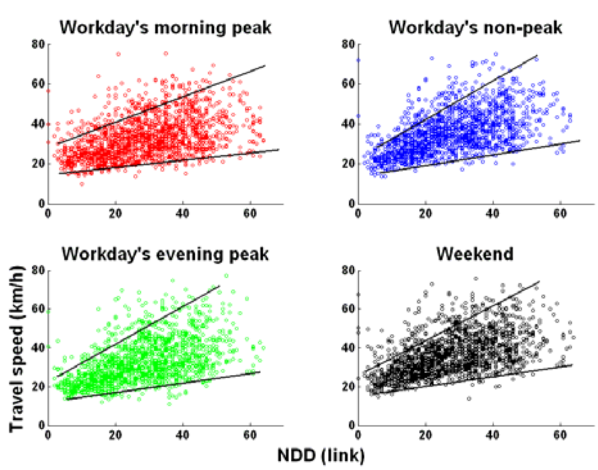

Fig.7. The relation graph of road link's NDD with travel speed

To take a closer look of the divergence, we sort road link's travel speed from low to high, and then trisect them into three groups: low speed, middle speed and high speed. Fig. 8 shows the results in different time period:

- Compared with the road links of middle speed and high speed, the road links with low speed mainly concentrate on locations near the urban center, and the number of low speed road link decreases significantly with the increase of distance. This means, traffic state is worse in urban center area, and becomes better in the outer area which in this study is about 15 intersections away from the network's center in workday's morning peak.
- The distributions of middle speed and high speed levels show more peaks and hollows than that of low speed level. This can be used to discover hot and cold regions in urban road network by traffic state.

The spatial distribution pattern of traffic state is influenced by both road link location and urban road network structure. Although this study only performs analysis in one city, similar result is expected in other cities.

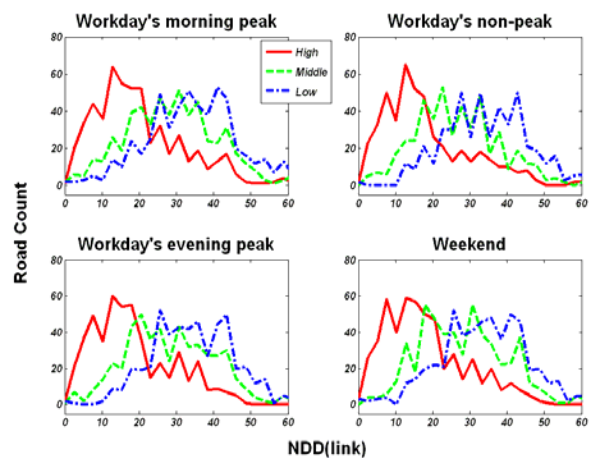

Fig.8 The distribution of different level's travel speed in network space

V. CONCLUSIONS AND FUTURE WORKS

Spatial analysis provides a means to study traffic state from road network level. This study proposes a spatial analysis model for urban transport network, and introduces spatial analysis theory and techniques, such as spatial autocorrelation analysis and trend surface analysis, to analyze the traffic data. Experiment based on real-world data illustrates that urban traffic state exists significant spatial dependency and heterogeneity, with divergence tendency.

On one hand, the distribution of urban traffic state is unstable; on the other hand, urban traffic state shows the feature of local aggregation, which is influenced by both road link location and road network structure. All the spatial characteristics of traffic state exhibit certain temporal features.

The interaction of road network spatial structure and its traffic state is a potentially important but largely unexplored area. This study takes spatial issue into consideration to analyze urban traffic state. It, for example, can be used to discover urban road network hot and cold regions, and to understand the overall performance of urban traffic system, which is of benefit to both dynamic traffic control and long-term urban and traffic planning.

REFERENCES

[1] W. Willinger, M. Taqqu, and A. Erramilli, “A bibliographical guide to self-similar traffic and performance modeling for modern high-speed networks”, Stochastic Networks: Theory and Applications, Roy. Stat. Soc. Lecture Notes Series, Vol. 4, 1996.

[2] Sun Shiliang, Zhang Changshui, and Zhang Yi, “Traffic Flow Forecasting Using a Spatio-temporal Bayesian Network Predictor”, *Artificial Neural Networks: Formal Models and Their Applications - ICANN 2005, LNCS 3697*, pp.273–278, 2005

[3] Kumagai M., Fushiki T., Kimita K. and Yokota T., “Spatial interpolation of real-time floating car data based on multiple link correlation I feature space, *Proc. of 13th World Congress of ITS,* London, CD-ROM, 2006.

[4] Yuan J. and K. Mills., "A cross-correlation-based method for spatial-temporal traffic analysis", Performance Evaluation Long Range Dependence and Heavy tail Distributions, Vol.61, issues.2-3, pp.163-180, 2005.

[5] Yue Y. and Yeh A. G., "Spatiotemporal traffic-flow dependency and short-term traffic forecasting", Environment and Planning B: Planning and Design, Vol.35, pp.762-771, 2008.

[6] Tobler W. R, "A computer model simulating urban growth in the Detroit region", Economic Geography, Vol.46, No.2, pp.234-240, 1970.

[7] Goodchild M.F, Anselin L., Appelbaum R. and Harthorn B., "Toward spatially integrated social science", International Regional Science Review, Vol.23, pp.139-159, 2000.

[8] William R. Black and Isabelle Thomas, “Accidents on Belgium’s motorways: a network autocorrelation analysis”, Journal of Transport Geography, Vol.6, No.1, pp.23-31, 1998.

[9] Moran. P., “The interpretation of statistical maps”, Journal of’ the Royal Statistical Society, Vol.10b, pp.243-251, 1948

[10] Cliff A. D. and Ord J. K., Spatial process: models and applications. London, Pion, 1981.

[11] Calderón Gema Fernández-Avilés, "Spatial Regression Analysis vs. Kriging Methods for Spatial Estimation", International Advances in Economic Research, Vol.15, No.1, pp. 44-58, 2009.

[12] Goodchild M.F., Spatial Autocorrelation (CATMOG 47). Norwich, U K: GeoBooks, 1986.

[13] A.E. Wren, “Trend Surface Analysis - A Review”, Canadian Journal of Exploration Geophysics, Vol.9, No.1, pp.39-44, 1973.

[14] Xie Z. and Yan J., "Kernel Density Estimation of traffic accidents in a network space", Computers, Environment and Urban Systems, Vol.32, No.5, pp.396-406, 2008.

[15] Yue Y., Zou H. X. and Li Q. Q.,”Urban road travel speed estimation based on low sampling floating car data”, Proceedings of the 9th International Conference of Chinese Transportation Professionals (ASCE), Harbin, China, August 5-9, 2009, pp.1719-1725.

[16] Anselin, Luc and Oleg Smirnov. "Efficient algorithms for constructing proper higher order spatial lag operators", Journal of Regional Science, Vol.36, pp.67-89, 1996.