International Journal of Civil Engineering

Transportation

Technical Note

Pattern-based short-term traffic forecasting for urban heterogeneous conditions

L. Zhang[1],*, Y. Ma[1]

Received: July 2013, Accepted: January 2014

Abstract

Short-term traffic flow forecasting plays a significant role in the Intelligent Transportation Systems (ITS), especially for traffic signal control and transportation planning research. Two mainly problems restrict the forecasting of urban freeway traffic parameters. One is the freeway traffic changes non-regularly under the heterogeneous traffic conditions, and the other is the successful predictability decreases sharply in multiple-steps-ahead prediction. In this paper, we present a novel pattern-based short-term traffic forecasting approach based on the integration of multi-phase traffic flow theory and time series analysis methods. For the purpose of prediction, the historical traffic data are classified by the dynamic flow-density relation into three traffic patterns (free flow, synchronized and congested pattern), and then different predict models are built respectively according to the classified traffic patterns. With the current traffic data, the future traffic state can be online predicted by means of pattern matching to identify traffic patterns. Finally, a comparative study in a section of the Third-Loop Freeway, LIULIQIAO, Beijing city, shows that the proposed approach represents more accurately the anticipated traffic flow when compared to the classical time series models that without integration with the traffic flow theory.

Keywords: *Traffic forecasting, Multi-phase traffic flow theory, Auto regressive integrated moving average (ARIMA).*

1. Introduction

With the growth of vehicles, traffic congestion on the freeway system is becoming more and more worrying. This is particularly obvious near big cities at morning and evening rush hours. To relieve traffic congestion, Intelligent Transportation Systems (ITSs) is introduced, which encompass a broad range of communication and electronic based technologies. With ITSs, freeway traffic is no more passive but interacts with the drivers to increase the global efficiency of transportation networks.

Short-term traffic prediction is one of the basic aspects for optimizing the transportation system operations. Many efforts on the short-term traffic flow forecasting ranging from time series models [1,2,3], nonparametric statistical methods [2,4], neutral networks [4], and Kalman filtering [5,6,7,8] can be traced in the literature. The review work of Vlahogianni et al. [9] has indicated that the nonparametric techniques have performances comparable to simple autoregressive integrated moving average (ARIMA) models.

* *Corresponding author: zhangliguo@bjut.edu.cn*

1 School of Electronic Information and Control Engineering, Beijing University of Technology, Beijing, China, 100124

Huynh et al. [10] implements an adaptive model in which the model parameters are updated online periodically (e.g., every 5 min) based on the prevailing traffic conditions.

Cetin and Comert propose an adaptive approach [11], which explicitly accounts for occasional regime changes by using online change-point detection algorithms.

Although these methods have given promising results, there are some practical issues that fail at being addressed, such as their reduced performance of multiple-steps-ahead prediction in the transitional conditions. A recent study shows that the performance decreases sharply as traffic state transfers from free flow to congested, causing errors as high as 30% [9]. This could be due to traffic data exhibit a highly fluctuating behavior with frequent and sudden transitions between different traffic states under the heterogeneous traffic conditions, and the behavior that cannot be modeled directly by these methods may encompass useful information to predict future values of traffic flow.

This paper proposes a pattern-based prediction framework that is consistent with the evolution of traffic flow theory and the ability of exploiting past traffic pattern information in order to enhance predictability. The proposed framework, which includes two main parts (see Fig. 1): offline part and online part, differs from previous studies in that it is addressed by referring the traffic pattern of traffic flow evolution. The term traffic pattern described with ARIMA model has been acquired from the past in order to predict the future value of traffic flow. First, historical traffic data is classified by the flow-density relation into three traffic states (free, synchronized and congested). Several historical traffic patterns then are built respectively according to the classified historical traffic data. The current traffic pattern can be easily identified by means of matching historical traffic patterns with the autoregressive parameters of real-time traffic data. Finally, it is to implement the matched pattern to provide a multiple-steps-ahead prediction of traffic flow. Findings indicate that pattern-based forecasting is more accurate, in the traffic flow states considered, when compared to classical ARIMA models that only operate under the time series consideration.

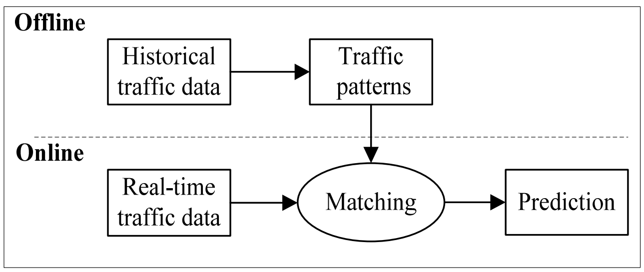


Fig. 1 Overview of the proposed framework

The remainder of the paper is organized as follows. The fundamental theoretical aspects of the proposed framework are presented in Section 2. Section 3 presents a comparative study on a section of the Third-Loop Freeway, LIULIQIAO, Beijing city, to test the proposed framework. The conclusions of this paper are given in section 4.

2. Pattern-based Short Term Traffic Flow Forecasting

2.1. Traffic state classification

The flow-density relation is one of the basic relations among traffic variables and provides a basic way to classify traffic states under heterogeneous traffic conditions. In this paper, traffic states are classified into free flow, synchronized and congested state by Kerner's three-phase traffic flow [12]. The different traffic states and transitions between the states are represented in a states diagram, see Fig. 2. The characteristics of these states encompass a distinct manner of traffic evolution with respect to deterministic and nonlinear attributes of traffic flow. This observation is significant to the predictability of traffic flow; the more deterministic the evolution of traffic states the more predictable it is.

As can be observed in Fig. 2, free flow state appears at low traffic demand. There occur transitions to other traffic states (congested state and synchronized state) depending on different traffic conditions. If the flow is rather high, traffic flow will change to the congested state. A transition to the synchronized state occurs when the traffic demand is relatively high.

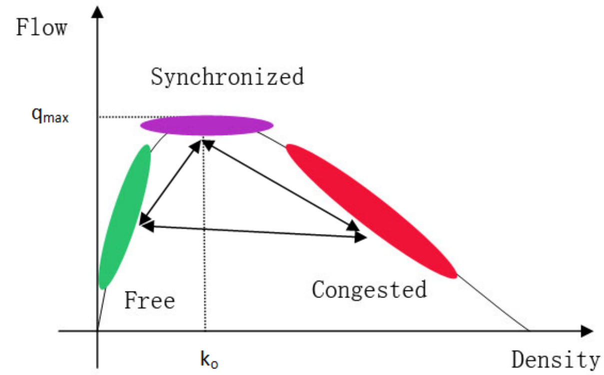


Fig. 2 Traffic states and transitions in the states diagram

The synchronized state can be formed in the transition between the free flow state and the congested state. The speed in the synchronized state is a little lower than in the free flow state, but still high. The flow on average is relatively high with nearly optimal density. In this state, the future value of traffic flow is difficult to predict in that the observed nonlinear determinism along with the occurrence of stochastic reflects a complex traffic flow reality.

In the congested state, the volume reduces, but the demand remains high; the vehicles slow down, the densities increase, and vehicles pack more closely together. During congestion the road system is operating in an inefficient manner, with increased vehicle delays, driver frustration, and greater potential for accidents. The traffic flow returns from the congested state to the free flow state directly or via the synchronized state and finally to the free flow state when the congestion dissolves.

The three traffic states are illustrated in Fig. 3, using 6 hours flow and density data collected from the studied area. As can be observed, historical traffic data has been successfully classified into three states by the flow-density relation. The historical traffic patterns for each traffic state can be built respectively according to the classified traffic data.

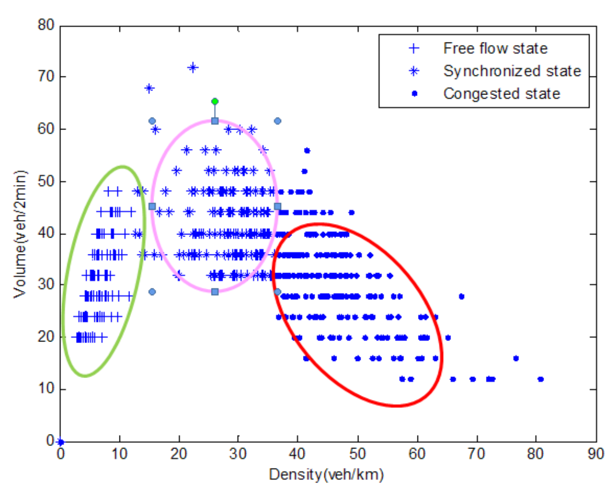


Fig. 3 Different traffic states in the flow-density relation

2.2. Building historical traffic patterns

Several historical traffic patterns of volume and density are developed for each traffic state according to the classified historical traffic data in the previous section. ARIMA models are flexible and widely used in short-term traffic flow forecasting [13, 14], which combines three processes: auto regressive (AR), differencing to strip off the integration (I) and moving averages (MA). Each of the three process types has its own characteristic way of responding to a random disturbance. The auto regressive process indicates weighted moving average over past observation, the integration process indicates linear trend or polynomial trend of the series and the moving averages process indicates weighted moving average over past errors.

The ARIMA (p,d,q) model of the time series $\{X_1, X_1 \ldots X_t\}$ is defined as [15, 16]

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q} \quad (1)$$

where, X_t is obtained by differencing the original time series d times, a_t is the white noise component at t, $\phi_i, i = 1, 2 \ldots p$ are the autoregressive parameters, $\theta_j, j = 1, 2, \ldots,\ q$ are the moving average parameters, p is the order of autoregressive parameters, q is the order of moving average parameters and d is the order of differencing.

The process of building a historical traffic pattern for the classified traffic data is achieved by four steps, see Fig. 4. The first step is model identification, in which use of the data and of any available information to identify potential models. The second step is model estimation, in which efficient use of the data to make inference about the parameters. It is conditioned on the adequacy of the selected model. The third step is model diagnostics, in which check the adequacy of fitted model to the data in order to reveal model inadequacies and to achieve model improvement. If valid then use the decided model, otherwise repeat the steps of identification, estimation and diagnostics. The last step is model index, in which use second-order autoregressive feature extractors of models to index the historical traffic pattern. The second-order autoregressive feature of historical traffic patterns of volume and density using 18 data sets is shown in Fig. 5.

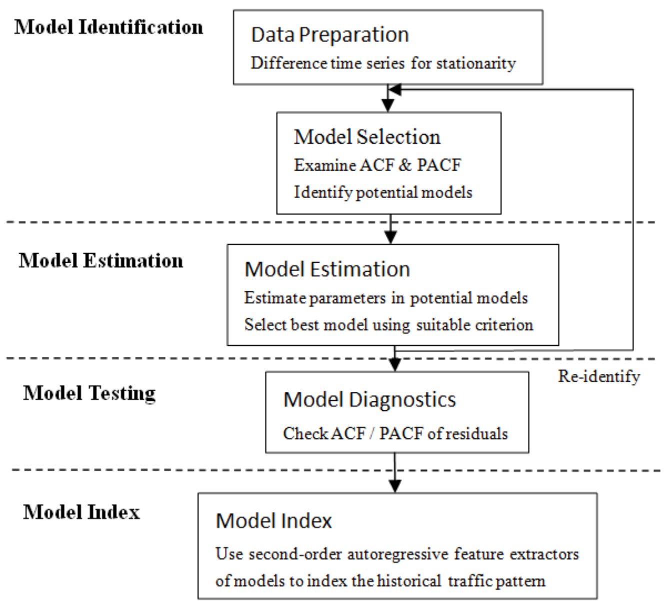


Fig. 4 Procedure of building traffic

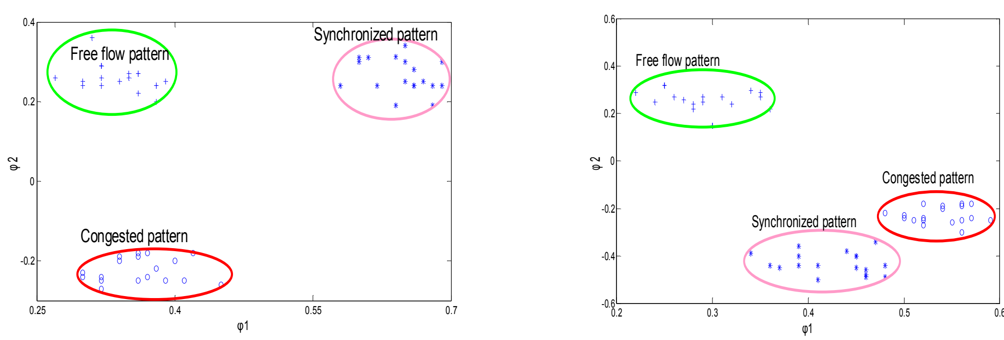


Fig. 5 Second-order autoregressive feature extractors of historical traffic patterns of volume (left) and density (right)

2.3. Pattern matching and forecasting

The future value of traffic flow can be quickly predicted with the real-time traffic data by means of pattern matching technique to identify the current traffic pattern. The pattern matching technique will be done in two steps. The first step consists in extracting the second-order autoregressive parameters of the current traffic data. In the second step we apply the Euclid distance, defined by equation 2, to measure the similarity between the historical traffic patterns and identified current traffic pattern.

$$D_E^2(\phi_T,\phi_R)=(\phi_T-\phi_R)^T(\phi_T-\phi_R) \tag{2}$$

here, ϕ_R is the autoregressive parameters of the historical pattern, ϕ_T is autoregressive parameters of the identified pattern. The historical traffic pattern that has smallest Euclid distance will be used in that it is the most similar with current traffic conditions.

3. Implementation and Findings

3.1. The traffic data

The proposed prediction framework is implemented using data collected from the study areas (on a section of the Third-Loop Freeway, LIULIQIAO, Beijing city, to model and predict traffic flow. The dataset consists of traffic speed and volume in 2-minute intervals between March 1 and March 15 of 2013. Density is calculated from the speed and volume values. The dataset is separated into two subsets: the historical set that will be used to build historical traffic patterns and the test set that is chosen to encompass various traffic flow states during a day in order to test the accuracy of predictions. The collected traffic data is not smoothed because the relation between the flow and the density can be changed during the transition from one state to another, making accurate predictions more of a challenge.

3.2. Prediction performance

Prediction performance is evaluated using the Root Mean Square Error (RMSE) and the Mean Absolute Percent Error (MAPE):

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_t-F_t)^2} \tag{3}$$

$$MAPE=\frac{1}{n}\sum_{i=1}^{n}\frac{|X_t-F_t|}{X_t} \tag{4}$$

here, X_t and F_t are the observed and the forecast values of observation t, respectively, and n is the total number of observations.

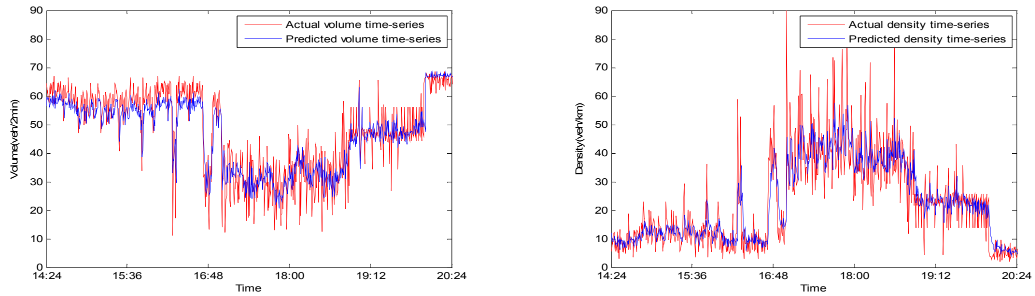
3.3. One-step prediction of traffic flow

For the purpose of prediction, several historical traffic patterns are developed for each traffic state according to the classified historical traffic data. The future value of traffic flow can be quickly predicted with the real-time traffic data by means of pattern matching technique to identify the current traffic pattern. A comparative study is conducted in order to evaluate the accuracy of the proposed pattern-based approach. The one-step prediction results for volume and density are demonstrated in Table 1. As can be observed, pattern-based approach is capable of predicting more accurately the future value of traffic flow in all traffic states when compared to classical ARIMA models that only operate under time series consideration. Moreover, there is no significant difference between pattern-based predicted and actual values of traffic flow in that the average MAPE is less than 10%. It is also observed that pattern-based approach predicts volume and density with similar levels of accuracy in all the three traffic states. Consequently, the pattern-based approach adapts to transitional traffic conditions and is consistent with the heterogeneous traffic flow.

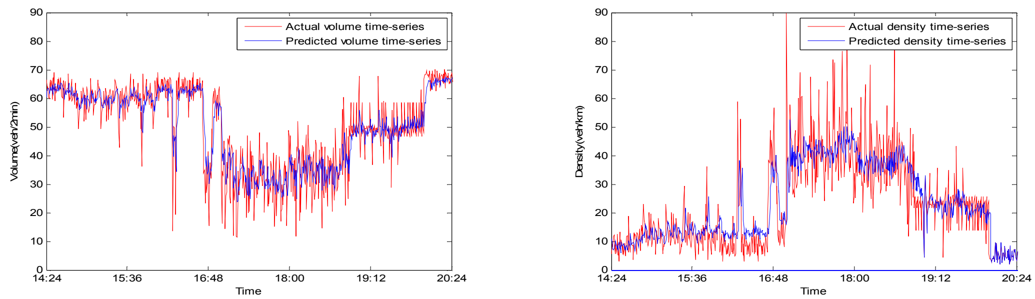
Table 1 Comparison of pattern-based one-step forecasting with ARIMA model

One-step forecasting					
Day		March 4 of 2013			
Address		A section of the Third-Loop Freeway, LIULIQIAO, Beijing			
		Patten-based ARIMA			Classical ARIMA
		Free	Synchronized	Congested	
Volume	RMSE	6.0516	5.8670	5.8670	7.4564
	MAPE(%)	8.25	9.86	8.80	11.97
Density	RMSE	5.8508	6.6120	5.8245	7.2023
	MAPE(%)	17.47	21.15	20.95	24.67

Fig. 6 shows the series of actual versus predicted values of volume and density for the afternoon peak of a typical workday. Three traffic states and their transitions are depicted in Fig. 6, the time series from 14:00 to 17:00 and from 20:00 to 21:00 shows the free flow state, from 17:00 to 19:00 shows the congested state, and from 19:00 to 20:00 shows the synchronized state. As can be observed, classical ARIMA models exhibit worse fit to actual values due to their tendency to focus on the mean values and miss the trends. Moreover, their prediction performance of short-term traffic flow decreases sharply as traffic states transfer from one to another. It is also observed that both pattern-based approach and classical ARIMA models exhibit decreased performance in synchronized flow state. This is probably due to the highly nonlinear characteristics of traffic flow, as well as the frequent transition between extreme deterministic and stochastic structure.



ARIMA



Pattern-based approach

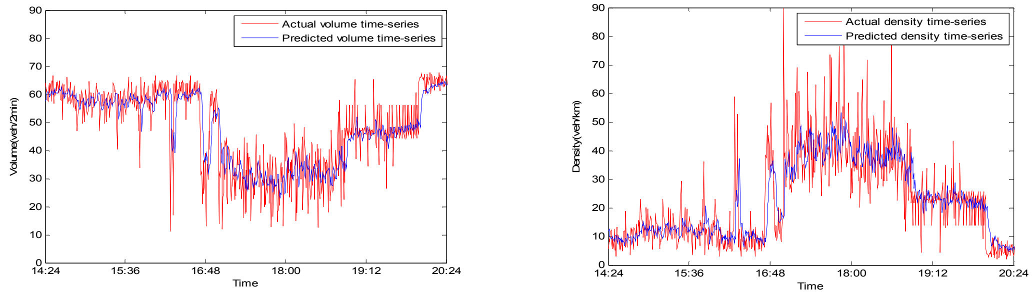
Fig. 6 Actual (blue) versus predicted (red) values of volume and density in one-step prediction

3.4. Multiple-step-ahead prediction of traffic flow

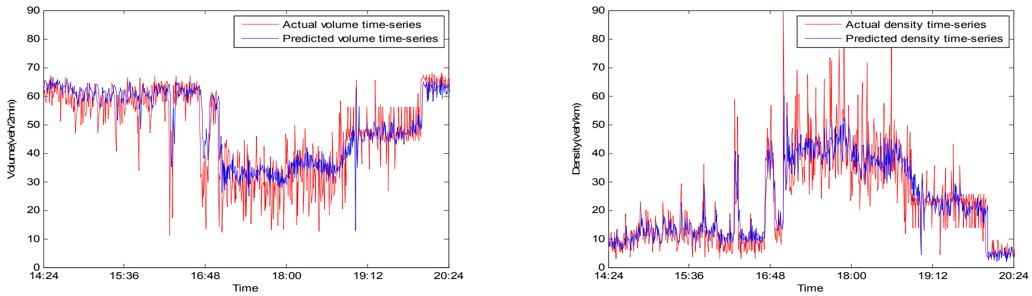
The final stage of analysis is to implement the proposed framework to provide a multiple-steps-ahead prediction of traffic flow. For the purpose of three-step-ahead prediction of traffic flow, the proposed framework is presented with volume and density data, as well as traffic pattern information on the deterministic and nonlinear features of traffic flow. The three-step-ahead prediction results for volume and density are depicted in Table 2. As can be observed, the three-step-ahead predictability of traffic flow by pattern-based approach slightly decreases in transitional conditions. The reason is that the hysteresis traffic data encompass information on the previous pattern and the information on the statistical features of traffic flow evolution is not presented to the network. As can be seen in Fig. 7, the hysteresis phenomena occurs when traffic state transfers from one to another, for example, from the synchronized state to free flow states, there are three steps hysteresis at 19:00. However, the classical ARIMA models predictability of traffic flow significantly decreases due to the predict model is not adapt to the variable traffic conditions.

Table 2 Comparison of pattern-based multi-step forecasting with ARIMA model

Multi-step forecasting					
Day		March 4 of 2013			
Address		A section of the Third-Loop Freeway, LIULIQIAO, Beijing			
		Patten-based ARIMA			Classical ARIMA
		Free	Synchronized	Congested	
Volume	RMSE	6.3941	6.1675	6.2578	8.0587
	MAPE(%)	10.16	11.56	10.70	12.39
Density	RMSE	2.1506	4.6159	3.6243	8.6902
	MAPE(%)	18.46	24.24	22.84	27.01



ARIMA



Pattern-based approach

Fig. 7 Actual (blue) versus predicted (red) values of volume and density in multiple-step-ahead prediction

4. Conclusion

This paper presents a methodological framework for enhancing accuracy of multiple-steps-ahead traffic flow prediction in transitional conditions. Forecasting is based on flow-density relation of traffic flow to classify traffic states and pattern matching technique to identify current traffic pattern, as well as classical ARIMA models to pattern-based prediction. Findings indicate that the proposed framework predicts the traffic flow for multiple steps into the future with enhanced accuracy when compared to classical ARIMA models that only operate under time series consideration with respect to both the RMSE and MAPE.

In view of the above, the proposed framework presents several interesting features: first, it improves on performance of short-term traffic flow forecasting under heterogeneous conditions. Second, it is consistent with the variable evolution of traffic flow and can claim applicability in different traffic flow conditions. Third, it can claim transferability in that it is based on the joint consideration of volume and density relation and is detached from any site specific geometric characteristics. Finally, the accurate prediction framework has limited dependence on the near past traffic data that can readily be integrated into an intelligent transportation system.

Acknowledgments: This research is supported by the grants NSFC (61374076, 61111130119), and the Doctoral Fund of Ministry of Education of China (No. 20091103120008).

References

[1] Billy M Williams, Lester A Hoel. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results, Journal of Transportation Engineering, 2003, No. 6, Vol. 129, pp. 664-672.

[2] Smith BL, Williams BM, Oswald RK. Comparison of parametric and nonparametric models for traffic flow forecasting, Transportation Research C, 2002, Vol. 10, pp. 303-321.

[3] Stathopoulos A, Karlaftis MG. A multivariate state space approach for urban traffic flow modeling and prediction, Transportation Research C, 2003, Vol. 11, pp. 121-135.

[4] Smith BL, Demetsky MJ. Traffic flow forecasting comparison of modeling approaches, Journal of Transportation Engineering, 1997, No. 4, Vol. 123, pp. 261-266.

[5] Okutani I, Stephanedes Y. Dynamic prediction of traffic volume through kalman filtering theory, Transportation Research B, 1984, Vol. 18, pp. 1-11.

[6] Whittaker J, Garside S, Lindveld K. Tracking and predicting a network traffic process, International Journal of Forecasting, 1997, Vol. 13, pp. 51-61.

[7] Wang Y, Papageorgiou M. Real-time freeway traffic state estimation based on extended kalman filter: a general approach,Transportation Research B, 2005, Vol. 39, pp. 141-167.

[8] Yang F, Yin Z, Liu HX, Ran B. Online algorithm for recursive short-term traffic prediction. In transportation research record, Journal of the Transportation Research Board, 2004, No. 1879, pp. 1-8.

[9] Vlahogianni EI, Golias JC, Karlaftis MG. Short-term traffic forecasting: overview of objectives and methods, Transport Reviews, 2004, No. 5, Vol. 24, pp. 533-557.

[10] Huynh N, Mahmassani HS, Tavana H. Adaptive speed estimation using transfer function models for real-time

dynamic traffic assignment operation. In transportation research record, Journal of the Transportation Research Board, 2002, No. 1783, pp. 55-65.

[11] Mecit Cetin, Gurcan Comert. Short-term traffic flow prediction with regime switching models, In transportation research record, Journal of the Transportation Research Board, 2006, No. 1965, pp. 23-31.

[12] Kerner BS. On-ramp metering based on three-phase traffic theory - part III, Traffic Engineering and Control, 2007, No. 3, Vol. 48, pp. 114-120.

[13] Nihan NL, Holmesland KO. Use of the Box and Jenkins time series technique in traffic forecasting, Transportation, 1980, No. 2, Vol. 9, pp.125-143.

[14] Vlahogianni EI. Karlaftis MG, Golias JC. Temporal evolution of short-term urban traffic flow: A non-linear dynamics approach, Computer-Aided Civil and Infrastructure Engineering, 2008, No. 7, Vol. 23, pp. 536-545.

[15] Wei W. Time Series Analysis Unvaried and Multivariate Methods, Addison-Wesley, Inc, Boston, Mass, 1990.

[16] Williams BM. Multivariate vehicular traffic flow prediction: evaluation of arimax modeling. In transportation research record, Journal of the Transportation Research Board, 2001, No. 1776, pp. 194-200.