2013 IEEE Intelligent Vehicles Symposium (IV)
June 23-26, 2013, Gold Coast, Australia

Short-Term Traffic Volume Prediction Using Classification and Regression Trees

Yanyan Xu, Qing-Jie Kong and Yuncai Liu

***Abstract*— Accurate short-term traffic flow prediction plays a fundamental role in intelligent transportation systems (ITS), e.g. advanced traffic management systems (ATMS). To generate accurate short-term traffic volume, nonparametric models have gained credit from quantities of researchers. On the basis of the common thought that future traffic states can be predicted according to the similar states in the historical traffic data, this paper presents a novel nonparametric-model-based method to predict the short-term traffic volume. The applied nonparametric model is the classification and regression trees (CART) model. In the application, the CART model first classifies the historical traffic states into plentiful categories. Afterwards, the linear regression model is built corresponding to each traffic state pattern. Finally, the model predicts the short-term traffic state through clustering the current state vector into the most congenial historical pattern and regression model. In the experiments, the proposed method is tested by using the 15 minutes average traffic volumes on freeways and is compared with the classic nonparametric methods k-nearest neighbors (k-NN) model, and the parametric method Kalman filter model. The results indicate that the CART-based prediction method outperforms the k-NN and Kalman filter methods in both the mean absolute percentage error and the mean absolute scaled error.**

I. INTRODUCTION

In the last several decades, the flourishing development of information technology has promoted the emergence and evolution of intelligent transportation systems (ITS) in modern traffic. As a major component of ITS, advanced traffic management systems (ATMS) focus on making efficient policies to reduce traffic congestion and to raise the road utilization ratio. Among the numerous elements of ATMS, short-term traffic flow prediction is always an essential technique but is difficult to be solved. Also, in traveler information service systems (TISS), accurate short-term traffic flow prediction can provide reliable guarantee of optimized guidance for travellers [1].

Since the early 1980s, based on the elaborate historical data, an increasing number of researchers attempt to predict traffic flow in less an hour instead of day time. In these works, the proposed methods mainly involve two categories: parametric-model-based and non-parametric-model-based techniques. At the beginning of ITS study, several forms of parametric models were applied into short-term traffic flow prediction, including autoregressive integrated moving average (ARIMA) models [2], [3], Kalman filtering [4], etc., as well as their alterations, e.g. subset autoregressive moving average models [5], seasonal ARIMA [6]. However, such parametric prediction models can not achieve satisfying outcomes, due to their unstable approximations to the nonlinear traffic conditions such as extreme peaks and rapid fluctuations.

Along with the evolution of the intelligence computation technologies during the last decade, various sorts of nonparametric algorithms were proposed to settle the issue of short-term traffic prediction. The nonparametric algorithms can often generate more stable and accurate prediction results than the parametric methods, owning to their plenty use of the historical traffic states [7]. One basic implementation of nonparametric regression is the k nearest neighbors (k-NN) algorithm [8], which predicts the short-term traffic state using its k most analogous historical states. In addition, with the rapid progress of data mining and machine learning since 1990s, a increasing number of intelligent algorithms were employed to build the relationships between the future traffic state with the states in the data sets. For instance, Dougherty *et al.* [9] developed an artificial neural network (ANN) approach to forecast the short-term inter-urban traffic. Fuzzy logic-based method was also applied into the traffic flow prediction [10]. As advanced machine learning methods, SVM [11], Gaussian Processes [12], etc., have also been extensively used in short-term traffic flow prediction.

The aim of this work is to find an accurate and efficient model to predict the traffic flow at a single location. Despite studying spatio-temporal models is much more popular among the researchers at present, the traffic flow prediction at a single station should not be ignored on whether freeways or urban road networks. In contrast with the spatio-temporal methods, the data fed into the single-location prediction model is easy to acquire and the model is more practical. Moreover, a number of algorithms based on the spatio-temporal model originate from the single-location model. For instance, Min and Wynter [13] exploited a multivariate spatial-temporal autoregressive model on the basis of Vector-ARMA technology. Their model took into account the spatial characteristics of the road network in a way that reflects not only the distance but also the average speed on the links. Additionally, boundary points or locations that serve as entrances to a network are isolated without any upstream detectors. This point was exemplified in Xu *et al.*'s [14] work, which assumed the future input flow rate

*This work is partly supported by China NSFC Program under Grant 61104160, National 863 Key Program under Grant 2012AA112307, and Shanghai STCSM Program under Grant 11231202801.

Yanyan Xu and Yuncai Liu are with the Department of Automation, Shanghai Jiao Tong University & Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai, 200240, China. Email: {*xustone, whomliu*}@*sjtu.edu.cn*. Qing-Jie Kong is with State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China, Email: qjkong.cas@gmail.com

was known. Consequently the traffic flow prediction at a boundary location is quite essential for the prediction for the whole network [7]. Therefore in this paper, we focus on the study on accurately and efficiently predicting traffic flow at single locations.

As Smith *et al.* [7] described, the future traffic states can be predicted through identifying the clusters of data space with the behavior similar with the current traffic state at a certain forecasting interval. Therefore, we investigate the application of classification and regression trees (CART) model to the single station short-term traffic volume prediction in our project, which is also a sort of nonparametric method. During the CART model building, we first classify the historical traffic states into a great number of patterns using the decision trees technology. Subsequently, the linear regression models are founded and the weights are stored in the leaf nodes of the trees model. Finally, the model predicts the future traffic state through assigning the current state vector to the most congenial historical pattern and regression model. In the experiment, the 15 minutes average traffic volumes obtained from five locations in Portland-area freeways are employed to evaluate the proposed CART-based prediction model.

The remainder of this paper is organized as follows: Section II presents the short-term traffic volume prediction based on the classification and regression trees model; Section III expounds the experiment traffic data sets and analyzes the test results; finally, some concluding remarks and directions for the future work are given in Section IV.

II. CART-MODEL-BASED TRAFFIC VOLUME PREDICTION

Similar to some other classic nonparametric prediction methods such as k-NN [7], four main challenges should be considered when developing a new classification model. They are definition of an appropriate state space vector, definition of various traffic patterns recognized from the historical observations, model of finding the best corresponding pattern to the current traffic state vector, and selection of a forecast generation method that can give the current conditions and the corresponding traffic pattern. The frame chart of the algorithm is shown in Figure 1.

As traffic volume is a sort of time series in nature, the state vector or independent variable input to the model can be presented as equation (1) shows.

$$x(t) = [V_t, V_{t-1}, \ldots, V_{t-d}] \tag{1}$$

where d is an appropriate number of time lags; V_t denotes the traffic volume during the current time interval; V_{t-1} is the traffic volume during the previous time interval.

Once the historical data set has been founded by using the foregoing state vectors, the classification model can be employed to cluster the traffic state vectors into multifarious traffic patterns, which are related to the values of the traffic volumes and their variation tendency. Essentially, the traffic patterns can not only demonstrate the change of values but also denote the traffic states (e.g., free flowing, rising up, dropping down, and congestion) during the $d+1$ intervals.

Based on core idea mentioned above, short-term traffic volumes can be predicted using the current traffic state vector and the nearest traffic patterns in the historical data set. In this section, the classification and regression trees (CART) [15] is selected to classify the historical observations into various traffic patterns and built a prediction model with each pattern.

A. Model of Classification and Regression Trees

CART is one of the most powerful data mining and machine learning methods for constructing prediction models from data. For the issue of traffic volume prediction, the leaf nodes in the tree structure represent a corresponding linear regression model with fully trained coefficients and the branches represent the conjunctions of the features that lead to those target variables.

1) Decision Tree: In the issue of traffic prediction, the target can be interpreted as creating a model that predicts the target variable y based on several input variables x_i that belong to a state vector X. Each interior node corresponds to one of the input variables, and there are edges to the sub-nodes for each of the possible input variables. Each leaf represents a value of the target variable, which is given by the values of the input variables through the path from the root to the leaf. Decision trees are represented by a set of questions which split the learning samples into smaller and smaller parts.

Suppose $X_i = \{x_1, x_2, \ldots, x_p\}$ is the input corresponding to the dependent variable y_i in the model. The set $S_{train} = \{X_1, X_2, \ldots, X_{N_d}\}$ denotes the training data set to be clustered in our experiment, where the capacity of the training data set equals to N_d. The purpose of the decision tree learning is to cluster the training data set into various patterns.

2) Tree Growing: Essentially, tree growing is to split the decision node. The branches stop growing when the splitting achieves a leaf node. Each decision node implements a test function $f(S)$ to split the corresponding data space S into two subsets. Therefore, for decision node m, we have

$$f_m(S_m) : x_i^j \geq \theta_{m0} \tag{2}$$

where x_i^j denotes the ith feature in data space S_m; θ_{m0} is the best split threshold that leads to a minimum error. Thence, the best binary split on the decision node divides the input space into two: $L_m = \{S_m | x_i^j > \theta_{m0}\}$ and $R_m = \{S_m | x_i^j \leq \theta_{m0}\}$.

Sequentially, the mean square error is utilized to measure the goodness of a split. Let $\hat{y}_m$ represent the predicted value of y_m. Then we have

$$E_m = \frac{1}{N_m} \sum_t (y_{m,t} - \hat{y}_{m,t})^2 \tag{3}$$

where N_m is the capacity of the data space on node m; the predicted value $\hat{y}_m$ is obtained by the weights: $\hat{y}_m = X_m \times w_m$. Similarly, the error after splitting the space S_m

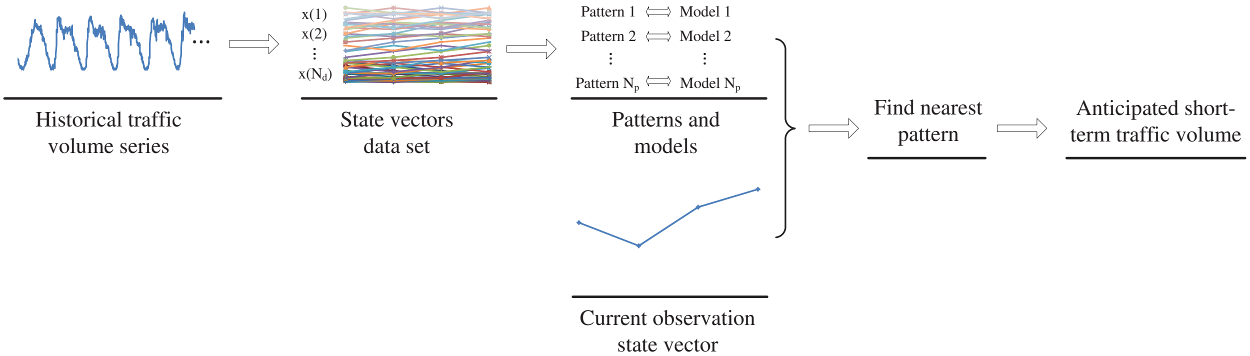

Fig. 1. Flowchart of the classification-based traffic prediction method.

into subsets L_m and R_m can be calculated by:

$$\begin{aligned} E'_m = & \frac{1}{N_{m,L}} \sum_t (y_{m,L,t} - \hat{y}_{m,L,t})^2 \\ & + \frac{1}{N_{m,R}} \sum_t (y_{m,R,t} - \hat{y}_{m,R,t})^2 \end{aligned} \tag{4}$$

The decrease in error is given as the difference between Equation 3 and Equation 4. The optimal split feature and threshold θ_{m0} are determined through calculating the minimal decrease error.

Once a branch reaches a node that meets the predefined conditions (the lowest predictive error or the minimal capacity), it stops to grow and the node is labeled as a leaf node. Afterwards, the data space in the leaf node is fitted with the linear regression method.

3) Tree Pruning: Frequently, a node doesn't split any more, if the training number is smaller than a certain percentage of the training set regardless of the impurity or error. This is because of the idea that any decision based on too few instances can cause variance and thus generate errors.

In the practice of the traffic volume prediction, the error decrease threshold θ_{error} and the least number of state vectors contained in a leaf node θ_N (the minimal capacity of each leaf node) are preestablished. Then, the tree stops growing when either of the two conditions is contented (the decrease error E_m is less than θ_{error} or capacity of the node N_{leaf} is not more than θ_N).

4) Prediction Generation: After the CART model is built by using the training data, we can employ the model to predict the short-term traffic volume with the newcome traffic state vector. The most straightforward approach to generating the prediction y from the independent variables X is to compute a simple linear regression model of the independent variables that belong to the traffic pattern corresponding to the given state vector. In other words, the training data set has been classified into leaf notes with corresponding space data $\{\boldsymbol{X}_i, \boldsymbol{y}_i\}$. The regression model $M_i : \boldsymbol{y}_i \leftarrow f(\boldsymbol{X}_i)$ is built in the leaf note and denotes that we use vector $\boldsymbol{X}_i$ to fit $\boldsymbol{y}_i$, as Equation 5 shows.

$$\begin{cases} M_1 : \boldsymbol{y}_1 \leftarrow f_1(\boldsymbol{X}_1) \\ M_2 : \boldsymbol{y}_2 \leftarrow f_2(\boldsymbol{X}_2) \\ \quad\vdots \\ M_n : \boldsymbol{y}_n \leftarrow f_n(\boldsymbol{X}_n) \end{cases} \tag{5}$$

Finally, the newcome state vector is classified using the existing CART model. Thence, the future short-term traffic volume can be obtained by using the corresponding weight.

$$y_{new} = X_{new} \times w_m \tag{6}$$

B. Model Design For Traffic Volume Prediction

As we are using the trees model to predict the traffic flow, the nearest pattern of a state vector can be determined through a series of decisions. Therefore, the remaining issues should be considered are the definition of the state vector and the regression model for prediction.

1) Definition of the state vector: The definition of the data space S is an important issue should be discussed before modeling, including the state vector X and the size of time lag d. In this paper, we focus on the change of traffic volume at a single station along the time line, wherefore the state vector X is defined as $\{v_t, v_{t-1}, \ldots, v_{t-d}\}$. Moreover, the selection of the time lag d influences the accuracy of the CART model and the prediction results. If d is too large, the state series can not reflect the sudden change of the traffic volumes. On the contrary, a too small state vector also can not represent the variation tendency of the traffic states. Therefore, we define the time lag d equals to 3 in our research. The state vector for training the CART model is shown as follows:

$$V = \{v_{t-3}, v_{t-2}, v_{t-1}, v_t, v_{t+1}\} \tag{7}$$

2) Regression model: For the purpose of obtaining precise predictive value of traffic volume, firstly, the current traffic state vector are assigned to a certain leaf node by the CART model, and then the desired value can be calculated through the regression model corresponding to the leaf node. The regression model $M_i : \boldsymbol{y}_i \leftarrow f(\boldsymbol{X}_i)$ corresponding to the traffic pattern should be defined to generate the predictive value. In our paper, the linear regression model is addressed to fit the relationship between the historical and the desired traffic states. The weights stored in the leaf nodes are calculated as follows:

$$w_i = (\boldsymbol{X}^T \times \boldsymbol{X}_i)^{-1} \times (\boldsymbol{X}_i^T \times \boldsymbol{y}_i) \tag{8}$$

III. Experimental Data and Test Results

A. Data Description

In this study, we employed the traffic volume data obtained from the PORTAL FHWA Test Data Set [16] to evaluate the proposed CART-based prediction model. PORTAL receives the traffic volume every 20 seconds from the dual-loop detectors, which are installed in the main line lanes and on-ramps on the Portland-area freeways.

These data sets are from five different stations located on Interstate 5 (I-5), Interstate 84 (I-84), Interstate 405 (I-405), United State 26 (US-26), and Interstate 205 (I-205). Figure 2 shows the distribution of those stations on the freeways. More details about the number of lanes on the freeways, the specific locations, and the lengths of the stations are illustrated in Table I.

The traffic volume data sets that we used in this study were collected between May 1 and June 4, 2011. The data of the previous four weeks were regarded as the training data set; the data of last one week were taken as the test data set to evaluate the developed prediction model. In addition, the raw 24-hour traffic volume data were aggregated into 15-minute intervals in each data set. The traffic volume is formatted as volume per lane per hour (VPLPH).

TABLE I

Properties of selected stations

Station	Freeway	Lanes	Milepost	Length
1. 33rd WB	I-84 WB	4	2.1	1.04
2. Alberta St SB	I-5 SB	3	304.08	0.48
3. Couch to I-405 NB	I-405 NB	2	2.3	0
4. Skyline Rd EB	US 26 EB	3	71.37	2.18
5. Glisan to I-205 SB	I-205 SB	3	21.12	1.53

Figure 3 shows the traffic volume at the 33rd WB station on freeway I-84 from 1 May to 28 May. From the curve of the traffic volume, we can effortlessly find that the traffic states of freeways on weekends are more smooth than that on weekdays. However, in our work, we make no difference on model design and traffic prediction between weekdays and weekends.

B. Test Results and discussions

Besides the commonly used mean absolute percentage error (MAPE), another different measure for forecasting error analysis is employed in this study to evaluate the performance of the proposed model, which is the mean absolute scaled error (MASE). MASE is a forecast accuracy measure proposed by Rob Hyndman [17]. Different from the traditional error measures, MASE is a sort of scaled error that takes account of the gradient of the actual values, therefore, in MASE, the prediction accuracy can be compared not only between different methods for same data sets but also between methods for different freeway stations. Similarly, a smaller MASE indicates better prediction. MAPE and MASE are defined as follows:

$$MAPE = \frac{1}{K}\sum_{k=1}^{K}\left|\frac{V_k - \hat{V}_k}{V_k}\right| \tag{9}$$

$$MASE = \frac{1}{K}\sum_{k=1}^{K}\left|\frac{V_k - \hat{V}_k}{\frac{1}{K-1}\sum_{k=2}^{K}|V_k - V_{k-1}|}\right| \tag{10}$$

where K is the total number of the anticipated traffic volumes in the test data set; V_k denotes the actual average traffic volume in 15 minutes; $\hat{V}_k$ is the prediction value produced by the proposed prediction model or the compared models.

Base on the five data sets listed in Table I, we compared the performances of the CART prediction model and the other two classic prediction methods, k-NN model [7] and direct Kalman filter approach. Specially, the prediction results at station 33rd WB are selected for a deeper analysis. For the purpose of more impartial evaluation, the minimal capacity of the each leaf node in CART model θ_N and the number of neighbors in k-NN model are both set as 20; the k-NN method is employed using the same state vector with the CART model.

Figure 4 shows the 15 minutes average traffic volumes observed at station 33rd WB during a week and the prediction curves generated by the three methods. Obviously, the two nonparametric methods outperform the Kalman filter method to a large extent, especially at the unstable rush hours. This phenomenon is benefited from the periodicity and similarity of the traffic states on the freeways. Furthermore, the nonparametric methods including the CART and k-NN approaches would predict more accurately while fed with more complete and plenty training data sets.

In Figure 4, the CART-based model and k-NN model both fit the actual values very well. In order to further analyze the performances of the CART based and k-NN model, we randomly selected one day in the test result for further analysis, as Figure 5 shows. We can find that the CART-based model surpasses the k-NN at most of the time in this day.

The error analysis of the three methods is given in Table II and III. The relative percentage improvements of the CART-based model over the k-NN and Kalman filter methods are also listed in the table. The error results illustrate that the CART-based model predicts more accurate volumes than the other two methods at all of the five stations. The average promotions over the k-NN are equal to 10.472% and 11.556% in MAPE and MASE, respectively. As to the Kalman filter method, the promotions are 30.104% and 34.812%.

Although the CART and the k-NN based model are both nonparametric prediction methods, they are different when applied in practice. The CART-based method builds the trees model offline, which is another advantage compared to the k-NN model besides the predictive accuracy. Particularly,

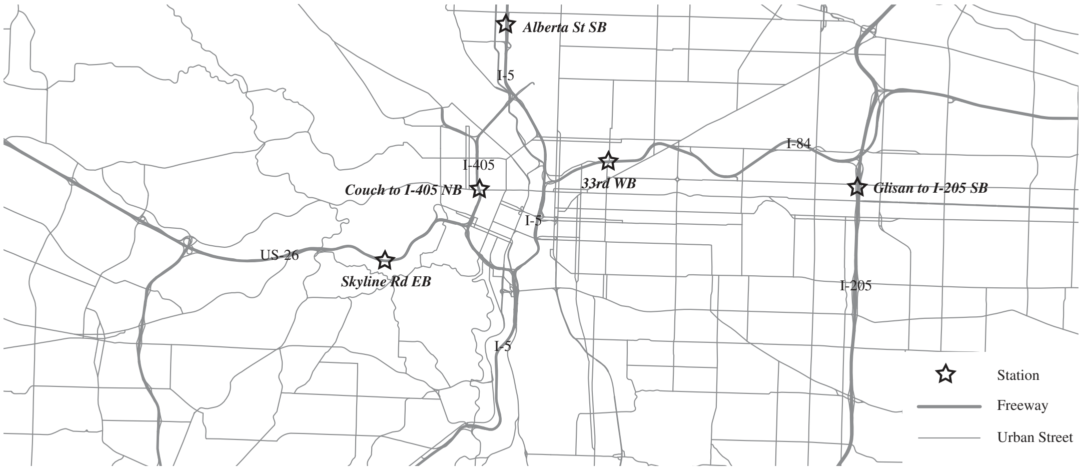

Fig. 2. Locations of stations on the freeways

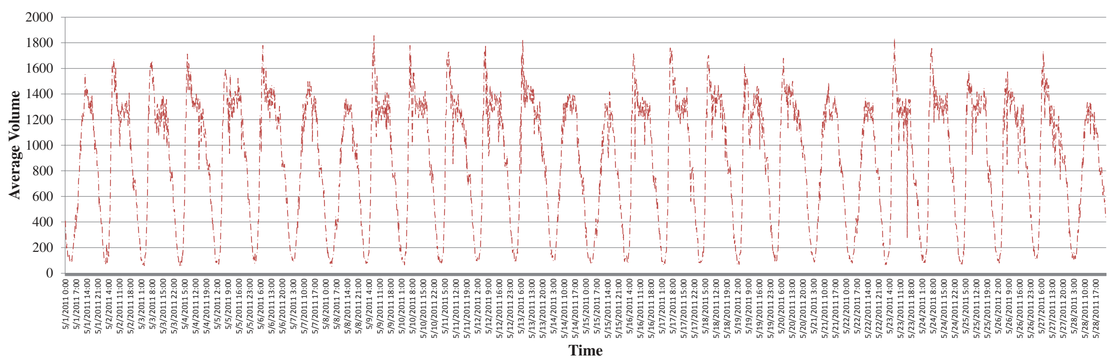

Fig. 3. Traffic volume of 33rd WB from 1 May to 28 May (VPLPH)

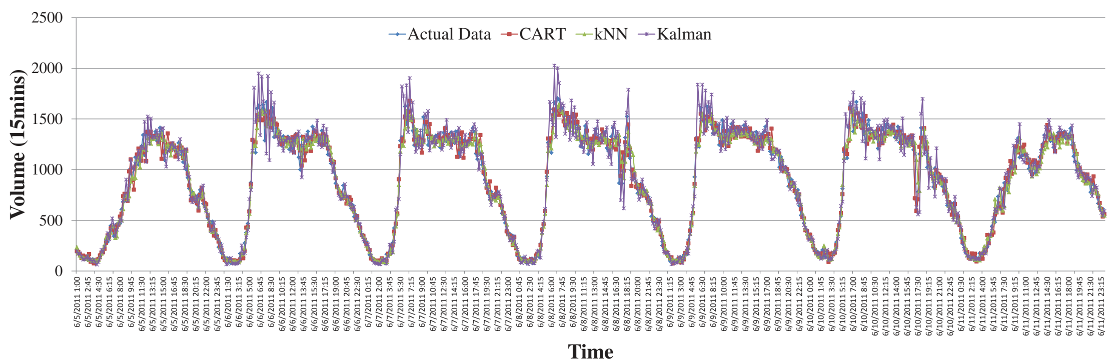

Fig. 4. Prediction of 15 minutes traffic volume at station 33rd WB from June 5 to 11.

when the historical data set is excessively huge, the CART-based model can easily find the optimal regression model corresponding to the current traffic state vector. However, the k-NN model has to search the huge historical data set to find the k nearest neighbors for prediction.

IV. Conclusions

This paper focuses on utilizing the classification and regression trees model to predict the short-term traffic volume at single locations. Before building the CART model, the traffic volume series are shaped into the state vectors and compose a data space. As training data set, the historical state data space is used to build a CART model. In the model building process, the data space is first clustered into a number of different subsets. In each subset, the linear regression model is applied to build the relationship between the historical state vector and the desired value subsequently. After such classification and regression model has been created, the future traffic state can be predicted through assigning the current traffic state vector to an optimal subset and calculating by using the corresponding regression model. The experiment employed the traffic volumes collected from five freeway locations distributed in the Portland-area. Based on the data sets, another nonparametric prediction method k-

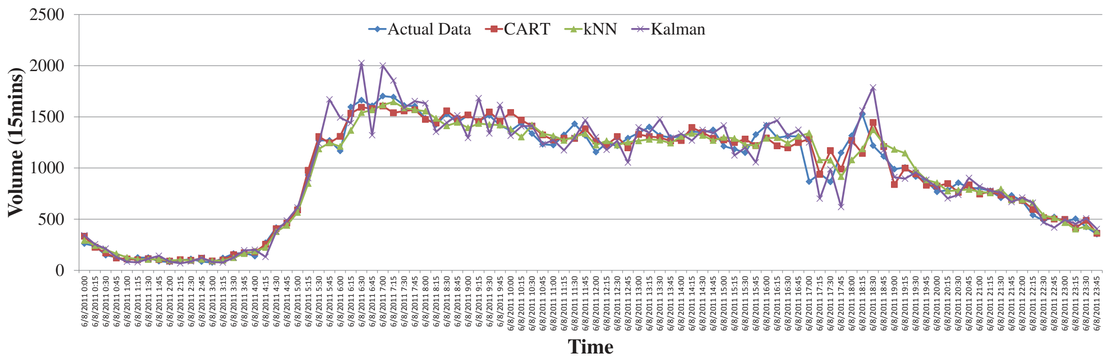

Fig. 5. Prediction of 15 minutes traffic volume at station 33rd WB on June 8.

Station	MAPE(%)			Promotion (%)	
	CART	k-NN	Kalman	p1*	p2*
1.	7.75	9.31	11.94	16.76	35.09
2.	10.99	11.57	14.8	5.01	25.74
3.	9.16	9.99	12.9	8.31	28.99
4.	7.54	8.45	11.17	10.77	32.5
5.	7.23	8.17	10.07	11.51	28.2
Average	**8.534**	**9.498**	**12.176**	**10.472**	**30.104**

* $p1 = |MAPE_{CART} - MAPE_{k-NN}|/MAPE_{k-NN}$
$p2 = |MAPE_{CART} - MAPE_{Kalman}|/MAPE_{Kalman}$

TABLE II
COMPARISON OF MAPE OF THE THREE MODELS.

Station	MASE			Promotion (%)	
	CART	k-NN	Kalman	p1	p2
1.	0.734	0.877	1.2	16.31	38.83
2.	0.832	0.865	1.24	3.82	32.9
3.	0.775	0.881	1.15	12.03	32.61
4.	0.742	0.827	1.17	10.28	36.58
5.	0.789	0.932	1.18	15.34	33.14
Average	**0.7744**	**0.8764**	**1.188**	**11.556**	**34.812**

TABLE III
COMPARISON OF MASE OF THE THREE MODELS.

NN and the classic parametric method Kalman filter model are compared with the CART-based model. The comparison results indicate that the CART-based model is able to generate more accurate predictive values than the k-NN and Kalman filter in both MAPE and MASE.

Another natural advantage of the CART-based model is that the model can be built offline in contrast to k-NN. Therefore, the online task of the CART-based method is nothing but predicting the future traffic volume by using the trained model, which is more efficient and suitable for the practical application. Furthermore, the CART-based prediction model is easy to be extended due to its learning procedure. Therefore, further studies can attempt to extend the CART model to a spatio-temporal method.

REFERENCES

[1] E. I. Vlahogianni, J. C. Golias, and M. G. Karlaftis, "Short-term traffic forecasting: Overview of objectives and methods," *Transport Reviews*, vol. 24, no. 5, pp. 533–557, 2004.

[2] M. Ahmed and A. R. Cook, "Analysis of freeway traffic time-series data by using box-jenkins techniques," *Transportation Research Board*, vol. 722, pp. pp. 1–9, 1979.

[3] M. V. D. Voort, M. Dougherty, and S. Watson, "Combining kohonen maps with arima time series models to forecast traffic flow," *Transportation Research Part C: Emerging Technologies*, vol. 4, no. 5, pp. 307–318, 1996.

[4] I. Okutani and Y. J. Stephanedes, "Dynamic prediction of traffic volume through kalman filtering theory," *Transportation Research Part B: Methodological*, vol. 18, no. 1, pp. 1–11, 1984.

[5] S. Lee and D. B. Fambro, "Application of subset autoregressive moving average model for shortterm freeway traffic volume forecasting," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1678, pp. pp. 179–188, 1999.

[6] B. M. Williams, P. K. Durvasula, and D. E. Brown, "Urban freeway travel prediction: application of seasonal arima and exponential smoothing models," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1644, pp. 132–141, 1998.

[7] B. L. Smith, B. M. Williams, and R. K. Oswald, "Comparison of parametric and nonparametric models for traffic flow forecasting," *Transportation Research Part C: Emerging Technologies*, vol. 10, no. 4, pp. 303–321, 2002.

[8] G. A. Davis and N. L. Nihan, "Nonparametric regression and short-term freeway traffic forecasting," *ASCE Journal of Transportation Engineering*, vol. 17, no. 2, pp. 168–178, 1991.

[9] M. S. Dougherty and M. R. Cobbett, "Short-term inter-urban traffic forecasts using neural networks," *International Journal of Forecasting*, vol. 13, no. 1, pp. 21–31, 1997.

[10] L. Dimitriou, T. Tsekeris, and A. Stathopoulos, "Adaptive hybrid fuzzy rule-based system approach for modeling and predicting urban traffic flow," *Transportation Research Part C: Emerging Technologies*, vol. 16, no. 5, pp. 554–573, 2008.

[11] Y. Zhang and Y. Xie, "Forecasting of short-term freeway volume with v-support vector machines," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2024, pp. 92–99, 2008.

[12] Y. Xie, K. Zhao, Y. Sun, and D. Chen, "Gaussian processes for short-term traffic volume forecasting," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2165, pp. 69–78, 2010.

[13] W. Min and L. Wynter, "Real-time road traffic prediction with spatio-temporal correlations," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 4, pp. 606–616, 2011.

[14] Y. Xu, Q.-J. Kong, S. Lin, and Y. Liu, "Urban traffic flow prediction based on road network model," in *Proc. 9th IEEE International Conference on Networking, Sensing and Control*, Beijing, China, 2012, pp. 334–339.

[15] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*. Belmont, California: Wadsworth, 1984.

[16] (Portland State University. http://portal.its.pdx.edu/. Accessed May. 15, 2012) The portal fhwa traffic data set. Portland State University. [Online]. Available: http://portal.its.pdx.edu/

[17] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International Journal of Forecasting*, vol. 22, no. 4, pp. 679–688, 2006.