An Ensemble Cascading Extremely Randomized Trees Framework for Short-Term Traffic Flow Prediction

Fan Zhang[1,2], Jing Bai[3*], Xiaoyu Li[3] , Changxing Pei[1] and Vincent Havyarimana[4]
[1]State Key Laboratory of Integrated Service Networks, Xidian University
Xi'an 710071, China
[2]PLA 61267 Troops, Beijing, 100000, China.
[3]School of Artificial Intelligence, Xidian University, Xi'an, 710071, China.
[4]Department of Applied Sciences, Ecole Normale Supérieure
6983, Bujumbura, Burundi
[e-mail: fzhang09@126.com, xiaobaijing1982@163.com, xiaoyulsj427@163.com,
chxpei@xidian.edu.cn, havincent12@gmail.com]
*Corresponding author: Jing Bai

Abstract

Short-term traffic flow prediction plays an important role in intelligent transportation systems (ITS) in areas such as transportation management, traffic control and guidance. For short-term traffic flow regression predictions, the main challenge stems from the non-stationary property of traffic flow data. In this paper, we design an ensemble cascading prediction framework based on extremely randomized trees (extra-trees) using a boosting technique called EET to predict the short-term traffic flow under non-stationary environments. Extra-trees is a tree-based ensemble method. It essentially consists of strongly randomizing both the attribute and cut-point choices while splitting a tree node. This mechanism reduces the variance of the model and is, therefore, more suitable for traffic flow regression prediction in non-stationary environments. Moreover, the extra-trees algorithm uses boosting ensemble technique averaging to improve the predictive accuracy and control overfitting. To the best of our knowledge, this is the first time that extra-trees have been used as fundamental building blocks in boosting committee machines. The proposed approach involves predicting 5 min in advance using real-time traffic flow data in the context of inherently considering temporal and spatial correlations. Experiments demonstrate that the proposed method achieves higher accuracy and lower variance and computational complexity when compared to the existing methods.

---

1. Introduction

Over the past decades, traffic congestion has become increasingly serious. Dynamic traffic management is a common method for reducing congestion in fast developing intelligent transportation systems (ITSs) and advanced traffic management systems (ATMSs) [1][2][3][4][5][6]. Accurate and timely traffic flow information is crucial for traffic regulation and vehicular navigation. The short-term traffic flow prediction problem is to forecast the traffic flow at a road surveillance point for the near future with the current traffic flow and several sensor readings such as speed and occupancy. Most of the proposed traffic flow prediction methods[7][8][9] require considerable training time as well as high computational complexity under the assumption that the traffic flow data fluctuates in an acceptable range. Therefore, how to perform short-term traffic flow prediction under non-stationary conditions, such as the presence of car collisions, road congestion, device breakdown and complicated urban environments [10][11], is a challenging task. In addition, achieving high prediction accuracy as well as less time consumption and computational complexity during prediction is a critical issue.

1.1 Related Work

Several researchers have used a wide variety of approaches for predicting short-term traffic flow. Existing traffic flow prediction approaches can be divided into parametric approaches and nonparametric approaches. Parametric approaches are based on time-series methods, such as the seasonal autoregressive integrated moving average (SARIMA) model that is extended from the autoregressive moving average (ARMA) model and the autoregressive integrated moving average (ARIMA) model. They are founded on stochastic system theory and use the patterns of the temporal variation of traffic flow for prediction.

Tahmasbi et al.[12] adopted stochastic differential equations (SDEs) for traffic flow prediction and used Hull-White model to estimate the parameters theoretically. Y. Hou et al. [13], presented four traffic flow forecasting models for urban work zones using random forest, regression tree, multilayer feed-forward neural network, and nonparametric regression. The results showed that the random forest model yielded the most accurate traffic flow forecasts. Recently, some researchers realized that traffic data are increasing exponentially [14] and proposed a deep learning approach using stacked autoencoders (SAEs) to learn generic traffic flow features considered that existing traffic flow prediction are shallow in architecture and training in a greedy layerwise fashion. Meanwhile, W. Huang et al. [15] proposed a deep architecture that consists of a deep belief network (DBN) at the bottom and a multitask regression layer at the top. Both deep learning models have a superior performance on the basis of experiments. However, they often require a high computational complexity, and system deployment is complicated.

Because traffic data have characteristics of real-time variability and high frequency, several online and incremental learning frameworks have been proposed. For example, incremental methods based on k-nearest neighbors (KNNs), artificial neural networks (ANNs), support vector regressions (SVRs), and deep learning have been constructed. M. Castro-Neto et al. [16] proposed the online-SVR (OL-SVR) approach to predict short-term freeway traffic flow under typical and atypical conditions from the standpoint of applicability and showed that OL-SVR is suitable and useful in real-world operations and has better performance than other models. K. Y. Chan et al. [17] proposed a neural network (NN) model and used hybrid

exponential smoothing and Levenberg-Marquardt algorithm to improve the generalization capabilities. K. Kumar et al. [18] applied ANNs for short-term traffic flow prediction and obtained consistent performance in time intervals from 5 min to 15 min .

In summery, in the existing methods, either the computation cost is high or they have difficulty in handling the real-time variability and high-frequency problem of traffic flow prediction.

1.2 Contributions

In this paper, we propose an ensemble cascading prediction framework based on extremely randomized trees (extra-trees) and a boosting technique with high prediction accuracy and low computational complexity. The extra-tree method is a tree-based ensemble method. Its randomness extends the method in which splits are computed. As in random forests, a random subset of candidate features is used, but instead of looking for the most discriminative thresholds, thresholds are drawn at random for each candidate feature, and the best of these randomly generated thresholds is selected as the splitting rule. This usually reduces the variance of the model slightly more, which is more suitable for the prediction of short-term traffic flow prediction under non-stationary conditions. Moreover, the extra-trees algorithm uses boosting ensemble technique averaging to improve the predictive accuracy and to control overfitting.

The contributions of this paper are summarized as follows.

1) Ensemble learning of the fast randomized tree methods for traffic flow prediction incorporates the advantages of low bias-variance and computational complexity. Compared with other tree-based ensemble algorithms, our prediction framework is more accurate and efficient.

2) Our prediction framework is well fitted in both stationary and non-stationary conditions with breakdown data caused by detectors.

3) We employ an ensemble mechanism to improve our traffic flow prediction performance. We use extra-trees as fundamental building blocks in boosting committee machines, regressors are combined using the weighted median, and predictors that have more accurate results are weighted more heavily. Therefore, our prediction framework obtains higher accuracy.

The remainder of this paper is organized as follows. In Section 2, we introduce the fundamental building block of our framework, the extra-trees algorithm. In Section 3, we introduce our proposed ensemble cascading framework using the boosting technique. In Section 4, the data description, experimental design and results are described. The conclusions and future work are presented in Section 5.

2. Extremely Randomized Trees

In our traffic flow prediction framework, we mainly use extremely randomized trees (extra-trees) as the fundamental building blocks due to their higher accuracy, lower computational complexity and variance compared with other conventional methods such as decision trees (DT), AdaBoost regressions (ABR), and support vector regressions (SVR).

The Extra-Trees algorithm is an integration method which combines multiple unpruned decision or regression trees according to the classical top-down procedure. According to [19], extra-trees consists of strongly randomizing both attribute and cut-point choice while splitting

a tree node. In the extreme case, it constructs completely randomized trees whose structures are independent of the output values of the learning sample. The extra-trees algorithm avoids overfitting and has better accuracy and stability relative to the use of a single decision tree. There are two main differences with other tree-based ensemble methods. First, the extra-trees algorithm uses the whole learning sample rather than a bootstrap replica to grow the trees to minimize bias. Second,it randomly chooses cut-points to split nodes, which reduces variance more strongly than the weaker randomization schemes used by other methods [19].

The extra-trees algorithm, in essence, is a tree-based averaging algorithm. It is a perturb-and-combine technique specifically designed for trees, which means that a diverse set of regressors are generated by introducing randomness in the regressor construction. Then, the averaging prediction result of each regressor is taken as the prediction result of our ensemble algorithm. In extremely randomized trees, the randomness extends the way that splits are computed. In our traffic flow prediction case, near-future traffic flow data are timely predicted based on historical and current data, such as traffic flow (f_i), speed (s_i) and occupancy (o_i). All of these data are numerical attributes. Some trees are constructed from the traffic flow data set $\{X, Y\}$ of the s-dimensional input while in training. We use historical traffic flow data (f_{k-1}) and current traffic flow data (f_k), speed (s_k) and occupancy (o_k) as inputs, and their corresponding output is f_{k+1}, which represents the near future traffic flow. The extra-trees building procedure is shown in Algorithm 1.

Algorithm 1 Extra-Trees Building Algorithm

Input: Traffic flow vector $\{X, Y\} = [f_{k-1}, s_k, o_k, f_k]$.

Output: Decision tree t.

1. **If** $|X| < n_{\min}$,or all candidate attributes are constant, or the output variable is constant **then**

2. Return a leaf labeled by average output.

3. **end if**

4. Randomly select K attributes, $K = n$ in regression.

5. Generate K splits $\{s_1, \ldots, s_K\}$ according to the maximal and minimal value of K attributes.

6. Select a split s_b with the best $Score_R(s_b, S)$

7. Split $\{X, Y\}$ into $\{X_l, Y_l\}$ and $\{X_r, Y_r\}$ according to s_λ .

8. Build t_l =build_an_ExtraTree(X_l, Y_l) and t_r = build_an_ExtraTree(X_r, Y_r).

9. Create a node with the split s_b, attach t_l and t_r , then return a tree t .

According to [11], there are three parameters in extra-trees algorithm. They are the number of trees of the ensemble M , the minimum sample size for splitting a node $n_{\min}$, and the

number of attributes randomly selected at each node K. In general, M is set to be 100, $n_{\min}$ is set to be 5 in regression, and K is set to be n, where n is the number of attributes.

We use the relative variance reduction, and then the score measure for a sample S and a split s is defined as

$$Score_R(s,S) = \frac{\text{var}(Y) - \frac{|Y_l|}{|X|}\text{var}(Y_l) - \frac{|Y_r|}{|X|}\text{var}(Y_r)}{\text{var}(Y)} \tag{1}$$

where $S = \{X,Y\}$, $|X|$ is the length of the vector X, and $\text{var}(Y)$ is the variance of the output available at the current node. We use the highest score to make an annotation for the current node. $\{X_l, Y_l\}$ and $\{X_r, Y_r\}$ are two subsets from traffic data $\{X,Y\}$, and we use them as input data to respectively construct the left t_l and the right t_r subtrees.

3. The Proposed Ensemble Cascading Framework

In this paper, we propose an ensemble cascading framework for traffic flow prediction. The fundamental building blocks are extra-trees, and the techniques we employ to improve our prediction performance are exponentially weighted moving average (EWMA) and AdaBoost.

3.1 Exponentially weighted moving average

Considering the time-varying characteristic of traffic flow, we first use EWMA algorithm to preprocess the traffic data to improve our prediction accuracy. It is a moving average method that applies weighting factors that decrease exponentially. The most recent data have the most important influence on the prediction procedure, and the older data also give some weights. Thus, we consider the EWMA algorithm to do the preliminary predictions, which helps us to obtain better performance. Normally, there are two parameters in the EWMA algorithm: the specified 'span' parameter s, and the decay parameter, α. The relationship of these two parameters are formulated as

$$\alpha = 2/(s+1) = 1/(1+c) \tag{2}$$

where c is the center of mass. Given a span, the associated center of mass is $c = (s-1)/2$.

3.2 Ensemble by Adaboost

AdaBoost is one of the boosting techniques that can reduce prediction errors. We use extra-trees as fundamental building blocks in AdaBoost committee machines, and the individual regressor is a weak regressor [20]. The AdaBoost algorithm trains each weak regressor sequentially. In our work, each extra-tree is trained on different subsets of the original traffic flow training set and gives different predictions. The predicted values that differ from the observed values are defined to be the most in error, and then the sampling probability is adjusted so that these values are more likely to be selected as members of the training set for the second extra-trees. Thus, the more difficult it is to predict, the more likely that it will appear in the training set. This makes our prediction more accurate. During prediction, the input vector is processed by all the extra trees. The outputs of the extra-trees are aggregated, and the average yields the final prediction result.

4. Experiments

4.1 Dataset description

The data we used for traffic flow prediction were collected from Caltrans Performance Measurement System (PeMS) database. PeMS is the most widely used dataset in traffic flow prediction. The traffic data, including traffic flow, speed, and occupancy, are collected every 30 s from over 15000 individual detectors, which are deployed statewide in freeway systems across California [21]. The collected data were aggregated as counts of cars into a 5-min period for each detector station. In our prediction case, we randomly chose five detector station, as shown in **Table 1**. We used the data collected from 2015-11-23 to 2015-11-30. Since we aim to predict the traffic for the next 5 min, the time for one day is divided into 288 intervals. The data of the first seven days were selected as the training set, including 2016 intervals, and the remaining one-day data were selected as the testing set.

Table 1. Detail Information of the Selected Stations

Freeway	ID	LAT	LON	LEN	ABS PM
I800-S	400312	37.6774	-122.1189	0.32	0.32
SR120-E	1004510	37.7831	-121.2338	1.93	3.83
US101-N	769346	34.1567	-118.4070	0.54	14.88
1210-W	770187	34.2597	-118.3297	1.54	11.1
I5-N	1122528	32.8976	-117.2253	0.51	30.1

4.2 Performance metrics

Typically,we use error measures to evaluate the performance of machine learning methods. To evaluate the effectiveness of our proposed model, we use three performance indexes: root mean square error (RMSE), mean absolute percentage error (MAPE), and mean absolute error. They are formulated as

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \frac{\left| f_i - f_i^{'} \right|}{f_i} \times 100 \tag{3}$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(\left| f_i - f_i^{'} \right| \right)^2} \tag{4}$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} \left| f_i - f_i^{'} \right| \tag{5}$$

where f_i is the observed traffic flow, $f_i^{'}$ is the predicted traffic flow, and n is the total number of observations.

4.3 Comparison results

In this subsection, we mainly discuss our experimental design and results. We first give an illustration of the real and predicted traffic flow for one day on I-5 North, as shown in **Fig. 1**. We can see that there are two rush hours in one day of traffic flow, from 5:00 AM to 10:00 AM and 5:00 PM to 10:00 PM. The purpose of the traffic flow prediction is to relieve traffic pressure during peak hours. Thus, for each of the five randomly selected freeways, we consider the rush hour from 5:00 AM to 10:00 AM as a typical scenario.
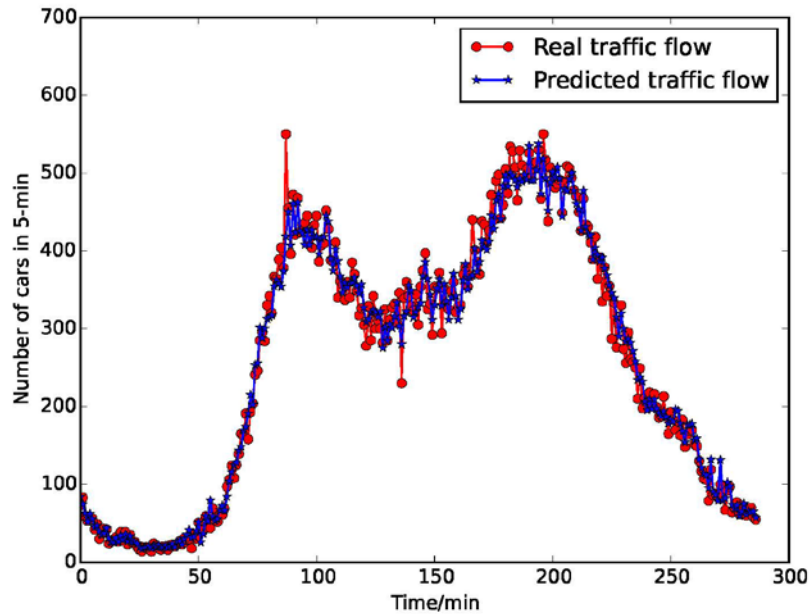
Fig. 1. The real and predicted traffic flow in one day on I-5North

To evaluate the performance of the proposed method, we compared it with the performance of other integrated learning frameworks, deep learning algorithms, and classical machine learning algorithms. Specifically, we compared the performance of the proposed EET model with that of the DT, ABR, and SVR methods and the LSTM [22-24].

The DT is constructed by a sequence of binary splits of the training set into terminal nodes. We set the n_{min} to 5; if the node size $N_t < n_{min}$, we declare the node terminal node. However, the DT has the disadvantage of high variance. The ABR is an ensemble method of DT, and we set the number of trees, M, to 100. This ensemble method reduces the variance of DT but is still not a satisfactory approach. For SVR, we employ a radial basis function as the kernel. Although SVR has been extensively used in the field of ITS, it has drawbacks as well. In our proposed EET model, we set M to 100, n_{min} to 5, and K to n. **Fig. 2** to **Fig. 6**. show the real and predicted traffic flow comparison of DT, ABR, SVR, LSTM and EET during the rush hour of 5:00 AM to 10:00 AM on US101-N. From the line chart, we can see that the real traffic flow value has the lowest point at time index 20. There may be a traffic jam at this point. It is obvious that only the proposed EET and LSTM have predicted this sudden drop. Other methods, including DT, ABR and SVR, are affected by this incident, so they deviate from the actual flow. Although the deep learning algorithm has strong feature extraction, the key to traffic flow prediction is to address the non-stationary property of traffic flow data. LSTM offsets many real values at adjacent moments. Only the proposed EET method is capable of eliminating this influence. The results illustrate that the prediction result using EET is relatively close to the real traffic flow, even at the non-stationary times.
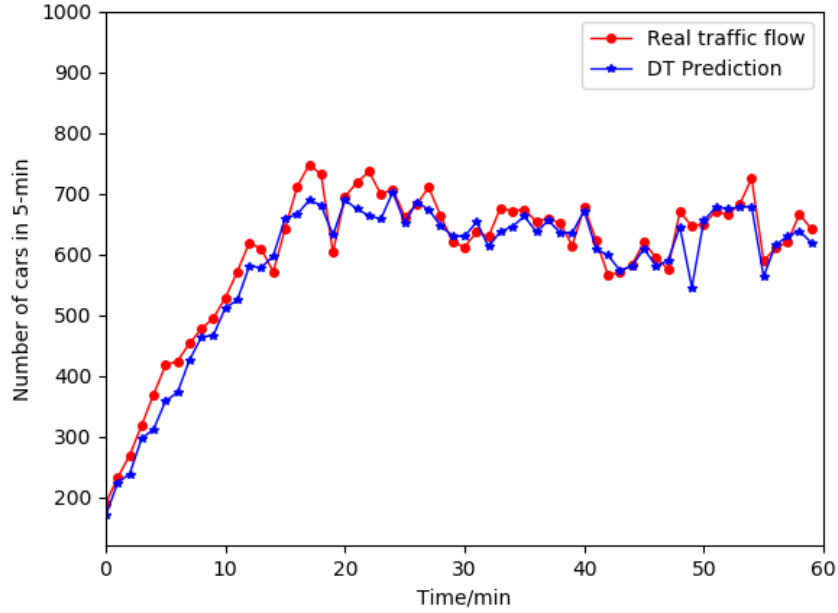
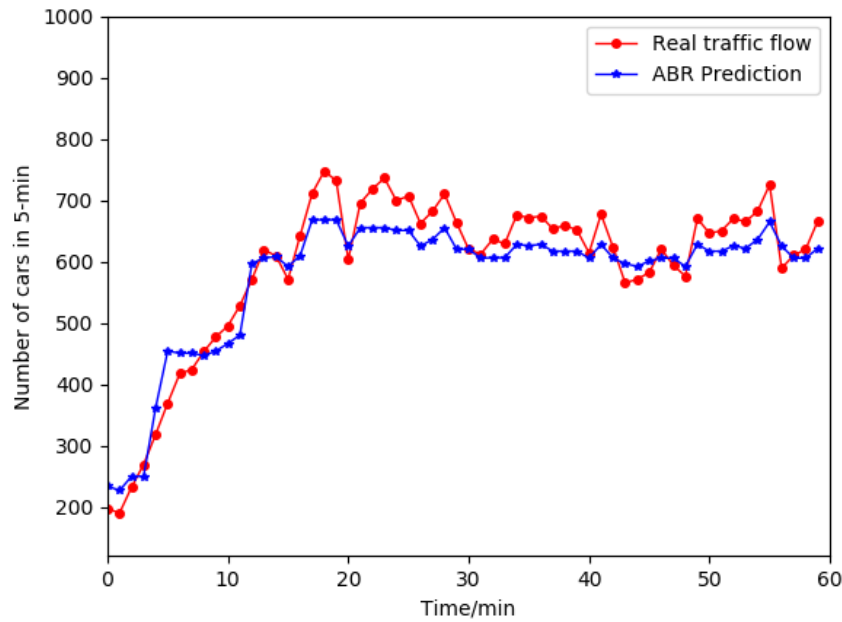**Fig. 2.** The real and predicted traffic flow during the rush hour from 5:00 AM to 10:00 AM using DT on US101-N

Fig. 3. The real and predicted traffic flow during the rush hour from 5:00 AM to 10:00 AM using AdaBoost on US101-N
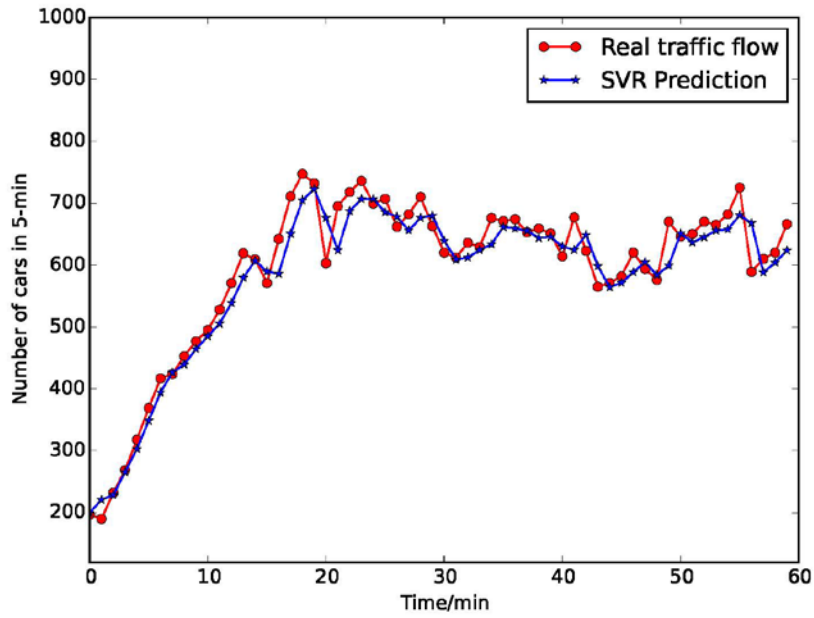
Fig. 4. The real and predicted traffic flow during rush hour from 5:00 AM to 10:00 AM using SVR on US101-N
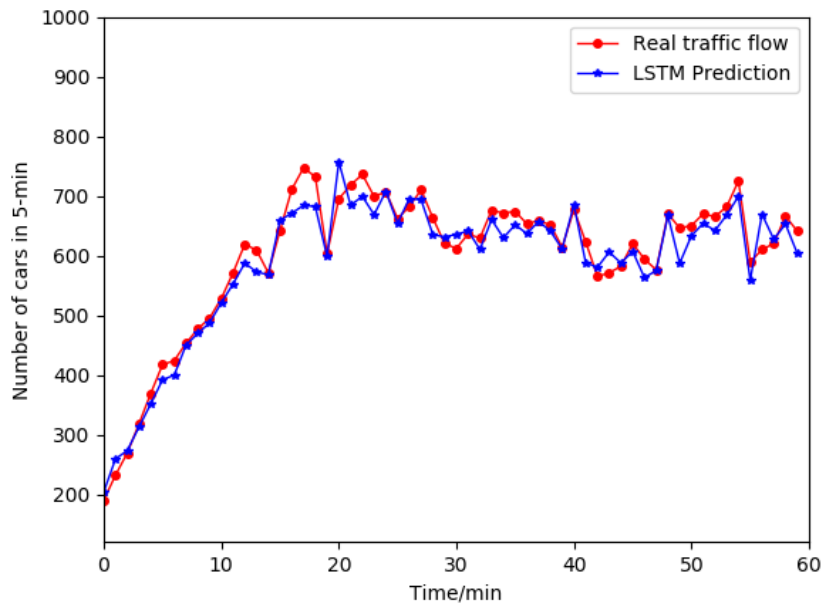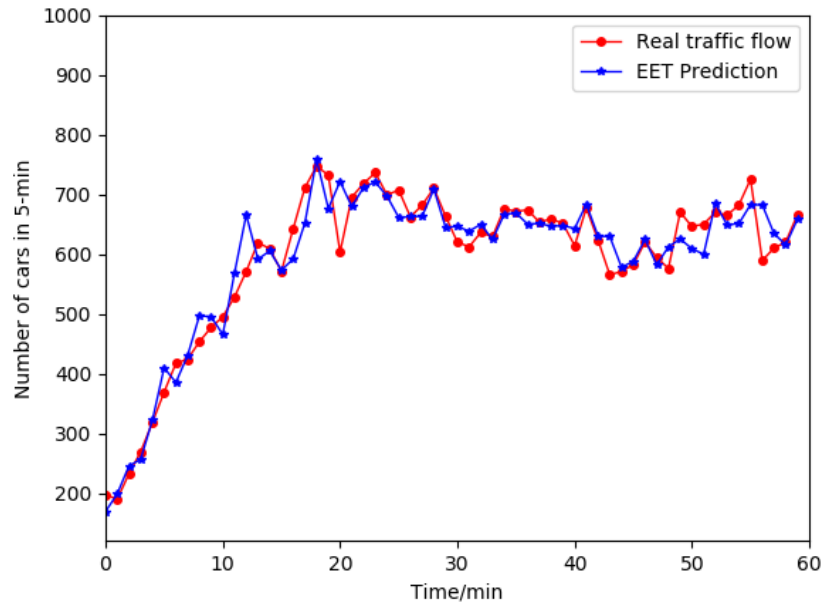

Fig. 5. The real and predicted traffic flow during the rush hour of 5:00 AM to 10:00 AM using LSTM on US101-N

Fig. 6. The real and predicted traffic flow during the rush hour of 5:00 AM to 10:00 AM using EET on US101-N

Table 2 to Table 4 show the MAPE, RMSE value and the MAE value calculated by different prediction methods at four detectors after running 5 times. From the previous introduction of the evaluation metrics, we know that MAPE is the representative of the relative error of the predicted values and that RMSE and MAE represent the absolute error. Since the test data set contained a midnight time period, the number of vehicles at this time was almost zero. Therefore, the overall relative MAPE becomes large, especially at points SR120-E. Therefore, it can be concluded from the longitudinal comparison of the tables that our experimental results are logical and representative.

By comparing MAPEs horizontally, we find that our EET predictions are better than the other methods except for individual rows. Specifically, at point US101-N, the effect of improvement is obvious. This shows that EET is more suitable for the prediction of non-stationary sequences. In Table 3 and Table 4 EET was also significantly better than the other algorithms, and the actual error decreased by approximately 3% to 5% compared with the other optimal algorithms. By analyzing the predicted results from different perspectives, it can be concluded that the proposed EET method is effective for short-term traffic flow prediction. The advantages of our EET exist and are obvious, both in terms of accuracy and stability.

Table 2. Performance Comparison of the MAPE for DT, ABR, SVR, and our EET

Freeway	DT	ABR	SVR	EET
I800-S	11.6136	7.7881	6.9461	**4.4532**
SR120-E	16.6372	33.1950	30.7212	**19.8808**
US101-N	7.2163	5.4948	4.2114	**3.9856**
1210-W	13.5136	9.2195	5.9235	**5.1549**
I5-N	11.0195	9.7389	6.2946	**5.2077**
Average	12.0000	13.0873	10.8194	**7.7220**

Table 3. Performance Comparison of the RMSE for DT, ABR, SVR, and our EET

Freeway	DT	ABR	SVR	EET
I800-S	101.9109	49.6517	43.6884	**35.1792**
SR120-E	12.9196	23.0978	14.9811	**8.7994**
US101-N	50.4465	37.8556	31.0053	**30.3252**
1210-W	52.1821	33.4888	19.9989	**19.1765**
I5-N	44.4462	36.7954	25.3114	**25.8024**
Average	52.3811	36.1779	26.9970	**23.8565**

Table 4. Performance Comparison of the MAE for DT, ABR, SVR, and our EET

Freeway	DT	ABR	SVR	EET
I800-S	62.8833	41.2063	35.7084	**24.0083**
SR120-E	7.9833	21.6247	12.5880	**5.2250**
US101-N	41.0167	31.3045	24.5090	**22.7383**
1210-W	36.9333	26.1918	15.4284	**14.3416**
I5-N	29.7000	27.0907	18.0513	**15.2616**
Average	35.7033	29.4836	21.2570	**16.3150**

5. Conclusion

In this paper, we propose an ensemble cascading prediction framework based on extra-trees using a boosting technique, called EET, to predict short-term traffic flow under non-stationary conditions. To the best of our knowledge, this is the first time that an ensemble framework is applied using extra-trees as building blocks for traffic flow prediction. We employ the EWMA algorithm to improve the prediction accuracy, which plays a significant role in our prediction procedure. The results show that the performance of our proposed EET is superior to conventional DT, ABR, SVR and LSTM approaches under non-stationary conditions both in accuracy and stability.

For future work, we will use more traffic flow data to learn more features of the traffic flow and acquire more accurate prediction results. Furthermore, we will focus on some problems in real data, such as missing data and data noise, which influences the prediction performance. This will help us to build a robust prediction system.

References

[1] Y. S. Jeong, Y. J. Byon, M. M. Castro-Neto, and S. M. Easa, "Supervised weighting-online learning algorithm for short-term traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 4, pp. 1700–1707, 2013. Article (CrossRef Link).

[2] Zhu Xiao, Xiangyu Shen, Fanzi Zeng, Vincent Havyarimana, Dong Wang, Weiwei Chen, and Keqin Li, "Spectrum Resource Sharing in Heterogeneous Vehicular Networks: A Non-Cooperative Game-Theoretic Approach with Correlated Equilibrium," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 9449-9458, 2018. Article (CrossRef Link).

[3] Vincent Havyarimana, Damien Hanyurwimfura, Philibert Nsengiyumva, Zhu Xiao, "A novel hybrid approach based-SRG model for vehicle position prediction in multi-GPS outage conditions,"*Information Fusion*, vol.41, pp.1-8, May, 2018. Article (CrossRef Link).

[4] D. G´omez, J. F. Mart´ınez, J. Sendra, and G. Rubio, "Development of a decision making algorithm for traffic jams reduction applied to intelligent transportation systems," *Journal of Sensors*,2016,(2016-9-25), vol. 2016, no. 3, 2016. Article (CrossRef Link).

[5] Z. Xiao, P. Li, V. Havyarimana, H. Georges, D. Wang, and K. Li, "GOI: A novel design for vehicle positioning and trajectory prediction under urban environments," *IEEE Sensors Journal*, vol. 18, no. 13, pp. 5586 – 5594, 2018. Article (CrossRef Link).

[6] L. L. Rui, Y. Zhang, H. Q. Huang and X. S. Qiu, "A New Traffic Congestion Detection and Quantification Method Based on Comprehensive Fuzzy Assessment in VANET," *KSII Transactions on Internet and Information Systems*, vol. 12, no. 1, pp. 41-60, 2018. Article (CrossRef Link).

[7] A. A. Lo´pez, A´ lvaro Duque de Quevedo, F. S. Yuste, J. M. Dekamp, V. A. Mequiades, V. M. Cort´es, D. G. Cobe˜na, D. M. Pulido, F. I. Urzaiz, and J. G. Menoyo, "Coherent signal processing for traffic flow measuring radar sensor," *IEEE Sensors Journal*, vol. PP, no. 99, pp. 1–1, 2017. Article (CrossRef Link).

[8] J. Xiao, Z. Xiao, D. Wang, J. Bai, V. Havyarimana and F. Zeng, "Short-term traffic volume prediction by ensemble learning in concept drifting environments," *Knowledge-Based Systems*, 2018. Article (CrossRef Link).

[9] Z. G. Shen, W. L. Wang, Q. Shen and Z. C. Li, "Hybrid CSA optimization with seasonal RVR in traffic flow forecasting," *KSII Transactions on Internet and Information Systems*, vol. 11, no. 10, pp. 4887-4907, 2017. Article (CrossRef Link).

[10] Vincent Havyarimana, Zhu Xiao, and Dong wang, "A Hybrid Approach-based Sparse Gaussian Kernel Model for Vehicle State Estimation during the Free and Complete GPS Outages," *ETRI Journal*, vol. 38, No. 3, pp. 578-587, June, 2016. Article (CrossRef Link).

[11] Z. Xiao, V. Havyarimana, T. Li, and D. Wang, "A nonlinear framework of delayed particle smoothing method for vehicle localization under nongaussian environment," *Sensors*, vol. 16, no. 5, p. 692, 2016. Article (CrossRef Link).

[12] R. Tahmasbi and S. M. Hashemi, "Modeling and forecasting the urban volume using stochastic differential equations," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 1, pp. 250–259, 2014. Article (CrossRef Link).

[13] Y. Hou, P. Edara, and C. Sun, "Traffic flow forecasting for urban work zones," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 4, pp. 1761–1770, 2015. Article (CrossRef Link).

[14] Y. Lv, Y. Duan, W. Kang, Z. Li, and F. Y. Wang, "Traffic flow prediction with big data: A deep learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 865–873, 2015. Article (CrossRef Link).

[15] W. Huang, G. Song, H. Hong, and K. Xie, "Deep architecture for traffic flow prediction: Deep belief networks with multitask learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 2191–2201, 2014. Article (CrossRef Link).

[16] M. Castro-Neto, Y. S. Jeong, M. K. Jeong, and L. D. Han, "Online-svr for short-term traffic flow prediction under typical and atypical traffic conditions," *Expert Systems with Applications An International Journal*, vol. 36, no. 3, pp. 6164–6173, 2009. Article (CrossRef Link).

[17] K. Y. Chan, T. S. Dillon, J. Singh, and E. Chang, "Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and levenberg –marquardt algorithm," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 2, pp. 644–654, 2012. Article (CrossRef Link).

[18] K. Kumar, M. Parida, and V. K. Katiyar, "Short term traffic flow prediction for a non urban highway using artificial neural network," *Procedia - Social and Behavioral Sciences*, vol. 104, pp. 755–764, 2013. Article (CrossRef Link).

[19] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine Learning*, vol. 63, no. 1, pp. 3–42, 2006. Article (CrossRef Link).

[20] H. Drucker, "Improving regressors using boosting techniques," in *Proc. of Fourteenth International Conference on Machine Learning*, 1997, pp. 107–115. Article (CrossRef Link).

[21] P. P. Varaiya, "The freeway performance measurement system (pems), pems 9.0: Final report," *Path Research Report*, 2009. Article (CrossRef Link).

[22] Guo Y, Zhang J, Gao L, "Exploiting long-term temporal dynamics for video captioning," *World Wide Web-internet & Web Information Systems*, pp.1-15, 2018. Article (CrossRef Link).

[23] Song J, Gao L, Nie F, et al, "Optimized Graph Learning Using Partial Tags and Multiple Features for Image and Video Annotation," *IEEE Transactions on Image Processing*, vol. 25, no. 11 pp. 4999-5011, 2016. Article (CrossRef Link).

[24] Gao L, Guo Z, Zhang H, et al, "Video Captioning With Attention-Based LSTM and Semantic Consistency," *IEEE Transactions on Multimedia*, vol. 19, no. 9, pp. 2045-2055, 2017. Article (CrossRef Link).

Fan Zhang received his B.S. degree in Computer Science and Technology from Chinese Peoples Liberation Army in Xi'an academy of Telecommunications in 2004,M.S. degree in Software Engineering from Northwestern Polytechnical University and from 2009 he is a Ph.D candidate in Xidian University.His research interests include quality of service in wireless networks and Network Topology Discovery.

Jing Bai received her B.S. degree in Electronic & Information Engineering, from Zhengzhou University in 2004 and Ph. D degree in Pattern Recognition & Intelligent System from Xidian University in 2009. She has been an associate professor at Xidian University. Her research interests includes high-resolution Remote Sensing image processing and interpretation.

Xiaoyu Li received his B.S. degree in communication engineering from Fuzhou University, China, in 2016. He is currently pursuing the M.S. degree at Xidian University.

Changxing Pei is a professor at Xidian University and his research interests include quantum communication, network measurement, network topology discovery , anti interference technology for wireless communication.

Vincent Havyarimana received his B.S. degree in Mathematics from University of Burundi, Bujumbura, in 2007. He got ME and PhD degrees in Computer Science and Technology from Hunan University, Changsha, China in 2011 and 2016 respectively. Currently, he is a Lecturer and Head of Research Service at Ecole Normale Superieure (ENS)-Burundi. He is also an academic member board of IT-Consulting and Trading Company (ICOTRACO)-Burundi. His research interests include Wireless Communication, vehicular ad-hoc networks and Mobile Computing.