ORIGINAL ARTICLE

CrossMark

Urban Arterial Travel Time Prediction Using Support Vector Regression

Anna Mary Philip[1] · Gitakrishnan Ramadurai[1] · Lelitha Vanajakshi[1]

Abstract

Travel time is one of the most appropriate traffic measures that can be provided to road users for easy understanding and decision making. This paper presents a simple method to predict travel time on urban arterial roads. Travel time prediction methods include historical methods, model based methods, and data driven approaches. Of these, data driven approaches capture the uncertainty and non-linearity of traffic time series better and hence could be used for travel time prediction under Indian traffic conditions effectively. Among data driven approaches, support vector regression (SVR) models predict travel times with reasonable accuracy, especially when the amount of data is less or the variability in the data is high. However, studies on the use of SVR for travel time prediction under Indian traffic conditions are limited. In this paper, SVR technique was used to predict the travel time using data collected from Bluetooth sensors placed at specific locations on an urban arterial corridor in Chennai, India. The optimum number of inputs, appropriate kernel function, cost parameter and width of tolerance for the SVR model were determined. The results obtained show that the SVR performed better than an Artificial Neural Network model and moving average approach.

Keywords Travel time prediction · Support vector regression · Bluetooth data · Sensitivity analysis · Heterogeneous traffic

Introduction

Travel time is an important traffic metric essential for the implementation of Advanced Traveller Information Systems (ATIS) and Advanced Traffic Management Systems (ATMS). Travel time is one of the most appropriate traffic measure that can be provided to road users for easy understanding and decision making. Hence, accurate and timely travel time prediction is very important for the success of ATIS and ATMS applications.

Traditional methods for travel time prediction include model based methods and data driven methods. Data driven approaches are approaches for which the estimation and prediction of travel time does not require the application of traffic flow theory and identities. Travel time is obtained from the data itself by statistical and machine learning techniques [6]. Several data driven and model based approaches have been proposed for the prediction of travel time on urban arterial links [4, 11, 13, 20]. Existing data driven approaches mainly use machine learning techniques, such as linear regression [7, 12, 19], k nearest neighbours (kNN) [9], Markov Chains [10, 17], artificial neural networks (ANN) [1, 18, 20] and support vector machines (SVM) [11, 15, 16] to predict travel time along urban arterial links.

Linear regression is a simple data driven technique that has been shown to give reasonable travel time predictions [12]. Linear Regression has also been reported to have a significant advantage in terms of computational time and memory resources [7]. Variations of linear regression, such as time-varying coefficient linear regression, have also been shown to produce good travel time forecasts [19]. But, linear regression is not suited for modeling relationships between traffic variables which are non-linear. Non-linear relationships can be modeled using other data driven techniques such as ANN, kNN and support vector regression (SVR). ANN is a machine learning technique that has been widely used for travel time prediction. They have been shown to produce fast and accurate results [1] and to perform well under different traffic conditions [5, 20]. ANN models have

✉ Anna Mary Philip
annamp470@gmail.com

Gitakrishnan Ramadurai
gitakrishnan@iitm.ac.in

Lelitha Vanajakshi
lelitha@iitm.ac.in

[1] Department of Civil Engineering, Indian Institute
of Technology Madras, Chennai, India

also been shown to be useful for long term prediction of travel time [8]. Ensemble models of ANN have been shown to provide accurate and unbiased results [14, 18]. In [4], it was shown that ANN models could be used to account for driver-to-driver variability and predict the range of travel time values in a link. The performance of ANN method was compared with that of a SVR model in [15]. The performances of both models were shown to be comparable, but when data availability was less and when the data was not representative, SVR performed better than ANN. SVR has also been shown to be applicable for the prediction of various traffic flow parameters, such as speed, flow, and headway, under Indian conditions [13]. Travel time prediction using SVR in India has also shown favorable results [11]. In [16], it was shown that with proper parameter tuning, the performance of SVR model can be improved and accurate results can be obtained. The above studies have demonstrated that machine learning techniques can be used effectively to obtain accurate predictions of travel time. These models were shown to be capable of capturing the uncertainty and non-linearity of a traffic time series from data and could be used for traffic prediction effectively.

Of these, SVR has been reported to predict travel times with reasonable accuracy, especially when the amount of data is less or the variability in the data is high [15]. Many studies [15, 16] have applied SVR models for travel time prediction with promising results for homogeneous traffic conditions. Studies involving the application of SVR models for travel time prediction in India where the traffic is heterogeneous and without lane discipline are limited. The aim of this paper is to study the applicability of SVR in Indian traffic conditions and to estimate the optimum input and parameter set for maximum accuracy of the SVR model.

Study Site and Data Collection

Study site for the present study was a 2.8 km long urban arterial road section between *Little Mount* and *Madhya Kailash* in Chennai. Figure 1 shows the selected study site. The road connecting *Little Mount* to *Sardar Patel Road* is a one-way four-lane road and *Sardar Patel Road* is a four-lane divided road. The direction from *Little Mount* to *Madhya Kailash* was selected in the present study.

The traffic volume at the study corridor during peak hours was about 7000 vehicles/hour. This volume was composed of 62% two-wheelers, 28% cars and LCV, 7% autorickshaws and 3% HMV. The average speed of the vehicles in the study stretch was found to vary between 15 and 30 km/h during peak hours [3].

Bluetooth sensors were located at *Little Mount* and *Madhya Kailash*. Data for 11 weekdays in the month of August, 2014 were used for the present study. The data reported by Bluetooth sensors included the Bluetooth ID of the device in the vehicle passing the sensor location and the time at which the device (vehicle) passed the sensor location. By comparing the time at which each vehicle crossed the sensors, the travel time between those sensor locations was estimated. The number of travel time values obtained in a 5-min interval varied from 0 to 30, with an average of 11 values. The travel time values, thus collected, were averaged over 5 min intervals and were used for travel time prediction.

A sample plot of the variation of travel time on a day is given in Fig. 2. The mean travel time was 4 min with a standard deviation of 1.5 min.

Fig. 1 Study site (*Source* Google Maps)
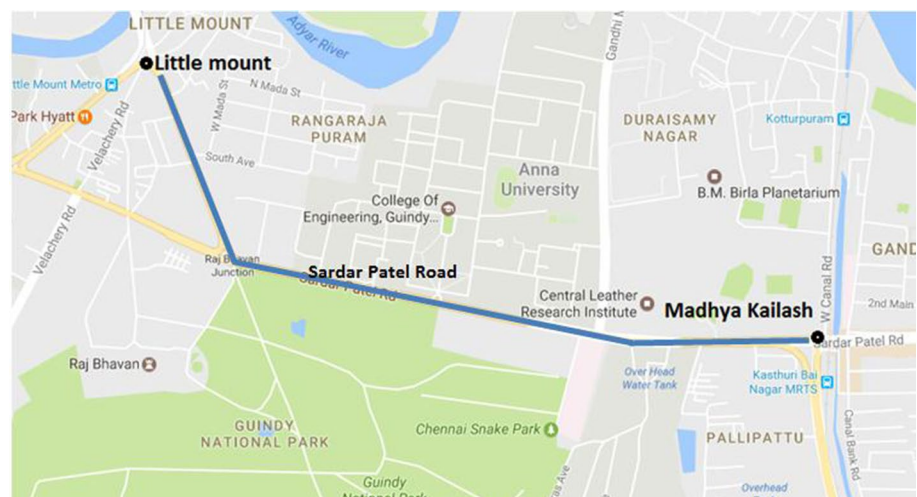
Fig. 2 Travel time variation on 18 August, 2014
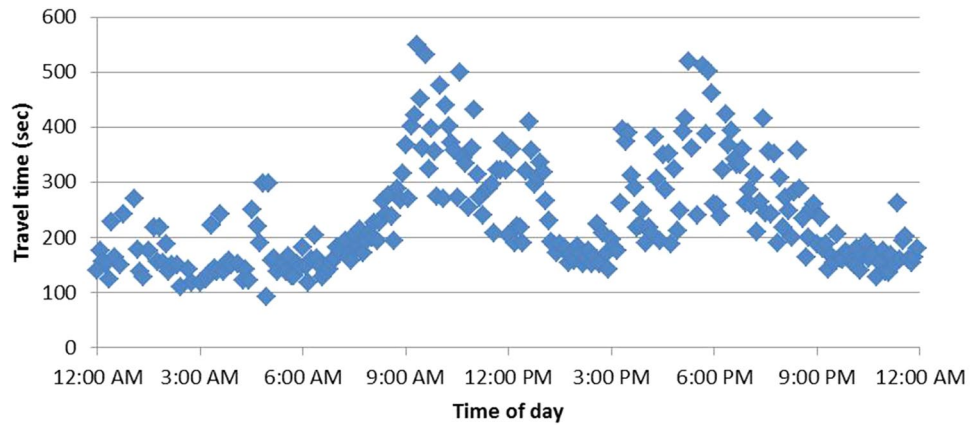
Methodology

In the present study, the 5 min averaged travel time values were taken as input for the prediction of next 5-min intervals' travel time. The number of inputs (previous time intervals' travel time values) and the various parameters of the model were varied to find the optimal values. The mean absolute percentage error (MAPE) and root mean square error (RMSE) were used as measures of prediction accuracy and are given by Eqs. 1 and 2.

$$MAPE = \frac{100}{n} \sum_{1}^{n} \frac{|A_t - P_t|}{A_t}, \tag{1}$$

$$RMSE = \sqrt{\frac{\sum_{1}^{n}(A_t - P_t)^2}{n}}, \tag{2}$$

where, n is the number of observations, A_t is the actual value and P_t is the predicted value

The present study used Support Vector Machine (SVM) technique for the problem of short-term prediction of travel time. To compare the performance, ANN and moving average method were used.

SVMs are supervised learning algorithms that analyse data for classification and regression. When data points are not linearly separable, SVM maps them onto a higher dimension space so that there is a hyperplane boundary between them. Data points belonging to mutually exclusive categories are classified by constructing a hyperplane or a set of hyperplanes in this higher dimension space. New data points are then mapped to this space and predicted to belong to a category, based on which side of the hyperplane they fall [2]. SVR is a version of SVM, which uses this method for regression, instead of classification.

The mapping of data points to high dimensional space is carried out with the help of kernel functions. A kernel

Table 1 Common kernel functions

Kernel	Function
Linear	xy
Polynomial	$[(xx_i) + 1]^d$
RBF	$exp\{-\gamma|x - x_i|^2\}$

function maps the data points into a higher dimensional space without the need to compute the coordinates of the data points in the higher dimensional space. The commonly used kernels in SVR are linear kernel, polynomial kernel and Radial Basis Function (RBF) as shown in Table 1.

If the data points are separable in the higher dimensional space, the SVR model becomes simple, since the hyperplane can be fixed at maximum margin from both the categories of data points. However, if the data points are not separable, the problem becomes complex. In this case, fitting the hyperplane separating all the points may lead to overfitting. To avoid this, certain data points are allowed to be on the wrong side of the hyperplane. A balance is achieved between the number of misclassified points and overfitting by using parameters C, known as cost parameter, and ϵ, known as the margin of tolerance.

The cost parameter denotes the weight or importance of each misclassified point. At higher values of C, more significance is attributed to misclassified points, leading to more number of correctly classified points. At lower values of C, more importance is given to avoid overfitting, which allows more number of misclassified points. The aim is to obtain an optimum level between misclassification and overfitting, so that the model can be used for prediction. The margin of tolerance ϵ denotes how far the margin is from the separating hyperplane. Larger the value of ϵ, farther the margin and smaller the penalty for wrongly classified points. Varying the values of C, ϵ and the kernel function can change the results. Tuning of these parameters was carried out to get optimum performance.

Implementation and Results

The 5 min averaged travel times were taken as input to build the SVR model. *fitrsvm* function in the Statistics and Machine Learning Toolbox of MATLAB was used for model estimation. Sensitivity analysis was carried out on the number of inputs to be used, kernel functions, and different SVM parameter combinations. The section below presents the findings of the sensitivity analysis.

Sensitivity Analysis for Number of Inputs and Kernel Function

Sensitivity analysis was carried out to study the effect of input data size and kernel functions used. Three kernel functions, RBF, linear and polynomial, were used for modelling. Input size (number of previous averaged travel time values) was varied from 2 to 12. Figure 3 shows the MAPE values obtained. The variation of RMSE values also showed a similar trend. Performances of the RBF and polynomial kernel were better than the linear kernel. Beyond 8 input values, there was a reduction in accuracy of the polynomial and RBF kernel models. Thus, the optimum input was taken as eight 5 min interval travel time values, indicating that eight previous 5-min interval data can best represent the next 5-min data. This corresponds to the previous 40 min travel time values. Polynomial kernel gave minimal error compared to RBF and Linear. Hence, polynomial kernel with eight previous interval data as input was taken for further analyses.
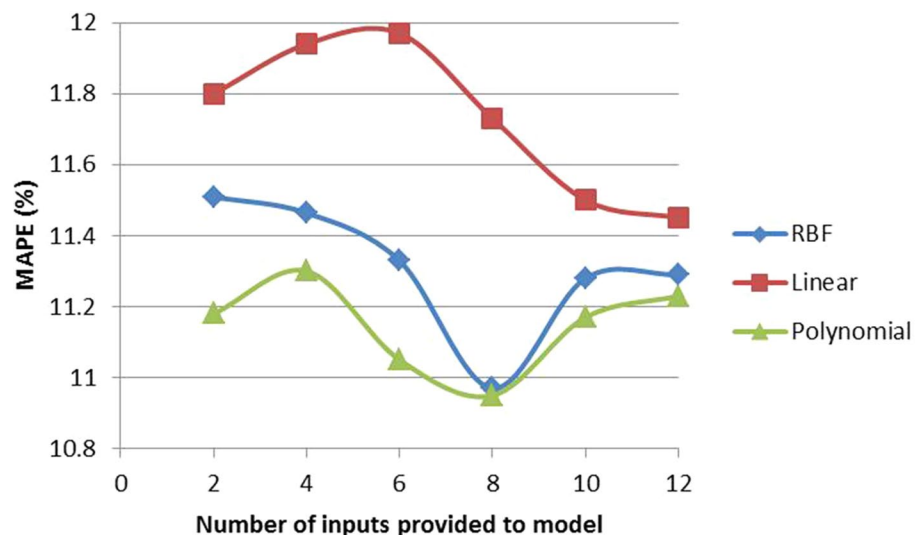
Parameter Tuning

Using the selected polynomial kernel SVR model with eight inputs, sensitivity analysis for the SVM parameters, C and ϵ was carried out. Cost parameter C values were varied from 1 to 100. Similarly, for ϵ, values ranging from 0.1 to 10 were used. The variation of MAPE values obtained by changing these parameters is shown in Fig. 4.

It can be seen that the MAPE value reduced significantly and subsequently increased as the value of C was increased, for all values of ϵ. Optimum C value was approximately 25. Increasing the value of ϵ resulted in increase of MAPE value. Hence, ϵ value of 0.1 was taken as the optimum.

Using the optimal model identified above, travel times were predicted. The prediction process has two stages, namely training and testing. In this study, 10 days' data were used for training and 1 days' data was kept for testing the results. Only weekday data were used for analysis.

The results obtained from the SVR model were compared with travel time values predicted using an ANN model and a baseline moving average method. An ANN model with three input layers and 10 hidden nodes in the second layer was used to predict the travel time. The number of hidden nodes was selected by a trial and error process. ANN model was modeled in MATLAB using backpropagation algorithm. Same input of eight previous interval travel time values was used here. In moving average method, the travel time for a time interval was estimated as the average of the travel times for the previous eight time intervals. This was used as a baseline method for comparison of performances of the SVR and ANN models.

A comparison of the travel time predictions from all the three approaches with the actual travel time is given in Fig. 5. Figure 6 shows the corresponding errors in prediction by the SVR, ANN and moving average methods.

Fig. 3 Variation of MAPE with number of inputs

Fig. 4 Variation of MAPE with C value
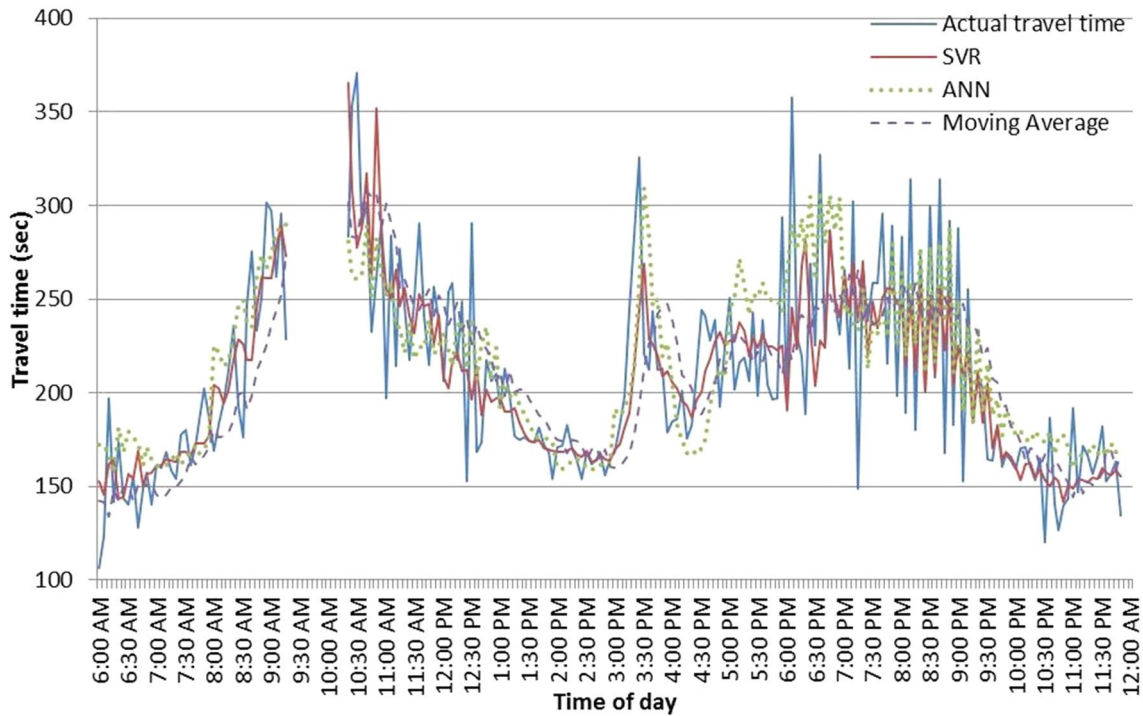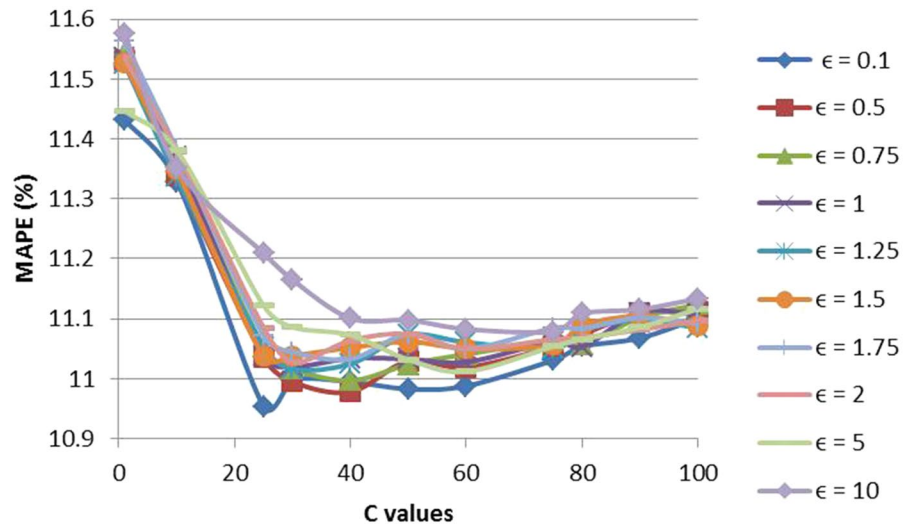
Fig. 5 Travel time prediction by SVR, ANN and moving average

The error values corresponding to the above plots have been summarized in Table 2. The Coefficient of variation (COV) of the error values are also shown. It can be seen that the errors are minimum for SVR followed by ANN and the highest errors are for the moving average method. Thus, it can be concluded that the SVR model was able to capture the travel time patterns and variations more efficiently than the other two methods.

Effect of Aggregation Interval

Variation in the performance of the developed model with varying aggregation intervals was also studied. In addition to the 5 min averaged travel time values described in "Study Site and Data Collection", 2, 10, 15, 20 and 25 min averaged travel time values were also used for travel time prediction. Time intervals less than 5 min were not used because it
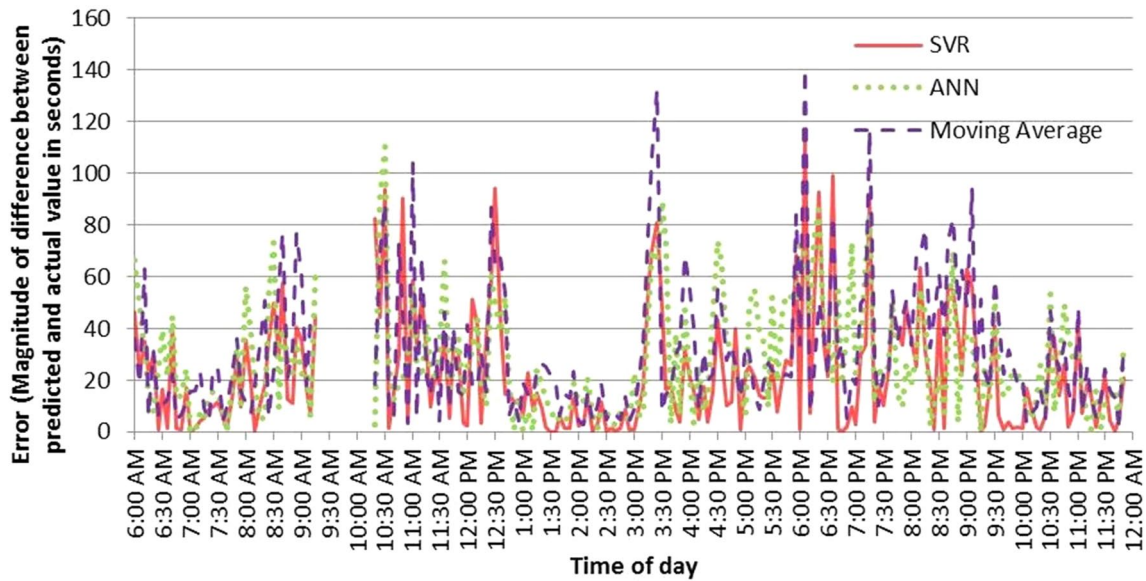
Fig. 6 Comparison of errors from SVR, ANN and moving average methods

Table 2 Error metrics of the models

Model	MAPE (%)	RMSE (s)	Maximum error (s)	COV
SVR	10.95	32	112	0.92
ANN	13.79	37	123	1.17
Moving average	16.36	43	137	1.12

resulted in many time intervals with insufficient number of data points. The optimal model parameters were used and the results obtained are shown in Fig. 7.

It could be seen that as the averaging interval increases, the prediction error decreases. This is due to the variations in travel time getting smoothened over larger interval, thus reducing the variations in the data. The minimum values of error were obtained at 15 min averaging interval. However, a more extensive study involving multiple study sections may be required to draw conclusions.

Conclusions

Travel time prediction was carried out for a 2.8 km urban arterial corridor in Chennai, using SVR, ANN and moving average method. Travel time for a 5 min interval was predicted from previous 5-min intervals' travel times. Travel times from prior 40 min were used to predict the travel time of the current interval. SVR, ANN and moving average models were built with 8 input values and the results obtained were compared. The best MAPE and RMSE values of 10.95% and 32 s was obtained from the polynomial kernel SVR model with a C value of 25 and ϵ value of 0.1. Thus, with suitable kernel functions and model parameters, the SVR model could be used to predict the travel time at the next instant quite accurately using previous travel time values. Overall, SVR is found to be a promising option for travel time prediction under Indian traffic conditions.
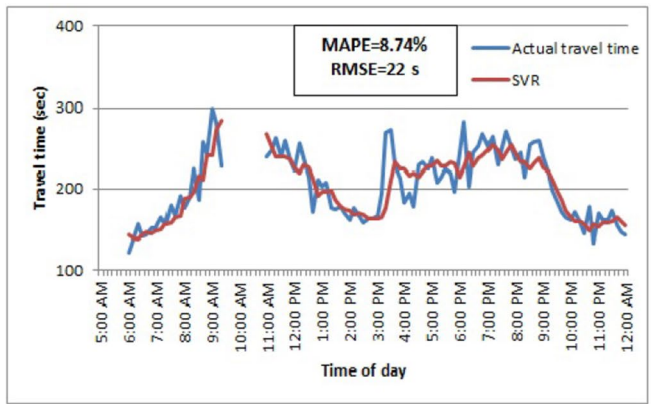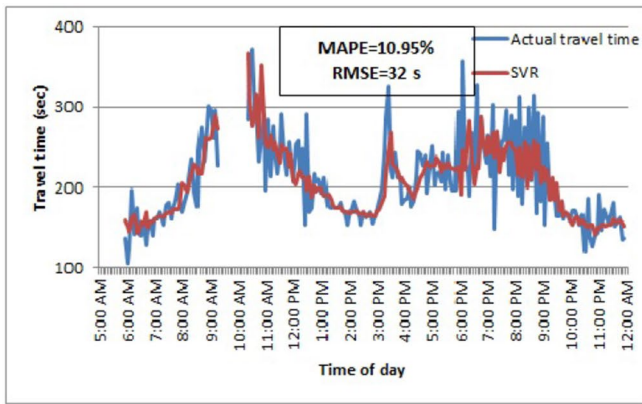
(a) SVR predictions for 5 minute interval data

(b) SVR predictions for 10 minute interval data
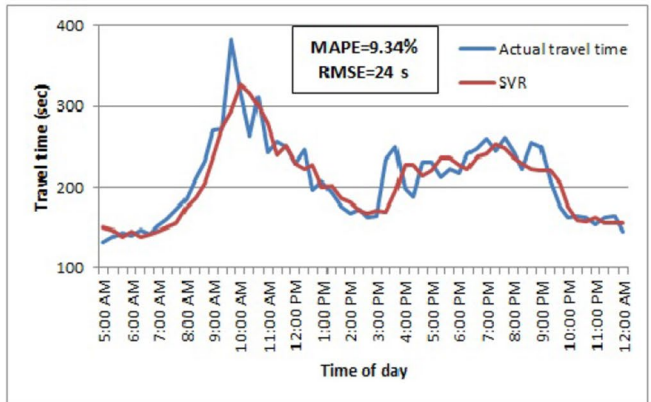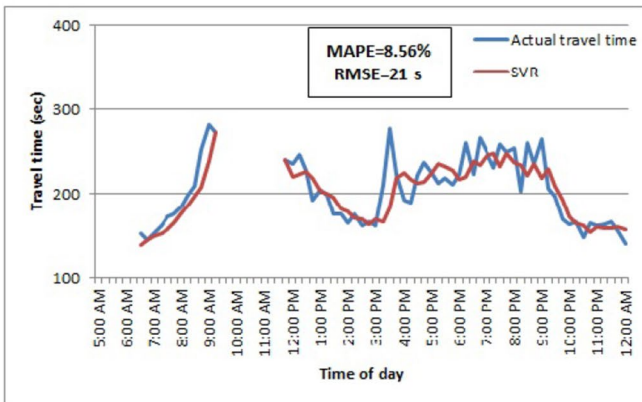
(c) SVR predictions for 15 minute interval data

(d) SVR predictions for 20 minute interval data
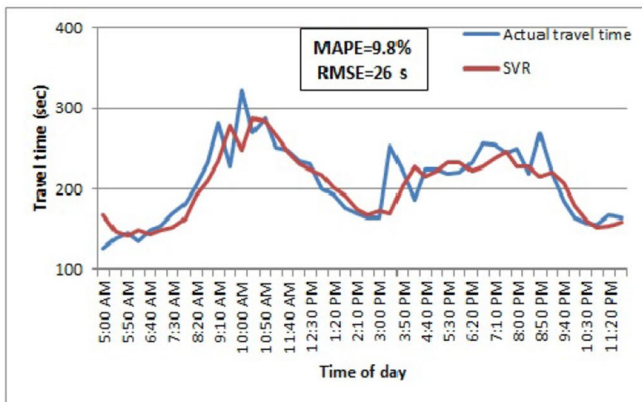
(e) SVR predictions for 25 minute interval data

Fig. 7 Travel time prediction with SVR for different time interval data

References

1. Dharia A, Adeli H (2003) Neural network model for rapid forecasting of freeway link travel time. Eng Appl Artif Intell 16(7):607–613
2. Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning. Springer, New York

3. IIT M (2014) Traffic data collection and analysis for the research study on Advanced Traveller Information System (ATIS) for Indian cities. Tech. rep. https://coeut.iitm.ac.in/Final%20Report%20R01.pdf

4. Li R, Rose G (2011) Incorporating uncertainty into short-term travel time predictions. Transp Res Part C Emerg Technol 19(6):1006–1018

5. Liang Z, Ling-Xiang Z (2006) Link travel time estimation model fusing data from mobile and stationary detector based on BP neural network. In: Communications, Circuits and Systems Proceedings, International Conference on, vol 3. IEEE, pp 2146–2149

6. Mori U, Mendiburu A, Alvarez M, Lozano JA (2015) A review of travel time estimation and forecasting for advanced traveller information systems. Transp A Transp Sci 11(2):119–157

7. Nikovski D, Nishiuma N, Goto Y, Kumazawa H (2005) Univariate short-term prediction of road travel times. In: Intelligent Transportation Systems. Proceedings. IEEE, pp 1074–1079, 13 Sep 2005

8. Park D, Rilett LR (1999) Forecasting freeway link travel times with a multilayer feedforward neural network. Comput Aided Civil Infrastruct Eng 14(5):357–367

9. Qiao W, Haghani A, Shao CF, Liu J (2016) Freeway path travel time prediction based on heterogeneous traffic data through non-parametric model. J Intell Transp Syst 20(5):438–448

10. Ramezani M, Geroliminis N (2012) On the estimation of arterial route travel time distribution with markov chains. Transp Res Part B Methodol 46(10):1576–1590

11. Reddy KK, Kumar BA, Vanajakshi L (2016) Bus travel time prediction under high variability conditions. Curr Sci 111(4):700

12. Rice J, Van Zwet E (2004) A simple and effective method for predicting travel times on freeways. IEEE Trans Intell Transp Syst 5(3):200–207

13. Theja P, Vanajakshi L (2010) Short term prediction of traffic parameters using support vector machines technique. In: Emerging Trends in Engineering and Technology (ICETET), 3rd International Conference on 2010 November 19. IEEE, pp 70–75

14. Van Lint J (2006) Reliable real-time framework for short-term freeway travel time prediction. J Transp Eng 132(12):921–932

15. Vanajakshi L, Rilett LR (2007) Support vector machine technique for the short term prediction of travel time. In: Intelligent Vehicles Symposium. IEEE, pp 600–605, 13 June 2007

16. Wu CH, Ho JM, Lee DT (2004) Travel-time prediction with support vector regression. IEEE Trans Intell Transp Syst 5(4):276–281

17. Yeon J, Elefteriadou L, Lawphongpanich S (2008) Travel time estimation on a freeway using discrete time markov chains. Transp Res Part B Methodol 42(4):325–338

18. Zeng X, Zhang Y (2013) Development of recurrent neural network considering temporal-spatial input dynamics for freeway travel time modeling. Comput Aided Civil Infrastruct Eng 28(5):359–371

19. Zhang X, Rice JA (2003) Short-term travel time prediction. Transp Res Part C Emerg Technol 11(3):187–210

20. Zheng F, Van Zuylen H (2013) Urban link travel time estimation based on sparse probe vehicle data. Transp Res Part C Emerg Technol 31:145–157