Short Term Prediction of Traffic Parameters Using Support Vector Machines Technique

Theja P. V. V. K
Under Graduate Student,
Department of Civil Engineering
Indian Institute of Technology Madras
Chennai, India 600036

Lelitha Vanajakshi
Assistant Professor
Department of Civil Engineering
IIT Madras, Chennai, India 600 036
Email: lelitha@iitm.ac.in

Abstract-**Accurate and precise prediction of traffic variables such as speed, volume, density, travel time, headways etc. is important in traffic planning, design, operations, etc. Short term prediction of these variables plays a very important role in Intelligent Transportation Systems (ITS) applications. Under Indian scenario, this short term prediction of traffic variables has gained greater attention with the recent interest in ITS applications such as Advanced Traveller Information systems (ATIS) and Advanced Traffic Management systems (ATMS).**
In the context of prediction methodologies, different techniques such as time series analysis, statistical methods, filtering techniques and machine learning techniques have been suggested in different studies in addition to the historic and real time approaches. However, for traffic conditions such as the one existing in India, with its heterogeneous and less lane disciplined traffic, many of these techniques may not bring the accuracy that was reported in literature under homogeneous traffic. There are only very limited studies on the application of these techniques for traffic conditions such as the one existing in India. The present study proposes the application of a recently developed pattern classification and regression technique called support vector machines (SVM) for the short-term prediction of traffic variables under mixed and less lane disciplined traffic conditions. An ANN model is also developed and a comparison of the performance of both these techniques is carried out.

Keywords- Intelligent Transportation Systems; Support Vector Machines; Short Term Traffic Prediction; Heterogeneous Traffic.

I. INTRODUCTION

In India, due to the increasing traffic demand, the need for implementing new systems for efficient utilization of infrastructure is increasing. This need resulted in the emphasis being shifted to implementation of Intelligent transportation systems (ITS). ITS applications are those which improve the efficiency of surface transportation systems and solve transportation problems by using modern information and communication technologies. With the recent interest in ITS systems like Advanced Traveller Information systems (ATIS) and Advanced Traffic Management systems (ATMS) in India, the importance of appropriate research pertaining to the field has increased greatly. One of the most important requirements of these systems is the ability to predict the nature of the traffic stream accurately. There are three fundamental macroscopic traffic variables for a traffic stream namely density, speed and volume. Ability to predict these variables helps to understand the nature of the changes that can occur in the transportation networks and thereby helps in mitigating congestion, accidents, etc.

As most of the ITS applications are real-time or adaptive in nature, it is even more important to have the knowledge of the stream characteristics beforehand. This helps in taking preemptive actions and planning the traffic management systems accordingly. Hence, there is a need to have a reliable technique for predicting the traffic variables. There are different techniques explored and reported for short term prediction of traffic parameters such as historic approach, real-time method, time series analysis, artificial neural networks (ANN) and more recently support vector machines (SVM). However, most of the studies reported the performance of these technique for traffic prediction using homogeneous traffic data. However, traffic conditions in many countries, such as India, are not homogeneous and lane disciplined. Traffic under conditions such as these is highly random in nature and it is difficult to model them mathematically. Due to this randomness, many of the prediction techniques do not perform satisfactorily. However, forecasting precision is very important in ITS applications since the reliability of such systems heavily depend on the accuracy of the information provided to the drivers. Hence, there is a need to explore better prediction techniques which may perform better even under traffic conditions with more variability. SVM is one such technique which came into light in the recent period.

Support Vector Machines (SVM) is a relatively new machine learning technique which is used for classification and regression purposes. The fact that SVM has better generalization ability from limited samples than the traditional techniques triggered exploring this technique for short term prediction of traffic parameters. Studies have reported the use of SVM for forecasting. However, their performance is not tested for Indian traffic conditions. Hence, there is a need to explore the use of SVM for predicting traffic parameters under Indian scenario.

The objective of this paper is to investigate the usefulness of Support Vector Machines (SVM) for short-term prediction of traffic parameters under Indian traffic conditions. The performance of SVM technique is then compared to that of ANN for a prediction horizon of one minute and five minutes.

The second author acknowledges the support of the Department of Science and Technology, Government of India under the grant SR/FTP/ETA-55/2007 for carrying out this work.

978-0-7695-4246-1/10 $26.00 © 2010 IEEE
DOI 10.1109/ICETET.2010.37

The following sections discuss the prediction models, implementation and results.

II. LITERATURE REVIEW

There are different methods which are reported for traffic parameter prediction purposes such as historic method, real-time method, time series analysis, statistical methods, machine learning methods etc. It is essential to understand the working principles behind each of these methods to know the limitations and advantages of using them. Hence, it is essential to study these methodologies and their characteristics.

Historic approach is based on the assumption that the historic data pattern will be representative of the future traffic parameters. The main limitation of this approach is that the temporal variation in traffic is assumed to be constant which is not practical. In the real time approach, the most recently obtained data profile is assumed to hold into the future. In this approach also the model can perform reasonably well under traffic conditions with less variation. However, in the case of atypical conditions, the performance of the approach is not satisfactory. The time series analysis is based on the assumption that successive values in the data file represent consecutive measurements. This generally reflects the fact that observations close together in time will be more closely related than observations further apart. Hence, the model becomes unreliable as we try to forecast further into the future. Models based on artificial neural networks are reported to have better prediction accuracy over the earlier models in most of the cases. However, they are found to be unreliable when the training data is not representative of the actual pattern and when the size of the training data is less.

Thus, all the above methods have their own limitations. It is found that most of these methodologies do not perform well under atypical conditions or when the amount of data available is less or non-representative. However, among the above mentioned methods, ANN is a viable option for highly random conditions where the data does not follow any kind of pattern. The main objective of this work being the application of SVM for short term prediction of traffic parameters under Indian traffic conditions, a detailed literature review was carried out in this area of short term prediction using SVM and is summarized below.

In a comparative analysis reported by Vanajakshi and Rilett [1, 2] for the short term travel time and traffic speed prediction, the usefulness of SVM is investigated under homogeneous traffic conditions by comparing its performance against that of historic approach, real-time approach and ANN. This analysis considered forecasts ranging from 2 mins to 1 hour into the future. These studies showed that for short term prediction, the performance of SVM and ANN is better than the others. When these two are compared with each other, it was observed that SVM is a better option if the amount of data is less or if the training data set is not a good representative sample of the testing data.

SVM is also integrated with other models and some of these integrations are reported to be efficient for their application in transportation. One such application is the use of SVM rough neural network for real time recognition of chaos in traffic flow. The study carried out by Pang and He [3] showed that it is highly feasible to recognize chaos in traffic flows by using real-time intelligent recognition system by SVM rough neural network. It was also proved that the correct recognition rate is higher than all other erstwhile known methods.

Another reported study was on traffic flow forecasting using SVM with fixed kernel parameters [4]. Bayesian inference was used to fix the SVM's kernel parameters. This improved the regression precision compared to existing methods like average method, ARMA, linear regression, nonparametric regression, neural networks etc. In a study by Cao *et al.* [5], ARIMA (auto-regressive integrated moving average) model and SVM model were combined with Judgemental Adjustment (JA) technique for traffic flow forecasting. It was reported that the results obtained by JA combined models are superior to those obtained using other models.

Usually in most of the transportation applications, a fixed SVM structure is used for the model. In a study carried out by Li [6], a time-dependent structure for SVM was proposed. In this, the SVM structure was determined by the input within the last one hour. Hence, the SVM structure changed temporally based on the training set used. This model was used to predict short term travel time for every 5 minute interval. As the SVM adapts itself to the different traffic environments, the prediction function was updated continuously.

In another work reported by Neto *et al.* [7], Online SVR (OL-SVR) was used to predict traffic flow under typical and atypical conditions. These atypical conditions included vehicular crashes, inclement weather, work zone, holidays etc. In training the OL-SVR, the training set was augmented one sample at a time and was updated after every step. In other words the newly obtained flow value is incorporated into the training set and it is updated accordingly. The performance of OL-SVR is compared with Gaussian maximum likelihood (GML), Holt exponential smoothing and ANN models. It was reported that GML performed slightly better under typical traffic conditions while OL-SVR performed best under atypical conditions.

The general conclusion that can be drawn from these studies is that SVM holds great promise and is slightly better in fuzzy conditions where other methods fail to capture the pattern of data. However, the above studies were not tested under heterogeneous conditions that exist in India, and it is the aim of the present study. A brief description of SVM and its working is detailed in the next section.

III. SUPPORT VECTOR MACHINES

Support Vector Machines (SVM) is a type of machine learning technique which is used for data classification and regression purposes. Many studies in different fields are reported on the use of SVM for classification and regression. The present study concentrates on regression and is discussed below [8].

Unlike in classification problem, where data is classifies into various classes, SVR tries to estimate a real valued output for any given input. The regression model using an e-insensitive cost function is explained below.

To start with, SVR constructs a set of hyper planes for the regression purpose. This hyper-plane is described by the equation

$$y = w.x + b \tag{1}$$

The process of constructing the hyper planes involves selecting the appropriate values for variables w and b. This learning is done using a set of known inputs and outputs. This process is known as training.

The training data of D dimensional input is of the form

$$\{x_i, y_i\} \text{ where } i = 1 \ldots n;\ y_i \in \Re;\ x_i \in \Re^D \tag{2}$$

This means that for an input vector x_i, the expected training target is y_i. Once a hyper plane is fixed, and if the value of the training target y_i corresponding to training input x_i is greater than a certain distance (ε) from the hyperplane, then penalty is allocated to this data point. Penalty is not allocated if the predicted value y_i is less than a distance 'ε' away from the actual value t_i, i.e. if $| t_i - y_i | < \varepsilon$. This region which is bound by ($y_i \pm \varepsilon$) is known as ε-insensitive region. The penalties are described by ε^+ or ε^- depending on whether they lie above or below the tube. The values of these penalties are decided by the cost function C.

The error function for SVM regression can then be written as:

$$C\sum_{i=1}^{n}(\varepsilon_i^+ + \varepsilon_i^-) + \frac{1}{2}\|w\|^2 \tag{3}$$

The value of w is modified so that the error is minimized subject to the constraints $\varepsilon_i^+ \geq 0$ and $\varepsilon_i^- \geq 0$. The hyperplanes are modified in each step until the optimality condition is reached. Optimized support vectors will have minimum error for given training data. This regression and the e-insensitive tube is illustrated in Fig 1.

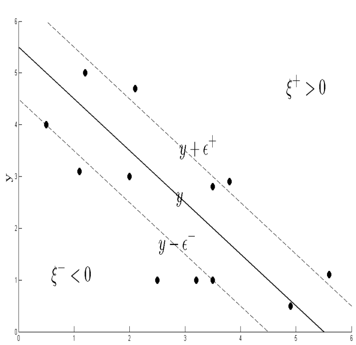


Figure 1. Regression with e-insensitive tube

Other types of loss functions include quadratic, laplace and Huber loss function. In the present study, the objective is short term prediction of traffic variables. The values of these traffic parameters are usually very widely distributed which results in the support vectors being sparse in nature. Hence, an SVM with e-insensitive loss function is used in the present study.

IV. Data

The three important traffic stream parameters that can be used to characterize a traffic stream are the stream speed, volume and density. The present study concentrates on the short term prediction of these important traffic parameters. Speed is the measure of the rate at which the position of vehicles changes in a traffic stream. There are two ways of measuring traffic stream speed and based on that it can be classified in to two types namely time mean speed (TMS) and space mean speed (SMS). Time mean speed is the average speed of a traffic stream passing a fixed point along a roadway measured over a fixed period of time. Space mean speed is the average speed of a traffic stream computed as the length of roadway segment divided by the total time required to travel the segment. In the present study, time mean speed of the vehicles measured in kmph is used. Volume or flow is the rate at which vehicles are passing a particular point on a roadway segment during a given interval of time. Count for every one minute interval is used for the volume prediction. Density is the average number of vehicles that occupy a fixed distance of a road, expressed in vehicles per mile or per kilometer. This being a spatial parameter, is difficult to measure and hence its microscopic measure namely space headway is used in the present study. Space headway is defined as the distance between the corresponding points of two consecutive vehicles. This directly indicates the density on a certain stretch of a roadway as:

$$K = 1000/h \tag{4}$$

K = density in veh/km
h = space headway in metres

Data collected using video footage from Rajiv Gandhi Salai in Chennai is used in the present study. This data comprised of one hour after noon peak video from five different days. Data is collected from a pedestrian over bridge facing the traffic stream direction. Extraction of the parameters was carried out manually in the laboratory. Volume data is obtained by counting number the vehicles crossing a fixed section of road on the video. Speed data is obtained by fixing two points at a known distance and measuring the time taken for a vehicle to move between the two points. Space headway was measured only between vehicles which were following each other in the same lane. If the vehicles are too far apart, then spacing between them is not measured. The spacing is taken as the distance between the front edge of the leading vehicle and the front edge of the following vehicle. The measured data on headway, speed and volume with time for the five days is shown in Fig 2, 3 and 4. It can be observed that the data for the five days is very random. In case of volume, day 2 data is observed to be different from that of the other days (Fig. 4).

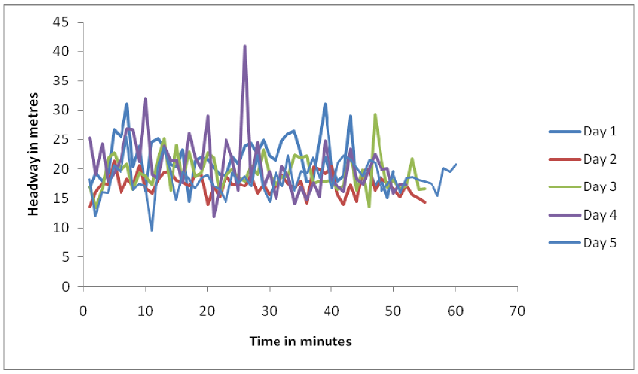


Figure 2. Headway distribution for the five days

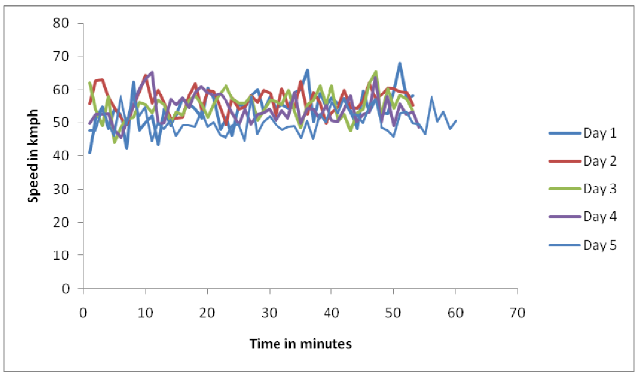


Figure 3. Speed distribution for the five days

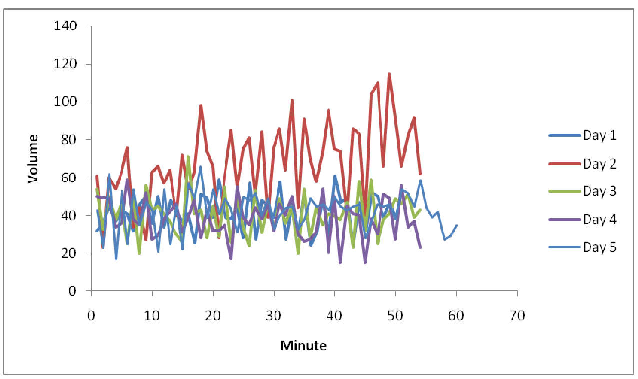


Figure 4. Volume distribution for five days

V. METHODOLOGY

An SVM/SVR toolbox developed by Gunn [9] is used for the implementation of SVR in the present study. A program was developed in Matlab for the purpose of prediction. An e-insensitive loss function was used for this model. There are different parameters involved in SVM, with each parameter affecting the performance in a unique way. A sensitivity analysis was carried out to study the effect of parameters on the performance of the model.

The data obtained is divided into every one minute discrete intervals. Four days data is used for training purposes and fifth day data is kept for validation. Training is the process of the machine understanding the patterns in the training data set. It usually requires a large amount of data. In the present study, about 200 data points, an hour of data from each of the four days, is kept for this purpose. This data is divided into input space and corresponding output space.

Training with an input dimension of n takes the parameter values for the first n time steps as the inputs and the successive value as the training output. In general if $x_1, x_2, x_3\ldots$ are the values of traffic variables, for a k dimensional input, $\{x_1, x_2\ldots, x_k\}$ is the input vector while x_{k+n} is the output for training the machine for a horizon of n time steps. In the present study, five dimensional inputs were used for training and testing. This was decided based on the assumption that the present data profile is affected by the data from the last five time steps. In the present study, prediction horizon of one minute and five minutes were attempted. In one minute forecasting, first five minute data is used as input while the sixth minute data is the output. In case of five minute forecasting, first five minute data is used as input while the tenth minute data acts as output. After training the machine with five dimensional inputs, the fifth day's data is used for the testing. The testing outputs are compared with the actual values available from the data and error analysis was carried out to test the prediction accuracy.

A. *Sensitivity Analysis*

Since SVM is a heuristic process, it is essential to find a proper implementation procedure and optimum parameter values which can be used for prediction. The SVR technique used in this study has two primary parameters namely cost function C and e-insensitivity region ε. Initially the default toolbox values were used. The default parameter values set were C=infinity and ε=0. A sensitivity analysis of these parameters was carried out to find the optimum values.

C was given a very large value (tending to infinity) initially so that the penalty is very high and ε is given a low value (ε =0.05) to have a very small ε-insensitive region. With the use of the above parameter values, it was found that the running time for training the machine is about 40 minutes. However, for ITS applications, this level of training time is not acceptable as most of the systems require real time output. Hence, a parametric study was carried out to improve the performance both in terms of time and accuracy.

It was found that using smaller values for C resulted in low computational time. However, using small values of C makes it less sensitive to outliers. Hence, an optimum value for C which gives an acceptable running time and accuracy has to be selected. A code was developed for this purpose to study the effect of C on the computational time and accuracy. Fig 5 shows a plot of change in training time with logC.

Similarly it was observed that values of ε have significant effect on the accuracy of prediction. Hence, the value of ε is varied and the corresponding prediction error is analysed. Fig 6 gives the variation of error with ε.

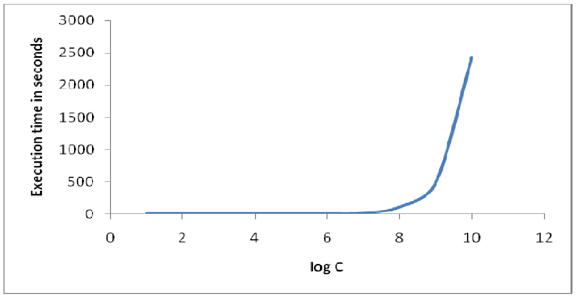


Figure 5. Variation of training time with logC

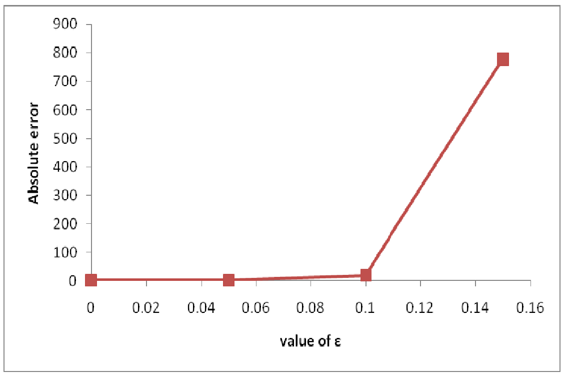


Figure 6. Absolute errors for different values of ε

From the results of the above study, the values of C equal to 10^6 for headway and speed and 10^4 for volume is selected for good performance. In case of ε, a value of 0.05 is found to be optimum for all the traffic parameters namely headway, speed and volume. The other parameters were also analyzed in the sensitivity analysis. However, it was observed that they have little or no effect on the computational time. A linear kernel is used in the prediction mechanism based on the literature survey.

B. ANN

The results of the SVM model need to be compared with that of other models to evaluate the feasibility of the model. Artificial neural networks (ANN) is one of the most popular methods reported for short term traffic prediction and hence is used in this study for comparison. A comparative study is carried out by analyzing the results from SVM and ANN models. A brief explanation of how ANN works is given below.

Artificial Neural Networks is a computational model that tries to simulate the structural and functional aspects of biological neural networks. A typical ANN involves a network of simple processing elements (neurons) which consists of nodes and connections. The nodes are connected to form a multi layered network. The connections in the network have certain weights which determine the relation between the input and output through each node. These weights of the connections are altered to produce a desired result. The weights of the connections are optimized using the inputs and outputs from training. When a new input is given to the trained network, it predicts the output based on its connections and weights. The input and output formats used for the prediction are similar to that of the SVM model.

VI. Results

The present study had about 200 data points for training from the first four days and about 50 data points from day 5 for testing. The accuracy of prediction was quantified using Mean absolute percentage error (MAPE), as defined below, as the performance measure.

$$MAPE = 1/n\sum_{i=1}^{n}\left|\frac{A_i - F_i}{A_i}\right| * 100 \qquad (5)$$

A_i – Actual value, F_i – Predicted value

Comparison between the actual and predicted outputs for one minute and five minute prediction horizon using SVM and ANN is carried out for all three traffic variables, namely headways, speed and volume. The percentage error for each minute data is then calculated using the outputs and the results are compared between SVM and ANN.

The comparison between one-minute and five-minute prediction of headway using SVM and ANN is discussed first. For one minute ahead prediction, the headways from 1-5 minutes are used as input and 6th minute's as output. This means that it is assumed that headways from last five minutes can be used for predicting the 6th minute's headway.

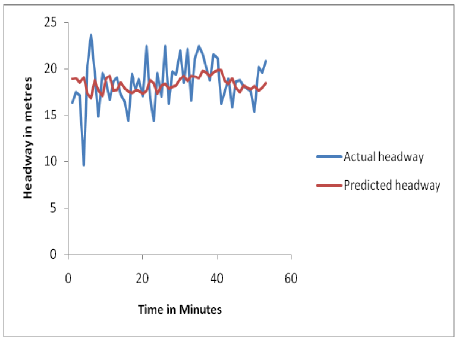
The comparison between the actual and predicted output using SVM and ANN are shown in Fig 7. The predicted output is found to have less variance compared to the actual value of headway in both cases. The mean absolute error of prediction in the case of SVM is 2.05 meters and the MAPE is 11.9%. The ANN prediction is observed to have slightly higher error with a mean absolute error of 2.4 meters and MAPE of 14.3%

For five minute ahead prediction, the headways from 1-5 minutes are used as input for 10th minute's output. The comparison between the actual and predicted headway in this case is shown in Fig 8. The mean absolute error of prediction using SVM for this case is 1.96 meters and the MAPE is 11.34%. In this case also ANN had slightly higher error with a mean absolute error of 2.43 meters and MAPE of 14.3%.

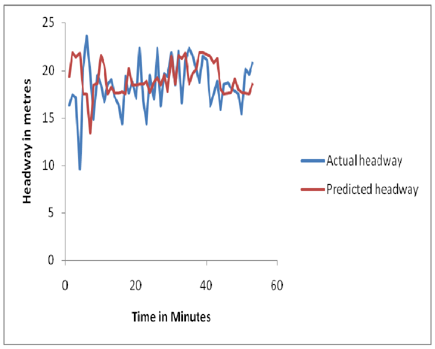
Similar predictions were carried out for speed and volume values. The MAPE for each of these predictions are shown in Table 1. It can be observed that SVM performed better than ANN in most of the cases except for five minute volume prediction with a MAPE of 23.5% against SVM's 24.4%. Overall performance of SVM is found to be better than ANN.

TABLE 1 MAPE FOR DIFFERENT PREDICTION INSTANCES

Prediction horizon	Traffic parameter	MAPE for SVM	MAPE for ANN	Preferred Model
One minute prediction	Headway	11.9	14.3	SVM
	Speed	7.7	10.56	SVM
	Volume	20.3	26	SVM
Five minute prediction	Headway	11.34	14.32	SVM
	Speed	9.2	11.7	SVM
	Volume	24.4	23.5	ANN

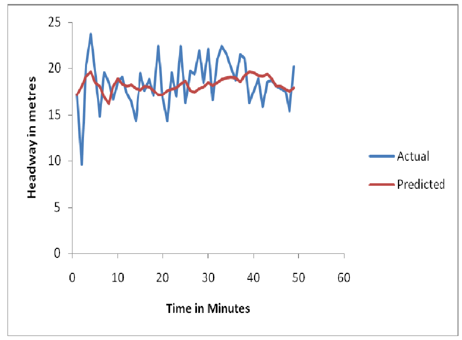


SVM

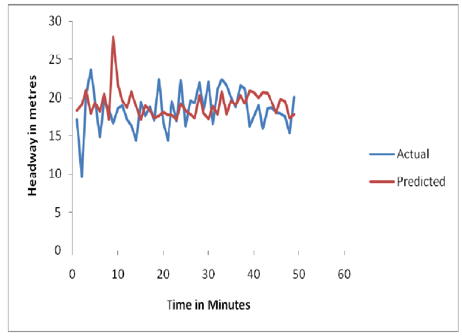


ANN

Figure 7. Comparison of actual and predicted headway – one minute ahead



SVM



ANN

Figure 8. Comparison of actual and predicted headway - five-minute ahead

VII. Conclusions

This paper investigated the usefulness of SVM for the short term prediction of traffic parameters namely speed, space headway and volume under heterogeneous traffic conditions. The SVM model used was a support vector regression with ε-insensitivity loss function with linear kernel. A sensitivity analysis was carried out to find the optimum parameters of SVR in terms of accuracy and running time. A comparison of the performance was carried out with an ANN model. The ANN model used was a multi-layer feed forward neural network with back propagation, based on literature.

The analysis was carried out for one minute and five minutes ahead prediction. Four days data was used as training set and one day data was left for cross validation to evaluate the prediction errors. The results obtained showed less prediction error with SVR in most cases. Overall, it was found that SVM is a viable alternative for short-term prediction of traffic parameter under Indian traffic conditions.

References

[1] Vanajakshi, L., Rilett, L. R. (2007). “Support Vector Machine Technique for the Short Term Prediction of Travel Time” *Intelligent Vehicles Symposium,* Istanbul, Turkey, 600-605

[2] Vanajakshi, L., Rilett, L. R. (2004). “A comparison of the performance of artificial neural networks and support vector machines for the prediction of traffic speed” *Intelligent Vehicles Symposium,* University of Parma, Parma, Italy, 194-199

[3] Pang, M.B., He, G. G. (2008). “Chaos Real-Time Recognition of Traffic Flow by Using SVM Rough Neural Network” *Wireless Communications, Networking and Mobile Computing, 4th International Conference*, Delian, 1-4

[4] Sun, Z., Pan, J., Duan, Q. (2008) 'Study on A New Traffic Flow Forecasting Method' *Natural Computation, Fourth International Conference*, Jinan. 349-353

[5] Cao, K., Liu, X. S., Cao, F., Zhao, M., Yu, S. W. (2006) 'A Dynamic Traffic Network Monitoring Algorithm' *Intelligent Transportation Systems Conference*, Toronto, 481-486

[6] Li, Q. (2009) 'Short-time Traffic Flow Volume Prediction Based on Support Vector Machine with Time-dependent Structure' *Instrumentation and Measurement Technology Conference,* Singapore, 1730-1733

[7] Neto, M. C., Jeong, Y. S.,jeong, M. K., Han, L. D. (2009) 'Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions' *Expert Systems with Applications: An International Journal*, 6164-6173

[8] Drucker, H., Burges, J. C., Kaufman, L., Smola, A., Vapnik, V. (1997). "Support Vector Regression Machines" *Advances in Neural Information Processing Systems*, MIT Press, 155-161

[9] Gunn, S.R. (1998). “Support Vector Machines for Classification and Regression” Available: http://users.ecs.soton.ac.uk/srg/publications/pdf/SVM.pdf