Short Time Traffic Speed Prediction Using Data from a Number of Different Sensor Locations

Ülkem Yıldırım, Zehra Çataltepe

Istanbul Technical University
Computer Engineering Department
Maslak, Istanbul, Türkiye

{ulkem.yildirim,cataltepe}@itu.edu.tr

Abstract

In this study we predict traffic speed on Istanbul roads using RTMS (Remote Traffic Microwave Sensor) speed measurements obtained from the Istanbul Municipality web site from 327 different sensor locations. We do speed predictions 5 minutes to an hour ahead and use SVM (Support Vector Machine) and kNN (k Nearest Neighbor) methods for speed prediction. First of all, for speed prediction at a certain sensor location, we compute the most important past speed measurements for better accuracy using feature selection methods. We also find out which other sensors could be used to predict the speed at a certain sensor location and show that especially for nearby/correlated sensors, it is possible to get better results using related sensor measurements in addition to the sensor being predicted. We also show that only using the correlated sensors, it is possible to get good accuracy. This result could be very useful when a sensor breaks down or needs to be calibrated. In all our experiments, we find out that SVM produces better results than kNN.

Keywords: short time travel prediction, traffic speed prediction, knn, svm

1. Introduction

As the cities get bigger and more crowded, traveling from one point to another becomes more difficult. In addition to the usual rush hour traffic during specific times of the day, unexpected situations also increase the waiting time in the traffic and affects the quality of life in the city negatively. Although municipalities work hard in order to solve the traffic problem, especially in the old city neighborhoods it is very difficult to construct new roads and there is not much that can be done. Intelligent Transportation Systems (ITS) are one of the possibilities to remedy the traffic situation by means of traffic sign control systems, electronic ticketing, highway management systems. ITS aim to provide a more comfortable and secure traffic situation for the population. Short time traffic speed prediction aims at predicting the traffic speed at a certain location within a short time period, such as a couple of minutes to an hour. Using these speed predictions, it is possible to compute an estimate of how long it would take to go from one point to another as well as the shortest path that results in the minimum travel time. The short time traffic prediction also helps drivers choose less crowded routes and results in a more homogenous traffic and less congestion. In predicting the speed at a short time ahead, the speed measurements from the past are used as inputs and the current speed measurements are taken as inputs. Linear models are the first models used for this purpose [11,13,16,4]. Since linear models are not able to model the complex relationship between the past and present speed measurements, for speed predictions in 15 minutes or further in the future, they perform worse than taking just the average of the past speed measurements [5,6]. Kalman filters, which use time dependent parameters also fail to give correct results for future speed predictions [9]. Considering the fact that the input-output relationship is nonlinear, a number of studies used SVMs (Support Vector Machines) and ANNs (Artificial Neural Networks). Both methods are able to model arbitrarily complex input-output mappings. ANNs originated from the idea of modeling the workings of human neurons and have been used in a number of studies [7,8,13,10]. Since the SRM (Structural Risk Minimization) principle used in SVMs results in less overfitting and SVMs do not suffer from the local minima problems encountered in neural network training, SVMs usually result in better traffic speed prediction [2,15]. In addition, SVMs are also better than ANNs in their ability to cope with missing data.

The sensors used for traffic speed prediction may sometimes break or need calibration, when this is the case, using other sensors to predict speed at the sensor location would be beneficial. On the other hand, storing and processing past sensor data is costly, if predictions could be computed using as small number of sensor measurements from the past as possible, that would also reduce the costs. In this study, we use SVM and kNN (k

Nearest Neighbor) methods for short time speed prediction. We first show that, using feature selection methods, it is possible to predict speed using small number of sensors. We also predict speed at a sensor location using other relevant sensors with or without the measurements at the sensor location. For certain sensors, the predictions are as good as predictions that use the measurements of the sensor being predicted.

The rest of the paper is organized as follows: In Section 2, we provide details about the traffic speed data acquisition. In Section 3, the pattern recognition models and the feature selection methods used are summarized. Section 4 includes the experimental details and Section 5 concludes the paper.

2. Data Set

The traffic speed measurements were gathered from the Istanbul Municipality Traffic Control Center's web page (www.tkm.ibb.gov.tr). The Traffic Control Center broadcasts speed measurements taken from 327 RTMS sensors located all over Istanbul. The measurements are updated every 1 or 2 minutes. The Traffic Control Center also maintains and calibrates the sensors continuously in order to make sure that the data displayed correctly reflects the current road conditions.

In order to predict the speed at a certain point 5 minutes ahead, the measurements taken 5, 10, 15, ..., 60 minutes, 1 day and 1 week ago (we use the notation: -5, -10, -15, ..., -60, -1440, -10080) are used. For speed prediction an hour later, the speed measurements taken 65, ..., 120 minutes, 1 day and 1 week ago are used.

3. Pattern Recognition Methods

3.1. k-Nearest Neighbor (kNN) and Support Vector Machine (SVM)

kNN [1] method is one of the simplest methods of pattern recognition. First, it measures the distance between the test data point and all traing points. Then it chooses k (in this study k=25) data points closest to the test data point. When it is used for regression, kNN output for the test data point is the mean of the outputs for all those k training data points. In this way, 25 different speed measurements are used to estimate a future speed which most like current speed distribution. kNN is slow in prediction phase, because it needs to calculate distance to every training data point. Decreasing feature count or the using a number of prototypes could increase the speed of kNN.

SVM[3] is an another pattern recognition method which has been found to give very good results for different problems approximately for ten years. SVM projects the original input space to a high (sometimes infinite) dimensional input space through a dot product, and it finds the maximum margin classifier in that dimension. The projection enables SVM to model complex relationships between inputs and outputs. Even though the data may not be separable in the original input space, it could be separated in the projected space. SVM can be used both for classification and regression problems. In this work, mySVM, which is a C++ tool by Stefan Rüping is used [12].

3.2. Feature selection

Feature selection methods are one of subgroups of dimensionality reduction methods. Feature projection methods like PCA and ICA produce linear combinations of original features. Especially when acquiring or saving mesurements are expensive in terms of money or time, feature selection methods are prefered to feature projection methods. Forward and backward feature selection methods are wrapper feature selection methods. They do a deterministic search in the feature space, including or excluding one feature at each step. Forward and backward feature selection can be used when the number of original features are moderately large.

3.3. Performance Measurement of Pattern Recognition Methods

Let $i = 1..n$ show the prediction time, Y_i show the real observation value and Y_i^* show the model output for time i. Relative Mean Error (RME) is used as a performance measure of speed prediction:

$$RME = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i - Y_i^*}{Y_i}\right| \qquad (1)$$

4. Experimental Results

4.1. Speed prediction using past speed measurements on the sensor itself or neighboring sensors

The three RTMSs used in this section are selected on FSM Bridge which has very heavy traffic during rush hours. RTMS s13 and s59 are on the roads connecting Maslak and Levent to FSM Bridge respectively. RTMS s268 is on the FSM Bridge itself and measures the traffic incoming from both s15 and s59 (Please see Figure 1).

Figure 2 shows the percentage error of estimated speed 5 minutes ahead for s268 sensor, using backward feature selection together with SVM and kNN. s268 knn and svm show the case of using the speed measurements from the sensor s268 only. s268 knn multi and s268 svm

Figure 1*: The location of s13, s59 and s268 RTMS*

multi show the case of using measurements from all three RTMSs to predict speed at s268. According to the figure, SVM is more successful than kNN for single sensor or multiple sensor cases.

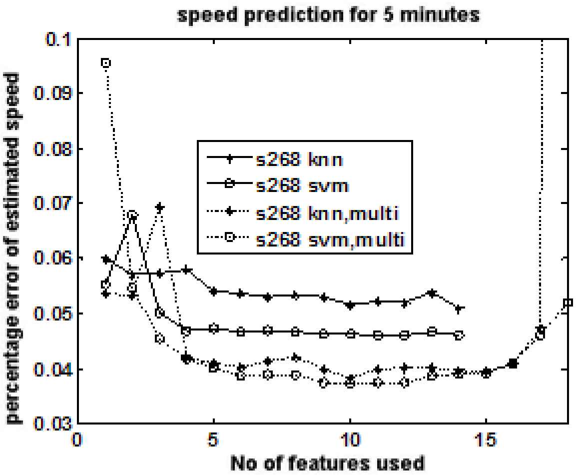

Figure 2*: Backward feature selection for speed prediction 5 minutes ahead on s268.*

Table 1: The most important feature for three RTMS location with SVM method.

Feature	S13	S59	S268
1	**-1440**	**-10080**	**-5**
2	**-10080**	**-1440**	**-1440**
3	**-30**	**-5**	**-10080**
4	**-10**	**-60**	**-10**
5	**-15**	**-35**	**-50**
6	**-40**	**-30**	**-35**
7	-20	-20	-60
8	-45	-50	-25
9	-25	-15	-30
10	-50	-45	-45
11	-55	-55	-40
12	-35	-40	-55
13	-60	-25	-15
14	-5	-10	-20

Approximately 6 speed features are enough for speed prediction and using more speed measurements does not much affect the percentage of error. Selected 6 speed features with backward feature selection are shown on Table 1 for s268, s13 and s59. For all 3 RTMSs, both long time speed measurements (-1440=1 day, -10080=1 week) and short time measurements (in 1 hour) are important in speed prediction.

4.2. Finding dependent sensors using correlations between sensors

In the previous section, it has been shown that using measurements from other sensors could improve speed prediction performance. Since sensors may need calibration or may be broken sometimes, it would be very helpful if sensors which are dependent on each other and hence could be used for speed prediction for each other could be found. Considering the fact that there are 327 different sensor locations, an automated method is needed to find sensors which could be helpful to each other for speed prediction. For all 327 sensor locations, the dependent sensors can be found by means of the correlations between speed measurements of these sensors. For each sensor location the most correlated sensor locations are found and speed measurements from this sensor are used for speed prediction together with the measurements from the sensor itself. Table 2 shows the highest correlated sensors for a number of sensors.

Table 2: Highest correlated sensor.

Sensor 1	Sensor 2
s95	s73
s266	ts266
s156	ts156
s73	s4
s95	s4
s174	s171

When the physical locations of these sensors are examined, they are found to be on the same road or on the roads connected to each other. For example s95, s73 and s4 which are higly correlated sensors follow one after another on the same highway as can be seen on Figure 3.

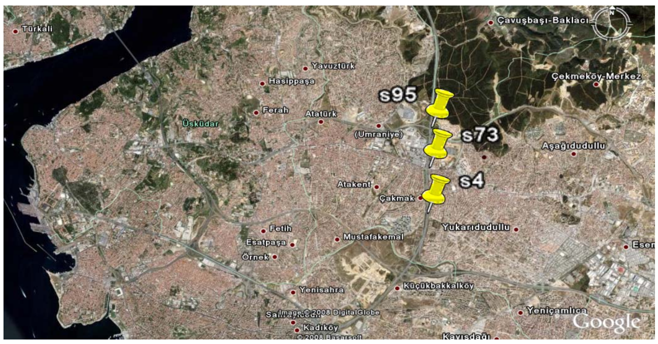

Figure 3*: The location of s95, s73 and s4 which are on the way of FSM Bridge – TEM Anatolian*

Similarly, as can be seen on Figure 4, s266 and ts266 are almost at the same point but measure speed at different directions. s266 measures the speed of the highway from Sarıyer to Maslak, ts266 measures the speed of the highway from Maslak to Sarıyer.

Figure 4*: The location of s266 and ts266 which are on the way of Sariyer - Maslak*

4.3. Speed prediction using correlated sensors in addition to the sensor itself

We performed a number of experiments using correlated sensors. Figures 5, 6 and 7 show the percentage error of speed estimation when the measurements from the sensor itself or the correlated sensor(s) in addition are used. In all the figures, SVM is used with the features obtained from the feature selection process.

In Figure 5, for the speed prediction on s95, s95, s95+s4 and s95+s4+s73 are used. Figure shows that using s4 in addition to s95 reduces the error, while incorporating s4 also does not result in any further improvement.

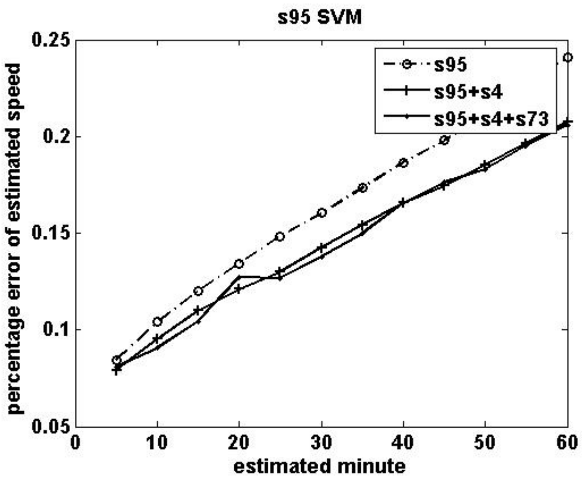

Figure 5: Prediction of s95 with SVM and with speed predictions further in the future (x axis).

Figures 6 and 7 also confirm the consistent usefullness of the correlated sensors for RTMS locations ts266 and s171 respectively.

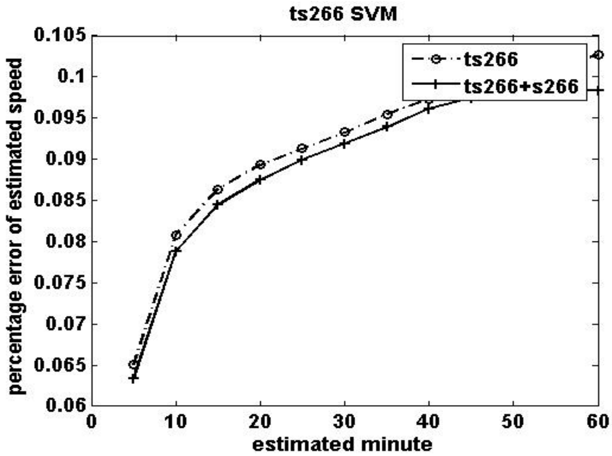

Figure 6: Prediction of ts266 SVM with increasing prediction time lap.

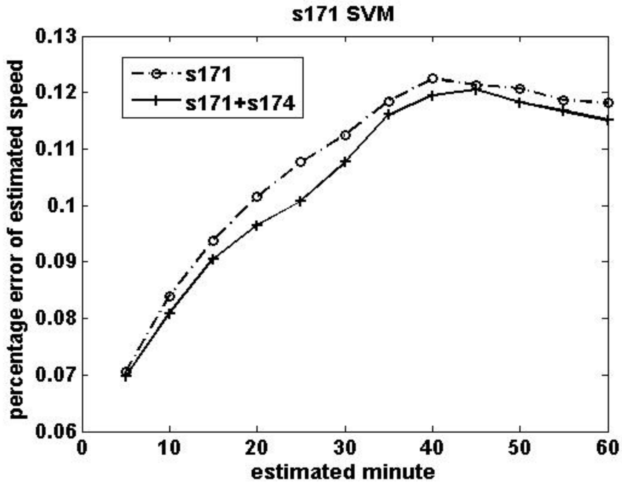

Figure 7: Prediction of s171 speed.

4.4. Speed prediction using the correlated sensors only

When a sensor needs to be calibrated, is broken or can not be reached, it would be very desirable to estimate the speed on that sensor by means of the measurements from correlated sensors. In this section the success of speed prediction using only the correlated sensors is examined. Figure 8 shows the estimation error when s95 itself, s73 or s73+s4 are used to estimate speed at s95. As can be seen from the figure, performance of using s73 only is only slightly worse than using s95 itself, while using s73+s4 gives better results for s95 speed prediction than using s95 itself.

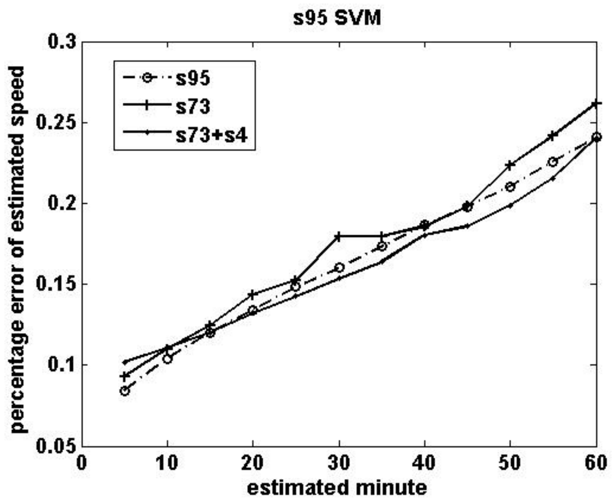

Figure 8: *Prediction of s95 speed.*

As shown in figures 9 and 10, using the correlated sensors only as opposed to the sensor itself results in only 3% more error for RTMSs s156 and ts266. Therefore, the correlated sensors are promising for sensor replacement and validation.

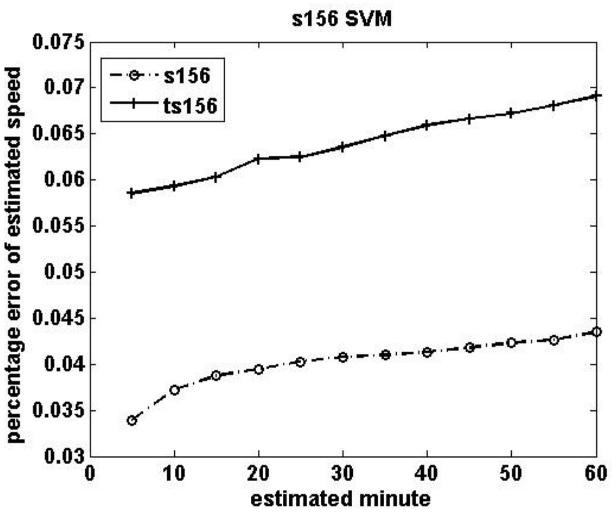

Figure 9: *Prediction of s156 speed.*

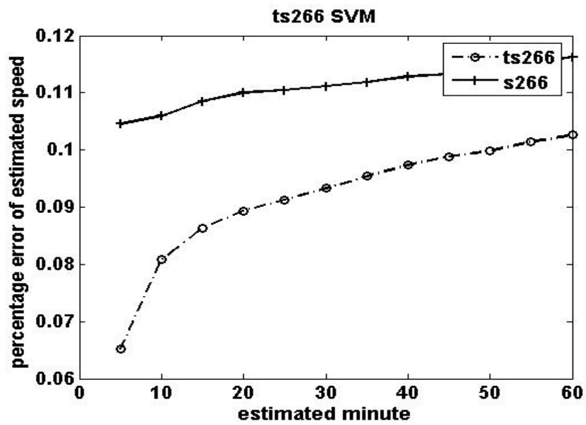

Figure 10: *Prediction of ts266 speed.*

5. Conclusions

In this study, we performed short time speed prediction using both SVM and kNN methods. It is seen that generally SVM produces better results than kNN method for this problem. When comparing best features with feature selection methods, it is seen that success ratio does not change when 6 or more features are used, therefore 6 speed measurements from the past are usually enough for speed prediction. The speed measurements from a day or a week before are usually among the 6 features selected. kNN actually performs worse as more features are used, whereas SVM is affected less by the redundant features.

As the speed predictions further in time are performed, both SVM and kNN's performance decrease and approach each other.

We also outline a correlation based method to identify sensors that could be used to improve the prediction performance for a specific sensor location. Using the correlated sensors in addition to the sensor itself results in better speed prediction performance. On the other hand, correlated sensors can also be used to predict speed for each other. This would be very desirable when the speed measurements from a sensor are unavailable or unreliable. Highly correlated sensors are usually found out to be physically also close to each other.

Acknowledgements

We would like to thank the Istanbul Municipality Traffic Control Center for making the speed measurements available on the internet and also for sharing their data at the earlier stages of our work.

References

[1] Alpaydin, E., ``*Introduction to Machine Learning*'', MIT Press, 2004.

[2] Bin, Y., Zhongzhen, Y. Baozhen, Y., "Bus Arrival Time Prediction Using Support Vector Machines", *Journal of Intelligent Transportation Systems,* Volume 10, Issue 4 October 2006, 151 – 158.

[3] Burges, C.J.C., "A Tutorial on Support Vector Machines for Pattern Recognition", *Data Mining and Knowledge Discovery,* Vol. 2, Number 2, 1998, 121-167.

[4] Chrobok, R., Kaumann, O., Wahle, J., Schreckenberg, M., "Three Categories of Traffic Data: Historical, Current, and Predictive", the *9th IFAC Syposium Control in Transportation Sytems*, 2000, 250-255.

[5] Hobeika, A.G. and Kim, C.K., "Traffic-flow-prediction systems based on upstream traffic", *Proceedings of Vehicle Navigation and Information Systems Conference,* 31 Aug-2 Sep 1994, 345 – 350.

[6] Kwon, J., Coifman, B., Bickel, P., ``Day-to-day travel time trends and travel time prediction from loop detector data", *Transportation Research Record*, (1554), 2000, 120-129.

[7] Lingras, P., and Mountford, P., "Time Delay Neural Networks Designed Using Genetic Algorithms for Short Term Inter-City Traffic Forecasting Engineering of Intelligent Systems", *14th International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems, IEA/AIE 2001,* Budapest, Hungary, June 4-7, 2001, Proceedings, 290-299.

[8] Mark, C.D., Sadek, A.W., Rizzo, D.,"Predicting experienced travel time with neural networks: a PARAMICS simulation study", *The 7th International IEEE Conference on Intelligent Transportation Systems*, 3-6 Oct. 2004, 906 - 911

[9] Park, D., and Ritett, L. R., "Forecasting Multiple-Period Freeway Link Travel Times Using Modular Neural Networks," *77th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 1998.

[10] Park, D. and Rilett, L. R., "Forecasting freeway link travel times with a multilayer feedforward neural network", *Computer-Aided Civil and Infrastructure Engineering,* 14(5), 357–367, 1999.

[11] Rice, J., van Zwet, E., "A simple and effective method for predicting travel times on freeways", *Intelligent Transp. Systems*, *IEEE Proceedings,* 227 -232, 2001.

[12] Ruping, S., "mySVM software", Available: http://www-ai.cs.uni-dortmund.de/SOFTWARE/MYSVM

[13] Sun, H., Liu, H., and Ran, B., "Short Term Traffic Forecasting Using the Local Linear Regression Model", *Transportation Research Record*, 2003.

[14] W.C.Van Lint, W.C., Hoogendoorn, S.P., and van Zuylen, H.J., "Robust and adaptive travel time prediction with neural networks," *Proceedings of the 6th annual TRAIL Congress* (part 2), December 2000.

[15] Wu, C.H., Ho, J.M., D.T., Lee, "Travel-time prediction with support vector regression", *Intelligent Transportation Systems*, 5(4), 276 – 28, Dec. 2004.

[16] Zhang, Z., Rice, J., and Bickel, P., "*Empirical Comparison of Travel Time Estimation Methods"*, Report for MOU 353, UCB-ITS-PRR-99-43, ISSN1055-1425, December 1999