Travel-time prediction with support vector regression — Source link

Chun-Hsin Wu, Jan-Ming Ho, Der-Tsai Lee

Published on: 01 Dec 2004 - IEEE Transactions on Intelligent Transportation Systems (IEEE)

Topics: Support vector machine

Travel-Time Prediction With Support Vector Regression

Chun-Hsin Wu, *Member, IEEE*, Jan-Ming Ho, *Member, IEEE*, and D. T. Lee, *Fellow, IEEE*

***Abstract*—Travel time is a fundamental measure in transportation. Accurate travel-time prediction also is crucial to the development of intelligent transportation systems and advanced traveler information systems. In this paper, we apply support vector regression (SVR) for travel-time prediction and compare its results to other baseline travel-time prediction methods using real highway traffic data. Since support vector machines have greater generalization ability and guarantee global minima for given training data, it is believed that SVR will perform well for time series analysis. Compared to other baseline predictors, our results show that the SVR predictor can significantly reduce both relative mean errors and root-mean-squared errors of predicted travel times. We demonstrate the feasibility of applying SVR in travel-time prediction and prove that SVR is applicable and performs well for traffic data analysis.**

***Index Terms*—Intelligent transportation systems (ITSs), support vector machines, support vector regression (SVR), time series analysis, travel-time prediction.**

I. Introduction

TRAVEL-TIME data are the raw elements for a number of performance measures in many transportation analyzes. They can be used in transportation planning, design and operations, and evaluation. Especially, travel-time data are critical pretrip and *en route* information in advanced traveler information systems. They are very informative to drivers and travelers to make decision or plan schedules. With precise travel-time prediction, a route-guidance system can suggest optimal alternate routes or warn of potential traffic congestion to users; users can then decide the best departure time or estimate their expected arrival time based on predicted travel times.

Travel-time calculation depends on vehicle speed, traffic flow, and occupancy, which are highly sensitive to weather conditions and traffic incidents. These features make travel-time predictions very complex and difficult to reach optimal accuracy. Nonetheless, daily, weekly, and seasonal patterns can still be observed at a large scale. For instance, daily patterns distinguish rush hour and late-night traffic and weekly patterns distinguish weekday and weekend traffic, while seasonal patterns distinguish winter and summer traffic. The time-varying feature germane to traffic behavior is the key to travel-time modeling.

Manuscript received December 1, 2003; revised August 1, 2004. This work was supported in part by the Academia Sinica, Taiwan, under Thematic Program 2001–2003. The Associate Editor for this paper was F.-Y. Wang.

C. H. Wu is with the Department of Computer Science and Information Engineering, National University of Kaohsiung, Kaohsiung 811, Taiwan, and with the Institute of Information Science, Academia Sinica, Taipei 115, Taiwan (e-mail: wuch@iis.sinica.edu.tw).

J.-M. Ho and D. T. Lee are with the Institute of Information Science, Academia Sinica, Taipei 115, Taiwan (e-mail: hoho@iis.sinica.edu.tw; dtlee@iis.sinica. edu.tw).

Digital Object Identifier 10.1109/TITS.2004.837813

Since the creation of support vector machine (SVM) theory by Vapnik of the AT&T Bell Laboratories [1], [2], there have been intensive studies on SVM for classification and regression [3]–[5]. SVM is quite satisfying from a theoretical point of view and can lead to great potential and superior performance in practical applications. This is largely due to the structural risk minimization (SRM) principle in SVM, which has greater generalization ability and is superior to the empirical risk minimization (ERM) principle as adopted in neural networks. In SVM, the results guarantee global minima, whereas ERM can only locate local minima. For example, in the training process of neural networks, the results give out any number of local minima that are not promised to include global minima. Furthermore, SVM is adaptive to complex systems and robust in dealing with corrupted data. This feature offers SVM a greater generalization ability that is the bottleneck of its predecessor, the neural network approach.

The rapid development of SVMs in statistical learning theory encourages researchers to actively apply SVM to various research fields. Traditionally, many studies focus on the application of SVM to document classification and pattern recognition [2]. For intelligent transportation systems (ITSs), there also are many works applying SVM to vision-based intelligent vehicles, such as vehicle detection [6], [7], traffic-pattern recognition [8], and head recognition [9]. These research results evidence the feasibility of SVM in ITS.

Recently, the application of SVM to time-series forecasting, called support vector regression (SVR), has also shown many breakthroughs and plausible performance, such as forecasting of financial market [10], forecasting of electricity price [11], estimation of power consumption [12], and reconstruction of chaotic systems [13]. Except for traffic-flow prediction [14], however, there are few SVR results on time-series analysis for ITS. Since there are many successful results of time-varying applications with SVR prediction, it motivates our research in using SVR for travel-time modeling.

In this paper, we use SVR to predict travel time for highway users. It demonstrates that SVR is applicable to travel-time prediction and outperforms many previous methods. In Section II, we describe the travel-time prediction problem more formally. In Section III, we introduce SVR briefly. In Section VI, we explain our experimental procedure. Then, we present the methods and results of different travel-time predictors in Sections V and VI, respectively. Section VII concludes this paper.

II. Travel-Time Calculation and Prediction

Travel time is the time required to traverse a link or a route between any two points of interest. There are two approaches

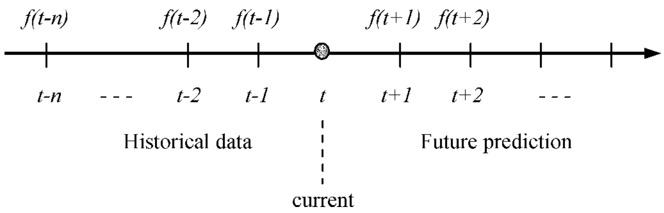


Fig. 1. Travel-time prediction problem. Assume the current time is t.

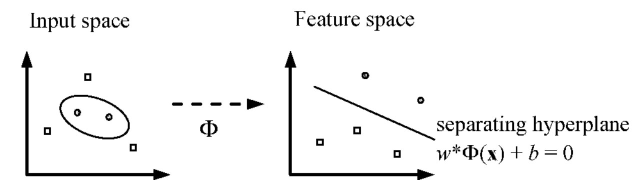


Fig. 2. Basic idea of SVM to solve the binary classification problem, separating circular balls from square tiles.

to calculating travel times: link measurement and point measurement [15]. In the link-measurement approach, link or route travel time is directly measured between two points of interest by using active test vehicles, passive probe vehicles, or license-plate matching. In the point-measurement approach, however, travel time is estimated or inferred indirectly from the traffic data measured by point-detection devices on the roadway or roadside, such as loop detectors, laser detectors, and video cameras. Generally speaking, link-measurement approaches can collect more precise and experienced travel-time data, but point-measurement approaches can be deployed more cost effectively to obtain real-time travel-time data.

There are three categories of traffic data: historical, current, and predictive [16]. Usually, travel-time prediction can be distinguished into two main approaches: statistical models and analytical models. Statistical models can be characterized as data-driven methods that generally use a time series of historical and current traffic variables such as travel times, speeds, and volumes as input. In Fig. 1, suppose that it currently is time t. Given the historical travel-time data $f(t-1), f(t-2), \ldots$, and $f(t-n)$ at time $t-1, t-2, \ldots, t-n$, respectively, we can predict the future values of $f(t+1)$, $f(t+2), \ldots$, by analyzing historical data set. Hence, future values can be forecast based on the correlation between the time-variant historical data set and its outcomes. Numerous statistical methods on the accurate prediction of travel time have been proposed, such as the ARIMA model [17], linear model [18]–[21], and neural networks [22]–[24].

The main idea of traffic forecasting in statistical models is based on the fact that traffic behaviors possess both partially deterministic and partially chaotic properties. Forecasting results can be obtained by reconstructing the deterministic traffic motion and predicting the random behaviors caused by unanticipated factors. On the other hand, analytical models predict travel times by using microscopic or macroscopic traffic simulators, such as METANET [25], [26], NETCELL [27], and MITSIM [28]. They usually require dynamic outside diameter (OD) matrices as input and the predicted travel times evolve naturally from the simulation results.

III. SVR

As shown in Fig. 2, the basic idea of SVM is to map the training data from the input space into a higher dimensional feature space via function Φ and then construct a separating hyperplane with maximum margin in the feature space. Given a training set of data $x_i \subset R^{\mathrm{n}}$, $i = 1, \ldots, l$, where l corresponds to the size of the training data and $y_i = \pm 1$ class labels, SVM will find a hyperplane direction w and an offset scalar b such that $f(x) = w * \Phi(x) + b \geq 0$ for positive examples and $f(x) = w * \Phi(x) + b \leq 0$ for negative examples. Consequently, although we cannot find a linear function in the input space to decide what type the given data is, we can easily find an optimal hyperplane that can clearly discriminate between the two types of data.

Consider a set of training data $\{(x_1, y_1), \ldots, (x_l, y_l)\}$, where each $x_i \subset R^{\mathrm{n}}$ denotes the input space of the sample and has a corresponding target value $y_i \subset R$ for $i = 1, \ldots, l$, where l corresponds to the size of the training data [4], [5]. The idea of the regression problem is to determine a function that can approximate future values accurately.

The generic SVR estimating function takes the form

$$f(x) = \big(w \cdot \Phi(x)\big) + b \tag{1}$$

where $w \subset R^{\mathrm{n}}$, $b \subset R$, and Φ denotes a nonlinear transformation from R^{n} to high-dimensional space. Our goal is to find the value of w and b such that values of x can be determined by minimizing the regression risk

$$R_{\mathrm{reg}}(f) = C \sum_{i=0}^{\ell} \Gamma\big(f(x_i) - y_i\big) + \frac{1}{2}\|w\|^2 \tag{2}$$

where $\Gamma(.)$ is a cost function, C is a constant, and vector w can be written in terms of data points as

$$w = \sum_{i=1}^{\ell} (\alpha_i - \alpha_i^*)\, \Phi(x_i). \tag{3}$$

By substituting (3) into (1), the generic equation can be rewritten as

$$\begin{aligned} f(x) &= \sum_{i=1}^{\ell} (\alpha_i - \alpha_i^*) \big(\Phi(x_i) \cdot \Phi(x)\big) + b \\ &= \sum_{i=1}^{\ell} (\alpha_i - \alpha_i^*)\, k(x_i, x) + b. \end{aligned} \tag{4}$$

In (4), the dot product can be replaced with function $k(x_i, x)$, known as the kernel function. Kernel functions enable the dot product to be performed in high-dimensional feature space using low-dimensional space data input without knowing the transformation Φ. All kernel functions must satisfy Mercer's condition that corresponds to the inner product of some feature space. The RBF is commonly used as the kernel for regression

$$k(x_i, x) = \exp\left\{-\gamma |x - x_i|^2\right\}. \tag{5}$$

Some common kernels are shown in Table I. In our studies, we have experimented with these three kernels.

TABLE I
COMMON KERNEL FUNCTIONS

Kernel	Function
Linear	$x*y$
Polynomial	$[(x*x_i)+1]^d$
Radial Basis Function (RBF)	$\exp\{-\gamma\lvert x-x_i\rvert^2\}$

The ε-insensitive loss function is the most widely used cost function [5]. The function is in the form

$$\Gamma\left(f(x)-y\right)=\begin{cases}|f(x)-y|-\varepsilon, & \text{for } |f(x)-y|\geq\varepsilon\\ 0, & \text{otherwise.}\end{cases}\quad(6)$$

By solving the quadratic optimization problem, the regression risk in (2) and the ε-insensitive loss function (6) can be minimized

$$\frac{1}{2}\sum_{i,j=1}^{\ell}(\alpha_i^*-\alpha_i)\left(\alpha_j^*-\alpha_j\right)k(x_i,x_j)-\sum_{i=1}^{\ell}\alpha_i^*(y_i-\varepsilon)-\alpha_i(y_i+\varepsilon)$$

subject to

$$\sum_{i=1}^{\ell}\alpha_i-\alpha_i^*=0,\quad \alpha_i,\alpha_i^*\in[0,C].\quad(7)$$

The Lagrange multipliers α_i and α_i^* represent solutions to the above quadratic problem, which act as forces pushing predictions toward target value y_i. Only the nonzero values of the Lagrange multipliers in (7) are useful in forecasting the regression line and are known as support vectors. For all points inside the ε tube, the Lagrange multipliers equal to zero do not contribute to the regression function. Only if the requirement $|f(x)-y|\geq\varepsilon$ (see Fig. 3) is fulfilled, Lagrange multipliers may be nonzero values and used as support vectors.

The constant C introduced in (2) determines penalties to estimation errors. A large C assigns higher penalties to errors so that the regression is trained to minimize error with lower generalization, while a small C assigns fewer penalties to errors. This allows the minimization of margin with errors, thus higher generalization ability. If C goes to infinity, SVR would not allow the occurrence of any error and results in a complex model, whereas when C goes to 0, the result would tolerate a large amount of errors and the model would be less complex.

Now, we have solved the value of w in terms of the Lagrange multipliers. For the variable b, it can be computed by applying the Karush–Kuhn–Tucker (KKT) conditions that, in this case, imply that the product of the Lagrange multipliers and constrains has to equal to 0

$$\begin{aligned}\alpha_i\left(\varepsilon+\zeta_i-y_i+(w,x_i)+b\right)&=0\\ \alpha_i^*\left(\varepsilon+\zeta_i^*+y_i-(w,x_i)-b\right)&=0\end{aligned}\quad(8)$$

and

$$(C-\alpha_i)\zeta_i=0$$

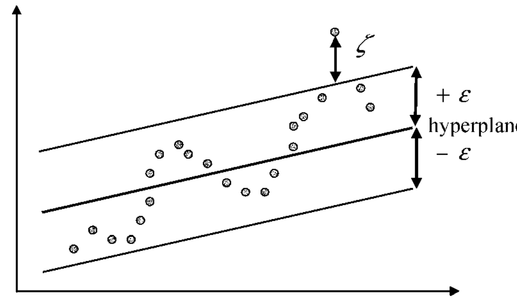


Fig. 3. SVR to fit a tube with radius ε to the data and positive slack variables ζ_i measuring the points lying outside of the tube.

$$(C-\alpha_i^*)\,\zeta_i^*=0\quad(9)$$

where ζ_i and ζ_i^* are slack variables used to measure errors outside the ε tube. Since α_i, $\alpha_i^*=0$, and $\zeta_i^*=0$ for $\alpha_i^{\alpha}\in(0,C)$, b can be computed as

$$\begin{aligned}b&=y_i-(w,x_i)-\varepsilon\quad \text{for } \alpha_i\in(0,C)\\ b&=y_i-(w,x_i)+\varepsilon\quad \text{for } \alpha_i^*\in(0,C).\end{aligned}\quad(10)$$

Putting it all together, we can use SVM and SVR without knowing the transformation. We need to experiment kernel functions; penalty C, which determines the penalties to estimation errors; and radius ε, which determines the data inside the ε tube to be ignored in regression.

IV. EXPERIMENTAL PROCEDURE

A. Data Preparation

The traffic data is provided by the Intelligent Transportation Web Service Project (ITWS) [29], [30] at Academia Sinica, a governmental research center based in Taipei, Taiwan. The Taiwan Area National Freeway Bureau (TANFB) constantly collects vehicle speed information from loop detectors that are deployed at 1-km intervals along the Sun Yet-Sen Highway. The TANFB web site provides the raw traffic information source, which is updated once every 3 min. The loop detector data is employed to derive travel time indirectly: the travel-time information is computed from the variable speed and the known distance between detectors.

Since traffic data may be missed or corrupted, we select a better portion of the dataset of the highway between February 15 and March 21, 2003. During this five-week period, there are no special holidays and the data loss rate is not over some threshold value, which could bias our results if not properly managed. We use data from the first 28 d as the training set and use the last 7 d as our testing set. We examine the travel times over three different distances: from Taipei to Chungli, Taichung and Kaohsiung, which cover 45-, 178-, and 350-km stretches, respectively. In addition, we examine the travel times of a 45-km distance between 7:00 and 10:00 AM further, since the travel time of a short distance in rush hour changes more dynamically. Fig. 4 shows the travel-time distribution of the short distance on a daily and weekly basis, respectively. We can find the daily

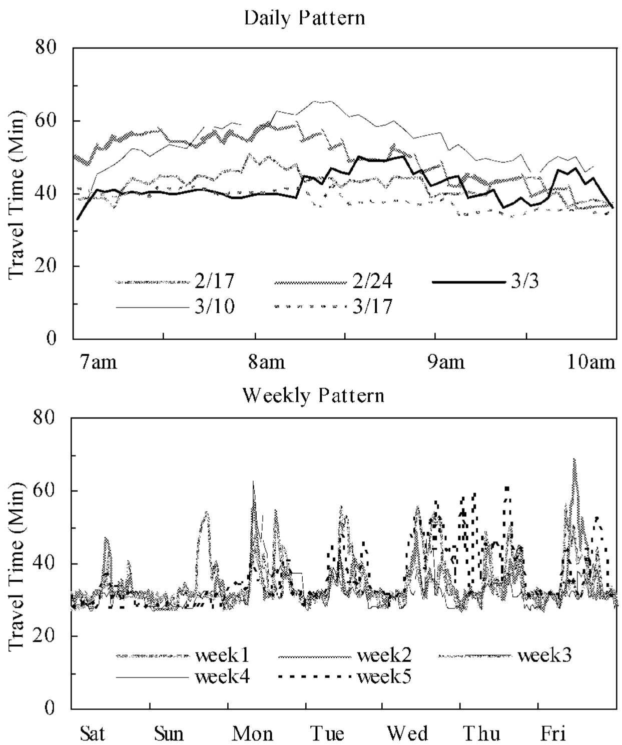


Fig. 4. Daily and weekly travel-time distributions traveling from Taipei to Chungli, a 45-km stretch, between 7:00 and 10:00 AM for five Wednesdays and five weeks between February 15 and March 21, 2003.

similarities and the instant dynamics from the daily and weekly patterns.

B. Prediction Methodology and Error Measurements

Suppose that the current time is t and we want to predict $y(t+l)$ at the future time $t+l$ with the knowledge of the value $y(t-n)$, $y(t-n+1), \ldots, y(t)$ for past time $t-n, t-n+1, \ldots, t$, respectively. The prediction function is expressed as

$$y(t+l) = f\big(t, l, y(t), y(t-l), \ldots, y(t-n)\big).$$

We examine the travel times of different prediction methods for departing from 7:00–10:00 AM during the last week between March 15 and March 21, 2003. Relative mean errors (RME) and root-mean-squared errors (rmse) are applied as performance indices

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

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i - Y_i^*}{Y_i}\right|^2}$$

where Y_1 is the observation value and Y_1^* is the predicted value.

V. Travel-Time Predicting Methods

To evaluate the applicability of travel-time prediction with SVR, some common baseline travel-time prediction methods are exploited for performance comparison.

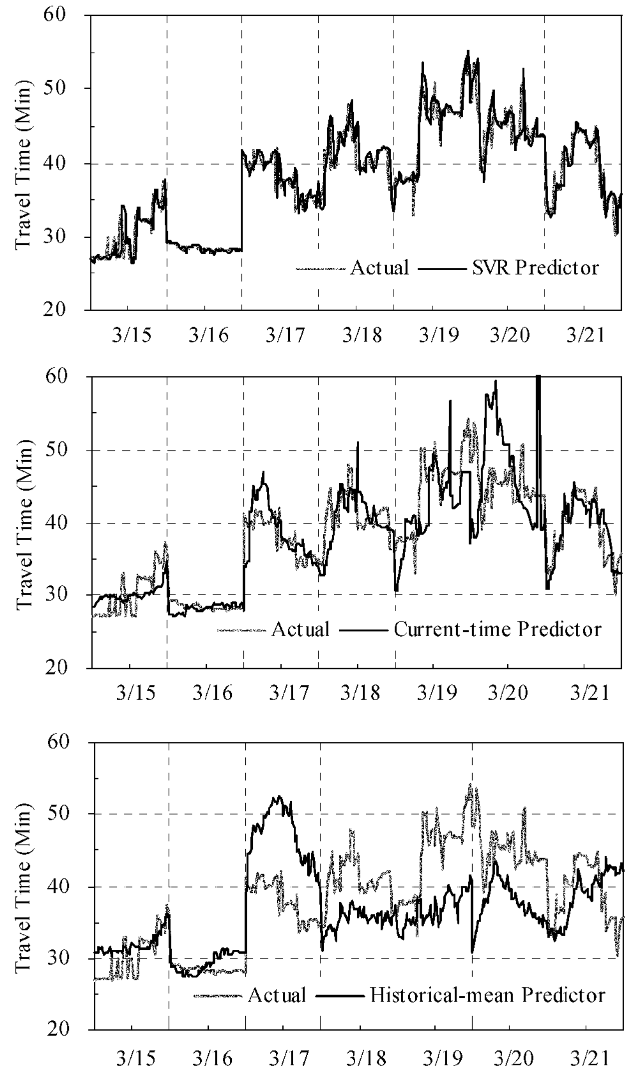


Fig. 5. Comparisons of predicted travel times over short distance in rush hour using different predicting methods.

A. SVR Prediction Method

As discussed previously, there are many parameters that must be set for travel-time prediction with SVR. We have tried several combinations and finally chose a linear function as the kernel for performance comparison with ε =0.01 and C =1000. In our experiences, however, the RBF kernel also performed as well as a linear kernel in many cases. The SVR experiments were done by running mySVM software kit with training window size equal to five [31].

B. Current Travel-Time Prediction Method

This method computes travel time from the data available at the instant when prediction is performed [24]. The travel time is defined by

$$T(t, \Delta) = \sum_{i=0}^{L-1} \frac{x_{i+1} - x_i}{\nu(x_i, t-\Delta)}$$

where Δ is the data delay, L is the number of sections, $(x_{i+1} - x)$ denotes the distance of a section of a highway, and $v(x_i, t-\Delta)$ is the speed at the start of the highway section.

TABLE II
PREDICTION RESULTS IN RME AND RMSE OF DIFFERENT PREDICTORS FOR TRAVELING DIFFERENT DISTANCES (ALL TESTING DATA POINTS)

RME	Current-time Predictor	Historical-mean Predictor	SVR Predictor
45 km (Taipei-Chungli)	9.29%	12.52%	3.91%
161 km (Taipei-Taichung)	3.88%	5.01%	1.71%
350 km (Taipei-Kaohsiung)	2.85%	2.56%	0.96%

RMSE	Current-time Predictor	Historical-mean Predictor	SVR Predictor
45 km (Taipei-Chungli)	28.75%	16.20%	6.79%
161 km (Taipei-Taichung)	9.98%	6.66%	2.57%
350 km (Taipei-Kaohsiung)	5.49%	3.42%	1.33%

TABLE III
PREDICTION RESULTS FOR THE TESTING DATA POINTS THAT HAVE GREATER PREDICTION ERRORS ($>= 5\%$) IN ANY ONE OF THE PREDICTORS

RME	Current-time Predictor	Historical-mean Predictor	SVR Predictor
45 km (Taipei-Chungli)	10.53%	14.31%	4.42%
161 km (Taipei-Taichung)	5.85%	7.81%	2.38%
350 km (Taipei-Kaohsiung)	6.13%	4.90%	1.21%

RMSE	Current-time Predictor	Historical-mean Predictor	SVR Predictor
45 km (Taipei-Chungli)	31.19%	17.55%	7.35%
161 km (Taipei-Taichung)	13.81%	9.00%	3.26%
350 km (Taipei-Kaohsiung)	10.29%	5.66%	1.63%

C. Historical Mean Prediction Method

This is the travel time obtained from the average travel time of the historical traffic data at the same time of day and day of week

$$\overline{T}(t) = \frac{1}{w}\sum_{i=1}^{w} T(i,t)$$

where w is the number of weeks trained and $T(i,t)$ is the past travel time at time t of historical week i.

VI. RESULTS

The experimental results of travel-time prediction over a short distance in rush hour are shown in Fig. 5. As expected, the historical-mean predictor cannot reflect the traffic patterns that are quite different from the past average and the current-time predictor is usually slow to reflect the changes of traffic patterns. Since SVR can converge rapidly and avoid local minimum, the SVR predictor performs very well in our experiments.

The results in Table II show the RME and rmse of different predictors for different travel distances over all the data points of the testing set. They show that the SVR predictor reduces both RME and rmse to less than half of those achieved by the current-time and historical-mean predictors for all different distances.

In our experiments, as the traveling distance increases, the number of free sections increases more than the number of busy sections, such that the travel time of long distance is dominated by the time to travel-free sections. So it is not surprising that all three of the predictors predict well for long distance (350 km), but this makes it difficult to compare the performances of the three predictors. For this reason, we specifically examine the testing data points where the predicted error of any predictor is larger than or equal to 5%. As shown in Table III, the SVR predictor not only improves the overall performance, but also significantly reduces the prediction errors for the cases where there are worse prediction errors in any one of the predictors.

VII. CONCLUSION

Support vector machine and SVR have demonstrated their success in time-series analysis and statistical learning. However, little work has been done for traffic data analysis. In this paper, we examine the feasibility of applying SVR to travel-time prediction. After numerous experiments, we propose a set of SVR parameters that can predict travel times very well. The results show that the SVR predictor significantly outperforms the other baseline predictors. This evidences the applicability of SVR to traffic data analysis.

REFERENCES

[1] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York: Springer, 1995.
[2] ——, "An overview of statistical learning theory," *IEEE Trans. Neural Networks*, vol. 10, pp. 988–999, Sept. 1999.
[3] S. R. Gunn, "Support vector machine for classification and regression," Tech. Rep., Univ. Southampton, Southampton, U.K., May 1998.
[4] K.-R. Müller, A. J. Smola, G. Rätsch, B. Schölkopf, J. Kohlmorgen, and V. Vapnik, "Predicting time series with support vector machine," in *Proc. Int. Conf. Artificial Neural Networks (ICANN'97)*, W. Gerstner, A. Germond, M. Hasler, and J.-D. Nicoud, Eds., 1997, pp. 999–1004. Springer LNCS 1327.
[5] ——, "Using support vector support machines for time series prediction," in *Advances in Kernel Methods*, B. Schölkopf, C. J. C. Burges, and A. J. Smola, Eds. Cambridge, MA: MIT Press, 1999, pp. 242–253.
[6] Z. Sun, G. Bebis, and R. Miller, "Improving the performance of on-road vehicle detection by combining Gabor and wavelet features," in *Proc. IEEE 5th Int. Conf. Intelligent Transportation Systems*, 2002, pp. 130–135.
[7] D. Gao, J. Zhou, and L. Xin, "SVM-based detection of moving vehicles for automatic traffic monitoring," in *Proc. IEEE 4th Int. Conf. Intelligent Transportation Systems*, 2001, pp. 745–749.
[8] J. T. Ren, X. L. Ou, Y. Zhang, and D. C. Hu, "Research on network-level traffic pattern recognition," in *Proc. IEEE 5th Int. Conf. Intelligent Transportation Systems*, 2002, pp. 500–504.

[9] R. Reyna, A. Giralt, and D. Esteve, "Head detection inside vehicles with a modified SVM for safer airbags," in *Proc. IEEE 4th Int. Conf. Intelligent Transportation Systems*, 2001, pp. 500–504.

[10] H. Yang, L. Chan, and I. King, "Support vector machine regression for volatile stock market prediction," in *Proc. Intelligent Data Engineering and Automated Learning*, H. Yin, N. Allinson, R. Freeman, J. Keane, and S. Hubbard, Eds., 2002, pp. 391–396. Springer LNCS 2412.

[11] D. C. Sansom, T. Downs, and T. K. Saha, "Evaluation of support vector machine based forecasting tool in electricity price forecasting for Australian National Electricity Market participants," *J. Elect. Electron. Eng. Austr.*, vol. 22, pp. 227–234, 2002.

[12] B. J. Chen, M. W. Chang, and C. J. Lin. Load forecasting using support vector machines: A study on EUNITE competition 2001. *EUNITE Competition for Smart Adaptive System* [Online] http://www.eunite.org

[13] D. Matterra and S. Haykin, "Support vector machines for dynamic reconstruction of a chaotic system," in *Advances in Kernel Methods*, B. Schölkopf, C. J. C. Burges, and A. J. Smola, Eds. Cambridge, MA: MIT Press, 1999, pp. 211–241.

[14] A. Ding, X. Zhao, and L. Jiao, "Traffic flow time series prediction based on statistics learning theory," in *Proc. IEEE 5th Int. Conf. Intelligent Transportation Systems*, 2002, pp. 727–730.

[15] S. Turner, W. Eisele, R. Benz, and D. Holdener, "Travel Time Data Collection Handbook," Federal Highway Administration, Report FHWA-PL-98-035, Mar. 1998.

[16] R. Chrobok, O. Kaumann, J. Wahle, and M. Schreckenberg, "Three categories of traffic data: Historical, current, and predictive," in *Proc. 9th Int. Fed. Automatic Control Symp. Control in Transportation Systems* , 2000, pp. 250–255.

[17] E. Fraschini and K. Ashausen, *Day on Day Dependencies in Travel Time: First Result Using ARIMA Modeling*: ETH, IVT Institut für Verkehrsplanung, Transporttechnik, Strassen-und Eisenbahnbau, Feb. 2001.

[18] X. Zhang, J. Rice, and P. Bickel, "Empirical comparison of travel time estimation methods,", Tech. Rep. Dept. Stat., Univ. California, Berkeley UCB-ITS-PRR-99-43, Dec. 1999.

[19] J. Rice and E. van Zwet, "A simple and effective method for predicting travel times on freeways," in *Proc. IEEE 4th Int. Conf. Transportation Systems*, 2001, pp. 227–232.

[20] H. Sun, H. Liu, and B. Ran, "Short term traffic forecasting using the local linear regression model," presented at the 82nd Annual Meeting of the Transportation Research Board, Washington, DC, Jan. 2003.

[21] X. Zhang and J. A. Rice, "Short-term travel time prediction using a time-varying coefficient linear model," Tech. Rep. Dept. Statistics, Univ. California, Berkeley, Mar. 2001.

[22] J. W. C. van Lint, S. P. Hoogendoorn, and H. J. van Zuylen, "Robust and adaptive travel time prediction with neural networks," presented at the *Proc. 6th Annual Transport, Infrastructure and LogisticsCongr.*, Delft, The Netherlands, Dec. 2000.

[23] S. Li, "Nonlinear combination of travel-time prediction model based on wavelet network," in *Proc. IEEE 5th Int. Conf. Intelligent Transportation Systems*, 2002, pp. 741–746.

[24] D. Park and L. R. Ritett, "Forecasting multiple-period freeway link travel times using modular neural networks," presented at the 77th Annu. Meeting Transportation Research Board, Washington, DC, Jan. 1998.

[25] A. Kotsialos, M. Papageorgiou, C. Diakaki, Y. Pavlis, and F. Middelham, "Traffic flow modeling of large-scale motorway networks using the macroscopic modeling tool METANET," *IEEE Trans. Intell. Transport. Syst.*, vol. 3, pp. 282–292, Dec. 2002.

[26] A. Messmer and M. Papageorgiou, "METANET: A macroscopic simulation program for motorway networks," *Traffic Eng. Contr.*, vol. 31, no. 549, pp. 466–470, 1990.

[27] R. Cayford, W. H. Lin, and C. F. Daganzo, "The NETCELL simulation package: Technical description," Univ. California, Berkeley, CA PATH Res. Rep. UCB-ITS-PRR-97-23, 1997.

[28] Q. Yang and H. N. Koutsopoulos, "A microscopic traffic simulator for evaluation of dynamic traffic management systems," *Transport. Res.*, pt. C, vol. 4, no. 3, pp. 113–129, 1996.

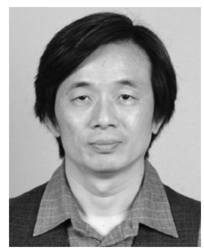
[29] C. H. Wu, D. C. Su, J. Chang, C. C. Wei, J. M. Ho, K. J. Lin, and D. T. Lee, "An advanced traveler information system with emerging network technologies," in *Proc. 6th Asia-Pacific Conf. Intelligent Transportation Systems Forum*, Taipei, Taiwan, 2003, pp. 230–231.

[30] C. H. Wu, D. C. Su, J. Chang, C. C. Wei, K. J. Lin, and J. M. Ho, "The design and implementation of intelligent transportation web services," in *Proc. IEEE Conf. E-Commerce*, 2003, pp. 49–52.

[31] S. Ruping. mySVM Software [Online] http://www-ai.cs.uni-dortmund.de/SOFTWARE/MYSVM/

Chun-Hsin Wu (S'92–M'01) received the B.S. degree in computer science and the Ph.D. degree in information engineering from the National Taiwan University, Taiwan, in 1992 and 1999, respectively.

He was a Postdoctoral Fellow and an Assistant Research Fellow with the Institute of Information Science, Academia Sinica, Taiwan, from 1999 to 2003. He currently is an Assistant Professor in the Department of Computer Science and Information Engineering, National University of Kaohsiung, Taiwan, and has been a Joint-Appointed Assistant Research Fellow with the Institute of Information Science, Academia Sinica, since August 1, 2004. His research interests include intelligent transportation systems, embedded systems, computer networks, peer-to-peer systems, and information retrieval.



Jan-Ming Ho (S'86–M'89) received the B.S. degree in electrical engineering from National Cheng Kung University, Taiwan, in 1978, the M.S. degree from the Institute of Electronics, National Chiao Tung University, Taiwan, in 1980, and the Ph.D. degree in electrical engineering and computer science from Northwestern University, Chicago, IL, in 1989.

He joined the Institute of Information Science, Academia Sinica, Taiwan, as an Associate Research Fellow in 1989 and was promoted to Research Fellow in 1994. His research interests are targeted at the integration of theoretical and application-oriented research, including mobile computing; environments for the management and presentation of a digital archive; management, retrieval, and classification of web documents; continuous video streaming and distribution; video conferencing; real-time operating systems with applications to continuous media systems; computational geometry; combinatorial optimization; very-large-scale integration (VLSI) design algorithms; and the implementation and testing of VLSI algorithms on real designs.

Dr. Ho is a Member of the Association for Computing Machinery (ACM) and is an Associate Editor of IEEE TRANSACTIONS ON MULTIMEDIA. He was Program Chair of the Symposium on Real-time Media Systems, Taipei, from 1994 to 1998, General Co-Chair of the International Symposium on Multi-Technology Information Processing in 1997, and General Co-Chair of IEEE Real-Time Technology and Applications Symposium 2001. He also was a Steering Committee Member of the VLSI Design/CAD Symposium and a Program Committee Member of several previous conferences, including the International Conference on Distributed Computing Systems 1999, the IEEE Workshop on Dependable and Real-Time E-Commerce Systems (DARE'98), etc.

D. T. Lee (S'75–M'78–SM'84–F'92) received the B.S. degree in electrical engineering from the National Taiwan University, Taiwan, in 1971 and the M.S. and Ph.D. degrees in computer science from the University of Illinois, Urbana-Champaign, in 1976 and 1978, respectively.

He has been with the Institute of Information Science, Academia Sinica, Taiwan, since July 1, 1998, where he currently is a Distinguished Research Fellow and Director. Prior to joining the Institute, he was a Professor in the Department of Electrical and Computer Engineering, Northwestern University, Chicago, IL, where he had worked since 1978. His research interests include the design and analysis of algorithms, computational geometry, very-large-scale integration (VLSI) layout, web-based computing, algorithm visualization, bio-informatics, digital libraries, and advanced IT for intelligent transportation systems. He has published over 120 technical articles in scientific journals and conference proceedings and also holds three U.S. patents and one R.O.C. patent. He is an Editor of *Algorithmica*, *Computational Geometry: Theory & Applications*, *ACM Journal of Experimental Algorithmics*, *International Journal of Computational Geometry and Applications*, and *Journal of Information Science and Engineering*. He is the Series Editor of the Lecture Notes Series on Computing for World Scientific Publishing, Singapore.

He is a Fellow of the Association for Computing Machinery (ACM), President of the Institute of Information and Computing Machinery, and Academician of Academia Sinica.