**SHIQUAN ZHONG**
E-mail: othersubmission@163.com
College of Management and Economics, Tianjin University
Tianjin, 300072, China

**JUANJUAN HU**
E-mail: hjj1008@bjjtxy.bj.cn
1. College of Architecture and Civil Engineering,
Beijing University of Technology
Beijing, 100022, China
2. Transport Management Institute,
Ministry of Transport of the People's Republic of China
Beijing, 101601, China

**SHUIPING KE**, Corresponding author
E-mail: ksppp2003@163.com
College of Management and Economics, Tianjin University
Tianjin, 300072, China

**XUELIAN WANG**
E-mail: wangxuelian1979@163.com
School of Management, Hebei University of Technology
Tianjin 300130, China

**JINGXIAN ZHAO**
E-mail: nzjx2005@163.com
School of Economics and Management,
Tianjin University of Science & Technology
Tianjin, 300222, China

**BAOZHEN YAO**, Corresponding author
E-mail: yaobaozhen@hotmail.com
School of Automotive Engineering,
Dalian University of Technology
Dalian 116024, China

# A HYBRID MODEL BASED ON SUPPORT VECTOR MACHINE FOR BUS TRAVEL-TIME PREDICTION

## ABSTRACT

*Effective bus travel time prediction is essential in transit operation system. An improved support vector machine (SVM) is applied in this paper to predict bus travel time and then the efficiency of the improved SVM is checked. The improved SVM is the combination of traditional SVM, Grubbs' test method and an adaptive algorithm for bus travel-time prediction. Since error data exists in the collected data, Grubbs' test method is used for removing outliers from input data before applying the traditional SVM model. Besides, to decrease the influence of the historical data in different stages on the forecast result of the traditional SVM, an adaptive algorithm is adopted to dynamically decrease the forecast error. Finally, the proposed approach is tested with the data of No. 232 bus route in Shenyang. The results show that the improved SVM has good prediction accuracy and practicality.*

## KEY WORDS

*bus travel time prediction; support vector machine regression; Grubbs' test method; adaptive algorithm;*

## 1. INTRODUCTION

With the development of public transportation system, the adoption of technologies such as bus tracking, location and communication has become popular. These technologies make it available to predict bus arrival time at stops and help to provide reliable service in public transport. The passengers can enable efficient scheduling of their trips using the available accurate bus travel time prediction to avoid waiting times. The transit operators can take proper control actions

(stop-skipping strategy, holding strategy etc.) to adjust bus operation when the bus operation is interrupted using accurate bus travel times.

Usually, variable travel times on links and dwell times at stops make arrival time of transit vehicles at stops in urban transit networks uncertain. Especially when buses are disrupted by stochastic factors, such as traffic accidents, traffic jams, it is a tough task to predict the bus travel time. The past decades have witnessed an increasing passion for the research of bus running time prediction. The literature focuses on time series [1] artificial neural network or support vector machine (SVM) [2-8] and Kalman filtering techniques [9, 10], etc.

The time series mainly rely on the similarity of the historical data and future data. The prediction result will apparently deviate with the variation of the historical data. Besides, the application of time series leads to an obvious lag. The artificial neural network is good at dealing with non-linear planning problems. It has been used for prediction of bus travel time in the recent decades. However, considering the structure confirming of neural network algorithm and local convergence problems, the artificial neural network is not widely used [11, 12]. Kalman filtering technique can be accommodated with irregular variation. It is effective in single-step prediction but lacks in accuracy in multi-step prediction [13] .

Recently, SVM has attracted researchers' attention as a new machine learning method. SVM was developed by Vapnik [14, 15], which is characterized by a specific type of learning algorithms. It has been successfully applied to solve some classic problems, such as incident detection [16], traffic-pattern recognition [17], passenger head recognition [18], and travel time prediction [7, 8, 19-21]. SVM is used to find regular pattern and makes use of them to analyse the unknowns. Based on collected statistics, SVM is available in small samples. As SVM has a strong capacity for learning, it is easy to make balance on data fitting and data generalization.

With the successful application of SVM, this paper proposes a bus travel time prediction model based on SVM. To predict travel times of links in an accurate and timely manner, it is essential to take the traffic conditions, e.g. traffic congestion, etc., into consideration. Since traffic congestions are complicated and difficult to measure, the speed of preceding/current bus on links is used to consider traffic conditions of links [7]. Three kinds of information, route link, velocity, the arriving time at a stop of former previous bus are input data to deal with some unexpected delays.

As the operation of a bus is sensitive to current traffic conditions at the downstream part of the used road network, Grubbs' test method is applied to eliminate the abnormal data from input data to avoid the deviation of results brought by unreasonable data. Many procedures to discard outliers have been proposed, such as the Grubbs' test [22, 23], the Tietjen-Moore test, and the Generalized Extreme Studentized Deviate (ESD) test [24]. In a consecutive procedure like the one by Grubbs, only one observation at a time can be tested as an outlier. In the presence of multiple outliers, these procedures must be used repeatedly until no further outliers can be detected. Tietjen-Moore test is a generalization of the Grubbs' test to the case of more than one outlier. It has the limitation that the number of outliers must be specified exactly. ESD procedure is the recursive version of Grubbs' test method. It requires only an upper bound on the suspected number of the outliers. Grubbs' test is usually used to detect outliers in small samples where the number of outliers is unknown.

SVM can obtain several computing techniques, such as optimization theory, kernel functions, etc. With these auxiliary tools, some problems like small samples, non-linearity, high dimension and optimal in part, can be solved effectively [18, 25-28]. However, there will inevitably come to quadratic programming when the standard SVM progresses the optimization problem, especially with large training data and training online. To solve the problems of large calculations and low real-time conducted by the proposed factors, an adaptive algorithm is put forward to minimize the prediction error.

The adaptive algorithm can improve the training speed of the SVM method to a certain extent. In addition, there remains the advantage of simpliness. Moreover, the adaptive algorithm overcomes the problem of lacking of an appropriate support vector in a standard SVM [8, 29]. Yao et al. [30] applied the adaptive algorithm to adjust the performance of SVM method. They confirmed that the adaptive algorithm is effective in the SVM improving programs. Based on equality constraint, Suykens and Vandewalle [31] proposed the least square support vector machines (LS-SVM). Using the Sheman-Morrison-Woodewalle equation, Chua [32] applied SVM in large samples. However, outcome data lacked stability. Depending on the LS-SVM, Yang et al. [25] proposed the adaptive algorithm. In their study, the approach can determine the number of support vectors.

Unlike previous studies, this paper focuses on bus travel-time prediction problem with an improved support vector machine. To make a better prediction performance, the proposed approach is combined with Grubbs' test method, SVM and an adaptive algorithm. However, there are also some limitations in this study. For example, the bus travel-time prediction of the target bus is less accurate in a specific link with increased bus traffic.

This paper is further organized as follows. Section 2 provides a brief introduction to improved SVM and explains the application of the Grubbs' test method in

data processing. Section 3 presents time prediction process with SVM and an adaptive algorithm to adjust the prediction results. Section 4 reveals results of a numerical test and analysis including performance evaluation of the proposed methodology. Lastly, the conclusions are stated in Section 5.

## 2. MODEL DEVELOPMENTS

In this paper, the improved SVM model, called SVMAA, is combined with a standard SVM model, Grubbs' test method and adaptive filtering method. The structure of the modelling process is presented in *Figure 1*.
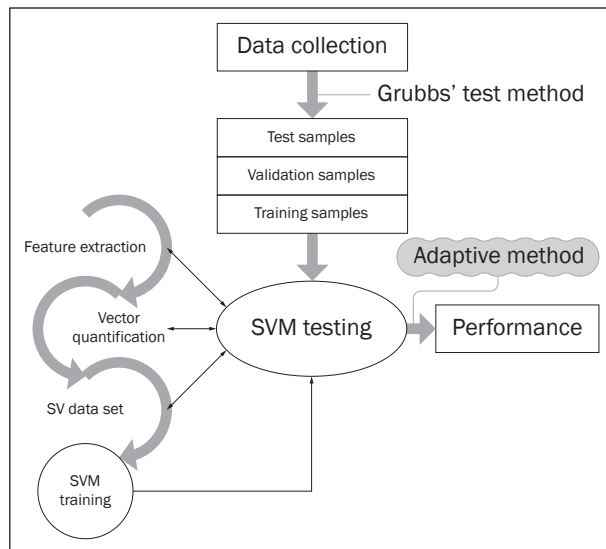


*Figure 1 – Structure of the hybrid model*

First of all, the needed data are collected with GPS technology and investigation. As not all of the data are effective in processing, the outliers among the samples are removed using the Grubbs' test. Then, the effective data set is divided into three parts chronologically: training samples, validation samples and test samples, of which the ratio is 70%, 20% and 10%, respectively. The training samples and validation samples are employed to optimize the parameters of the SVM model, while the test samples are applied to verify the prediction performance of the model. Since not all the results of SVM testing are appropriate, the adaptive filtering method is adopted to select the performance with large absolute error. The detail performance will be introduced in the following sections.

### 2.1 Outliers detection and removal based on Grubbs' test method

In this paper, the predicted bus travel time is predicted under the ordinary circumstance. Therefore, it is the key point to build a universal rule in the collected data set for bus running time prediction. However,

considering the unexpected factors which happen in public transport in some circumstances, such as traffic jams, bad weather, traffic accidents, etc., it will inevitably contain unexpected information, called outliers, which are beyond certain standard deviations (SD) from their respective means. As outliers in the data set may make a disturbance to the prediction accuracy, it is necessary to discard the outliers from the input data of SVM to avoid a disproportionate influence on data analysis although such outliers are sometimes important to the rational result, e.g. traffic accident identification.

Generally, the collected data can be divided into two categories, predictable and stochastic data. The predictable data are generated in ordinary circumstances as the input data of SVM in the bus travel-time prediction. The stochastic data are usually caused by artificial reasons and can be removed by the Grubbs' test methods. It always results in a small amount of special numbers when the data sets are collected in the wrong way.

Grubbs' test is based on the assumption of normality. That is, one should first verify that the data can be reasonably approximated by a normal distribution before applying the Grubbs' test. Grubbs' test detects one outlier at a time. This outlier is expunged from the dataset and the test is iterated until no outliers are detected. However, multiple iterations change the probabilities of detection, and the test should not be used for sample sizes of six or less since it frequently tags most of the points as outliers.

Moreover, when there is an accident in the route, the collected data of the traffic vehicle must be abnormal. The congestion caused by an accident will disrupt the normal operation of the accident bus and its following buses subsequently. That situation is thought as another kind of congestion. Thus, the interference data should be removed in the data set. Since Grubbs' test is simple and easy to apply, it is used to remove the outliers from the data set, assuming the distribution of bus speed on the predicted link is subjected to normal distribution during the studied periods.

Grubbs' test is defined for the hypotheses:

H0: There are no outliers in the data set;

Ha: There is at least one outlier in the data set.

The Grubbs' test statistics is defined in Ram [33]:

$$G = \frac{\max\limits_{i=1,L,n}\left|Y_i - \bar{Y}\right|}{g}, \tag{1}$$

where $\bar{Y}$ and $g$ denote the sample mean and standard deviation, respectively. The Grubbs' test statistics is the largest absolute deviation from the sample mean in units of the sample standard deviation. This is the two-sided version of the test. The Grubbs' test can be also defined as a one-sided test. To test whether the

minimum value is an outlier, the test statistics is:

$$G = \frac{\bar{Y} - Y_{min}}{g} ,$$ (2)

where $Y_{min}$ denotes the minimum value. To test whether the maximum value is an outlier, the test statistics is:

$$G = \frac{Y_{max} - \bar{Y}}{g} ,$$ (3)

where $Y_{max}$ denotes the maximum value. For the two-sided test, the hypothesis of no outliers is rejected at significance level $a$ if:

$$G > \frac{N-1}{\sqrt{N}} \sqrt{\frac{t^2_{a/(2N),N-2}}{N-2+t^2_{a/(2N),N-2}}} ,$$ (4)

where $t^2_{a/(2N),N-2}$ denotes the upper critical value of the $t$-distribution with $N-2$ degrees of freedom and a significance level of $a/(2N)$. For the one-sided tests, replace $a/(2N)$ with $a/N$.

## 2.2 Time prediction process with an adaptive algorithm

After selecting the collected data before time prediction, the SVM model is proposed to predict bus running time during the route links. As not all of the results of the SVM prediction are appropriate, the adaptive filtering method is adopted to select the performance with large absolute error.

### 2.2.1 Support Vector Machines for regression

Given a set of data, points, $(x_1, y_1)$, $(x_2, y_2)$... $(x_m, y_m)$, $x_i$ is the input vector, $x_i \in X \subseteq R^n$; $y_i$ is the desired value, $y_i \in Y \subseteq R^n$; $n$ is the number of training samples. SVM can map $x$ to a high-dimensional space H with a non-linear mapping function $f(x)$:

$$f(x) = \omega \times \varphi(x) + b .$$ (5)

where $\varphi(x)$ represents the high-dimensional feature spaces which are non-linearly mapped from the input space $x$. The coefficients $\omega$ and $b$ are estimated by minimizing the regularized risk function:

$$\frac{1}{2}\|\omega\|^2 + C\frac{1}{n}\sum_{i=1}^{n} N_\varepsilon(y_i, f(x_i)) .$$ (6)

In formula (6), the first term $\frac{1}{2}\|\omega\|^2$, called regularized term, is used to make the function as flat as possible, as well as to improve the controlling capacity. The second term $C\frac{1}{n}\sum_{i=1}^{n} N_\varepsilon(y_i, f(x_i))$ is an empirical risk function, which is defined with different loss function. To control the punishment of samples beyond the error, $C$ is settled to satisfy the condition $C>0$. *Figure 2*

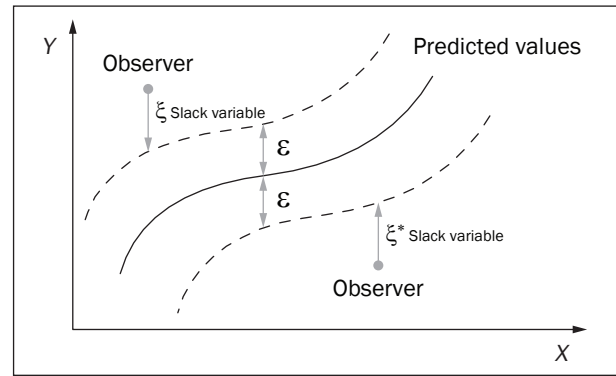presents the performance when applying ε (insensitive loss function).



*Figure 2 – The parameters for the support vector regression*

It is defined as follows for $N_\varepsilon(y_i, f(x_i))$:

$$N_\varepsilon(y_i, f(x_i)) = \max(|y_i - f(x_i)| - \varepsilon, 0) .$$ (7)

If $|y_i - f(x_i)| < \varepsilon$, then $N_\varepsilon(y_i, f(x_i)) = 0$; or else, $|y_i - f(x_i)| \geq \varepsilon$, then $N_\varepsilon(y_i, f(x_i)) = |y_i - f(x_i)| - \varepsilon$. With the introduction of non-negative slack variables, $\xi, \xi^*$, formula (6) can be described as:

$$\min\{\frac{1}{2}\|\omega\|^2 + C\frac{1}{n}\sum_{i=1}^{n}(\xi_i + \xi_i^*)\} ,$$ (8)

$$\text{s.t. } y_i - \omega\varphi(x_i) - b \leq \varepsilon + \xi_i ,$$ (9)

$$\omega\varphi(x_i) + b - y_i \leq \varepsilon + \xi_i^*, \ i = 1, \cdots, n ,$$ (10)

$$\xi_i^* \geq 0 .$$ (11)

Function (8) is a convex quadratic optimization problem. With the introduction of Lagrange Function, the following will be obtained:

$$\omega - \sum_{i=1}^{n}(a_i - a_i^*)x_i = 0 .$$ (12)

Therefore,

$$f(x) = \sum_{i=1}^{n}(a_i - a_i^*)\varphi(x_i)\varphi(x) + b .$$ (13)

With the introduction of kernel function $K(x_i, y_i)$, the formula (13) can be given in the following form:

$$f(x) = \sum_{i=1}^{n}(a_i - a_i^*)K(x_i, x) + b .$$ (14)

Using the above mentioned kernel function, all necessary computations can be performed directly in the input space, without having to compute the map $\varphi(x)$. There are some popular kernel functions, the linear kernel $K(x_i, y_i) = x_i * y_j$, polynomial kernel $K(x_i, y_i) = (x_i * y_j + 1)^r$, etc., where $r$ are the kernel parameters.

### 2.2.2 Adaptive algorithm

By using the SVM model, the potential rule of bus travel is developed through historical data. Though the outliers are eliminated by the Grubbs' test before performing the SVM model, the real-time capacity of the model is weak as it is using an offline training pattern. To solve this problem, an adaptive algorithm is proposed to adjust prediction results of bus travel time dynamically.

The adaptive algorithm introduces an adaptive factor, set as $g_h(m)\left(0 \le g_h(m) \le 1\right)$. It is $g_h(m)$ that is used to adjust the prediction results of bus running time by minimizing the covariance of prediction error. After real-time estimation, iterative update $g_h(m)$ and covariance of prediction error, the optimal value $g_h|(m)$ is obtained. Let $t_h(m)$ indicate the real running time of bus $m$ at route link $h$; let $t_h(m)^*$ indicate the predicted running time of bus $m$ at route link $h$, which is adjusted by the adaptive algorithm [34]. The adaptive algorithm is given with the following equations:

$$t_h(m)^* = (1 - g_h(m))t_h(k)^* + g_h|(m)\Delta h(m-1)^* , \qquad (15)$$

$$t_h(m) = \frac{\xi_h(m-1)}{\xi_h(m-1) + \xi_h(m-1)^*/g_h(k-1)} , \qquad (16)$$

$$\Delta_h(m\text{-}1)^* = t_h(m-1) - t_h(m-1)^* , \qquad (17)$$

$$\Delta_h(m\text{-}1) = t_h(m-1) - t_h(m-1)^* , \qquad (18)$$

$$\xi_h(m-1)^* = E((\Delta_h(m-1)^*)^2) , \qquad (19)$$

$$\xi_h(m-1) = E((\Delta_h(m-1))^2) , \qquad (20)$$

where $g_h(m)$ represents an adaptive factor; $\Delta_h(m-1)$ and $\Delta_h(m-1)^*$ indicate the error of travel time prediction of bus $m-1$ at route link $h$ with SVM model and an adaptive algorithm, respectively. $\xi_h(m-1)$ and $\xi_h(m-1)^*$ indicate the covariance of error of prediction running time of bus $m-1$ at route link $h$ with SVM model and an adaptive algorithm, respectively.

### 2.2.3 The hybrid model of travel-time prediction

As the prediction performance of SVM model may include some unexpected results, this paper combines the SVM model and the adaptive algorithm to predict the bus travel time. It is called the hybrid model of time prediction, denoted as SVMAA. In this model, the traditional SVM model is used to predict bus deadheading time at certain stops. The arrival time of the current bus at a certain stop is calculated according to the deadheading time and the arrival time for the former bus. Then, the outcomes of arrival time are selected with the adaptive algorithm. Lastly, the bus running time is worked out after the selection.

Among the variables that may contribute to the variation of bus travel, five input variables and an output

variable are used in the SVM model. In this model $d$ denotes the input variables, which consist of the five variables, the bus stop $k$ ($d_1$), the past second stop $k-2$ ($d_2$), the departure time at stop $k-2$ for bus $m-1$ ($d_3$), the departure time at stop $k-2$ for the objective bus $m$($d_4$), and the travel times of current segment $k-2{\to}k$ ($d_5$). Let $O$ denote output vector, the bus travel times on the route segment between two adjacent points. Variables $d_1$, $d_2$ are self-explanatory. The route segment variable identifies the section between the current stop and the next stop at which the arrival times are to be predicted. $d_3$ and $d_4$ are denoted as $T_{k-2}^{m-1}$ and $T_{k-2}^{m}$ respectively. The deadheading time which is departure time interval between bus $m$ and bus $m-1$ from stop $k-2$ can be presented as $T_{k-2}^{m} - T_{k-2}^{m-1}$. It varies with different bus stops. Assume that there are stops numbered from the origin terminal through the destination terminal. *Figure 3* describes variables $d_3$ and $d_4$ in more details.
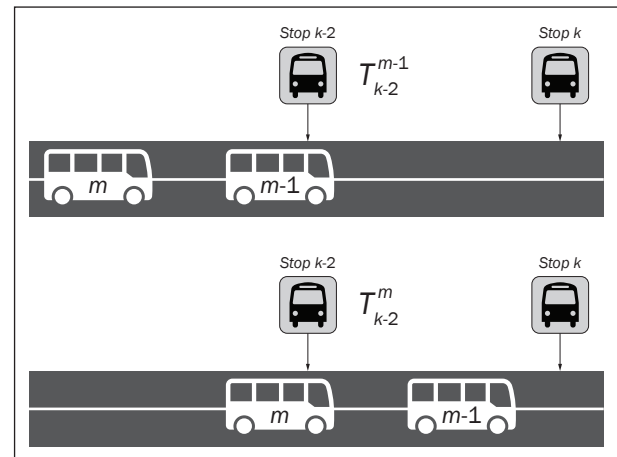


*Figure 3 – Illustration of $d_3\left(T_{k-2}^{m-1}\right)$ and $d_4\left(T_{k-2}^{m}\right)$*

Variable $d_5$ refers to travel times at segment $k-2{\to}k$ for each former bus (bus $m-1$, bus $m-2$,..., bus $m-r$). Variable $d_5$ is expected to estimate the traffic conditions of the current segment. The latest travel times on the predicted segment will be updated after a bus finishes its travel on the predicted segment. When vehicle m reaches stop $k-2$, the input variables of $d_5$ at segment $k-2{\to}k$ are travel times of the former r buses at the segment, denoted by $V_{m-1},...,V_{m-r}$, which are shown in *Figure 4*.
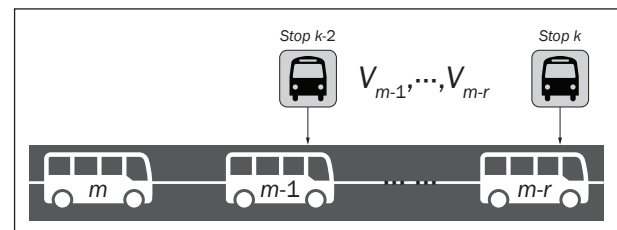


*Figure 4 – Illustration of $d_5\left(V_{m-1},\cdots,V_{m-\gamma}\right)$*

Using different kernel functions, one can construct different learning machines with arbitrary types of

decision surfaces. Let $\hat{t}_h(m)$ express the prediction deadheading time of bus $m$ and bus $m-1$ at stop $k$. *Figure 5* shows the progress of $\hat{t}_h(m)$ prediction with the SVM model.
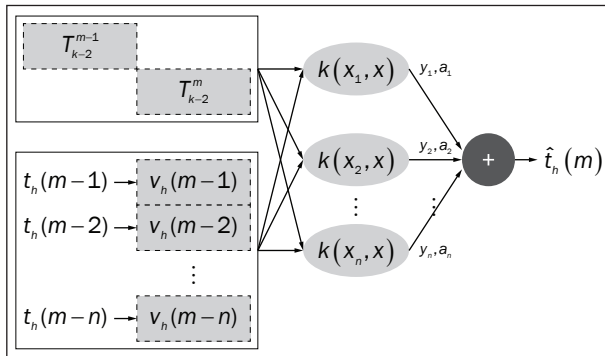


*Figure 5 – Progress of SVM prediction*

After the outcome of $\hat{t}_h(m)$, let $T_k^m$ denote the arrival time of a bus $m$ at stop $k$. Then, $T_k^m$ can be calculated as follows:

$$T_k^m = T_k^{m-1} + \hat{t}_h(m). \tag{21}$$

Using the adaptive algorithm, the prediction of the arrival time of bus at stop $k$ is dynamically regulated. The progress of the adaptive algorithm is depicted in *Figure 6*.
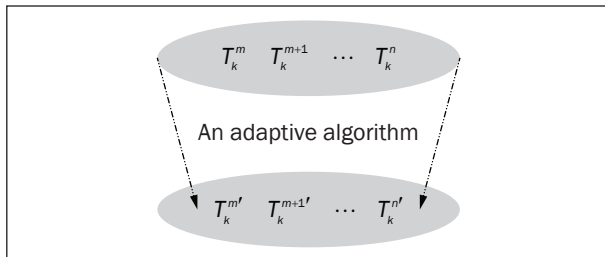


*Figure 6 – Progress of the adaptive algorithm*

After regulation of the SVM model performance, the final outcome of bus arrival time at stop $k$ is predicted, denoted as $T_k^{m'}$.

## 3. NUMERICAL TEST

The presented model for bus travel-time prediction has been tested with the data of transit route No. 232 in Shenyang City, China. The transit route goes from
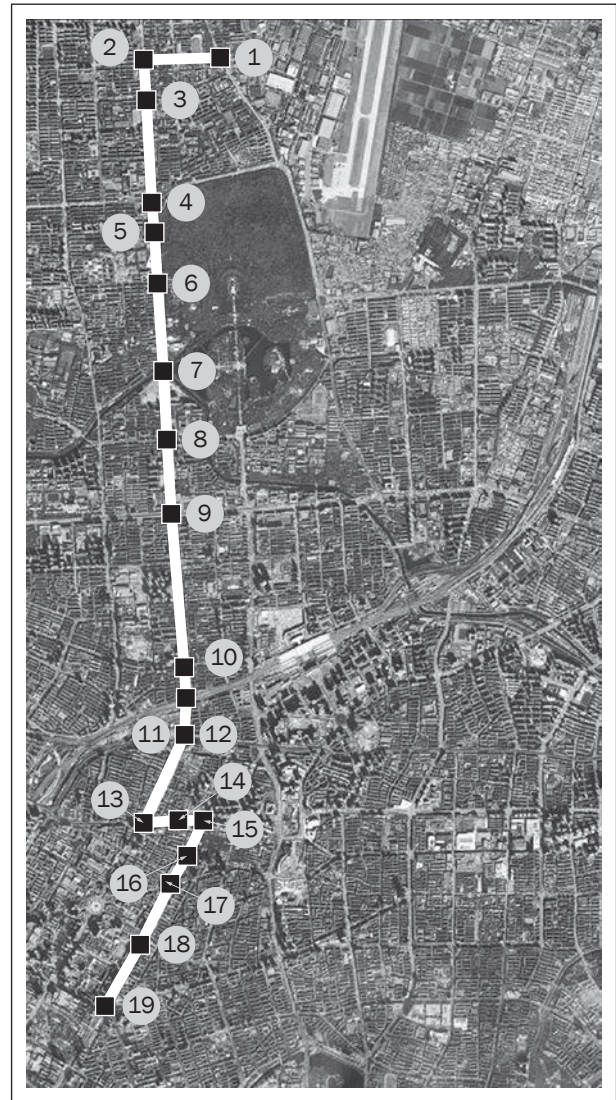


*Figure 7 – Sketch scheme of test route No. 232*

Ling North Street to the city center with 19 stops in total and 10.8 km per direction. In the numerical test, parts of the eastbound direction of the transit route are studied. Part of the routes and bus stops are selected to predict bus running time, which are representative stops on route No. 232, showed in *Figure 7*. The stops from Shenyang Station to Three Taizi are denoted as stop 1, stop 2, ..., stop 19, respectively. The length of each segment of No. 232 is described in *Table 1*. First, the data used in the models are described and then the results are obtained.

*Table 1 – Configuration of transit route No.232*

| Route segment | 1→2 | 2→3 | 3→4 | 4→5 | 5→6 | 6→7 | 7→8 | 8→9 | 9→10 |
|---|---|---|---|---|---|---|---|---|---|
| Length (m) | 435 | 441 | 1,000 | 214 | 619 | 380 | 1,100 | 540 | 352 |
| Route segment | 10→11 | 11→12 | 12→13 | 13→14 | 14→15 | 15→16 | 16→17 | 17→18 | 18→19 |
| Length (m) | 743 | 650 | 670 | 657 | 543 | 834 | 568 | 523 | 507 |

## 3.1 Data collection and processing

In this research, the bus speed is used to reflect the traffic conditions of links at the current time. However, it is difficult to measure the speeds directly. Hence, the approximate speeds are calculated using the link lengths and travel time, as a substitute for the real speed on the corresponding links.

To obtain the data of travel times of links, we conducted an on-board survey of all the No. 232 bus trips from October to November 2012. The collected data consist of the arrival times at 19 stops that are covered by the test beds in eastbound trip at peak period (6:30 A.M. – 7:30 A.M., PP) and off-peak period (10:00 A.M., 11:00 A.M., OP) on weekdays. There are 994 valid trips within this approximate one-month period. After calculating the approximate speeds on links, the data sets should be scaled. In the modelling process, the data sets are linearly scaled to the range from 0 to 1 to ensure the validity of all data and avoid numerical difficulties during the calculation. Before the procedure of model validation, group the speed data chronologically, using the first five weeks as sample data, 6th week and 7th week as the test set and 8th week as the forecasting set.

The data samples are divided into three parts: training samples (70%), validation samples (20%) and test samples (10%).

## 3.2 Model identification and results

The data can be mapped implicitly into a feature space and be enhanced quite efficiently by using a kernel function. Radial Basis Function is selected as the kernel function in this study. There are three parameters while using Radial Basis Function kernels: $c$, $\varepsilon$ and $\gamma$, which are calibrated by grid-search. In grid-search,

all pairs of $c$, $\varepsilon$ and $\gamma$ are tried and the one with the best performance is picked up. For the bus arrival time prediction problem, the three parameters are selected as ($2^{-2}$, $2^{-5}$ and 1.47). The process of the proposed prediction model is shown in *Figure 7*. This paper takes the average of absolute prediction error (MAPE) to test the accuracy of prediction results. MAPE can be described as follows:

$$MAPE = \frac{1}{J}\sum_{j=1}^{J}\frac{\left|t^d(j)-\tilde{t}^d(j)\right|}{t^d(j)}\times100\%,\qquad(22)$$

where $J$ is the number of test sample. $t_d(j)$ and $\tilde{t}^d(j)$ stand for the actual travel time and predicted travel time of sample $j$, respectively.

The predicted arrival time at the six bus stops of bus No. 232 in a weekday is shown in *Figure 8*. The predicted performances of the off-peak and peak time are compared in *Figure 8* as well. The performance in *Figure 8* reflects that MAPE of most stops are below 10%, which means the proposed approach has good prediction accuracy. It is interesting to find that the MAPE of predicted performance is decreasing from the first research stop (the north Shenyang station) to the last one (Songling cultural). The stop with large prediction error is the north Shenyang station. The main reason for the above problem is that the north Shenyang station is the second stop of the east-bound bus route. The prediction of bus travel time at the north Shenyang station is lacking in input data of the bus running time on the former route link, which results in a lower prediction accuracy.

However, a larger MAPE is shown in *Figure 8* at the peak time prediction. The average MAPE in the peak time is about 10.7%, while MAPE at off-peak time is 9.5%. Some bus stops with large passenger interchanges, such as the second province hospital and the Experimental Middle School, have large MAPE.
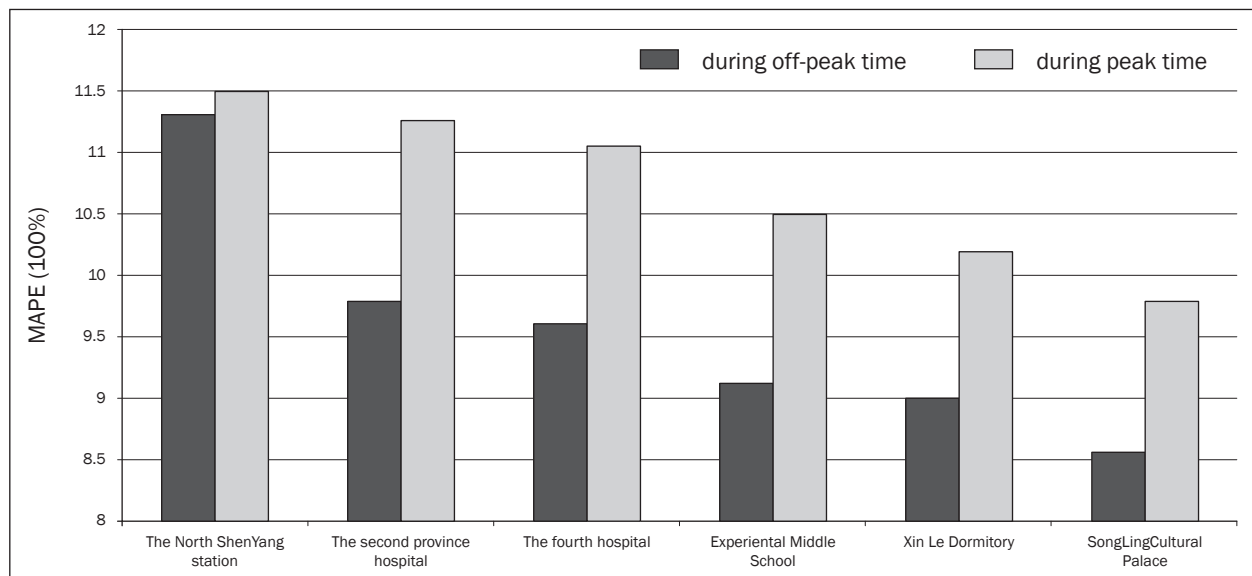


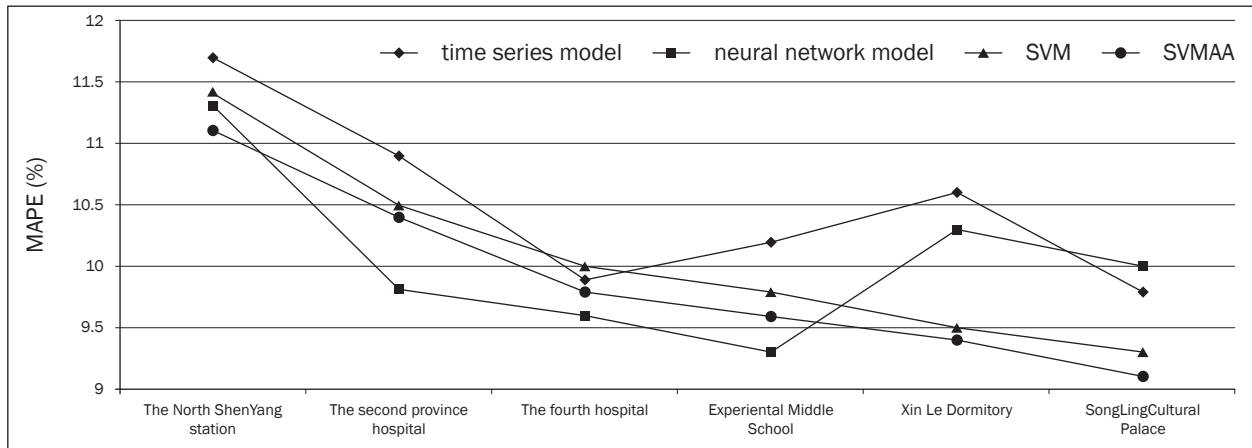*Figure 8 – Prediction results of bus arrival time at main stops*

*Figure 9 – Comparison of forecasting the bus travel time for four models*

This is because the use of the speed of previous buses at the second province hospital segment and the fourth hospital segment cannot effectively describe the current traffic condition since there is traffic congestion during the peak time. However, the average MAPE is still around 10%, which means that the proposed model can be used for bus travel-time prediction during peak time.

To prove the superiority of the proposed approach (SVMAA), we compared the predicted bus travel time (at the off-peak time) with the prediction results offered by other three models: time series model, neural network model and SVM model. Methods and parameter settings of the three models have been adopted in literature. After computing the bus running time of each objected stop with the four different models, the average prediction errors (MAPE) of each model are shown in *Figure 9*.

*Figure 9* describes that MAPE of SVMAA model in five stops is lower than the other three models except for the Shenyang station as lack of information. Taking account of the historical data and without the real-time information of a running bus, MAPE of predicted results of time series model and neural network model are the largest. Both of the SVM model and SVMAA model have better prediction accuracy as they are based on structural adventure and the VC dimensionality theory which is a measure of the capacity of a statistical classification algorithm [35]. SVM can decrease the predicted errors with good capacity of sam-

ples. Compared with SVM, the performance of SVMAA is better with the introduction of Grubbs' test method and an adaptive algorithm. The Grubbs' test method eliminates the outliers so as to obtain a better input. Finally, an adaptive algorithm will adjust the predicted result with the real-time information of the corresponding bus travel time.

## 4. CONCLUSIONS

It is hard to predict the bus running time accurately as the condition of bus operation is full of stochastic events. For the use of SVM a specific function is not needed, since SVM can reflect the relationship in non-linear and the real-time system of input and performance. With a good learning capacity, the SVM is suitable to predict bus travel time. Considering the real-time bus travel-time prediction, the Grubbs' test method is used to select the outliers. Besides, an adaptive algorithm is applied to real-time adjustment of the test result of prediction. Lastly, this paper takes the data of bus route No. 232 in Shenyang to test the proposed approach. The test consequence shows that SVMAA can predict bus travel time efficiently. In addition, to simulate road traffic conditions, the data of bus velocity in the test bus route are selected; it will give rise to a great prediction error for the bus with lower frequency. In the future research, the data of bus velocity of buses on multi-routes will be adopted to simulate road traffic conditions.

## ACKNOWLEDGEMENT

石泉1，胡娟娟2.3，柯水平1，王雪 4， 静 5，姚宝珍6
1．天津大学管理学院 天津300072
2．北京工 大学建工学院，北京 100022
3．交通 部管理干部学院，北京 101601
4．河北工 大学管理学院 天津300401
5．天津科技大学 管学院 天津300222
6．大 理工大学汽 工程学院 大 116026

## 摘要

### 应用基于支持向量机的混合模型进行公交车运行时间预测

在公共交通系 中有效的 行时 十分重要。 了 公交 行时 本文提出了一种改 的支持向量机 模型，并对 方法的有效性 行了 。改 的支持向量机模型 合了 支持向量机模型、格拉布斯 方法和公交 行时 的自适 算法。由于 差主要来源于数据 入，在 用 支持向量机模型之前，本文采用格拉布斯 方法来剔除异常数据。另外， 了减少历史数据在不同的 段对 支持向量机 果的影响，本文引入了自适 算法来动 地减少 差。最后，本文用沈阳市232公交 路的数据 行模型 。 果表明改 的SVM具有 好的 准确性和实用性。

### 关键词

公交 行时 ；支持向量机；格拉布斯 方法；自适 算法 ；

## REFERENCES

[1] Smith BL, Demetsky MJ. Short-Term Traffic Flow Prediction: Neural Network Approach. Transportation Research Record. 1995;1453:98-104.

[2] Chien SIJ, Ding Y, Wei C. Dynamic Bus Arrival Time Prediction with Artificial Neural Networks. Journal of Transportation Engineering. 2002;128(5):429-438.

[3] Chen M, Liu XB, Xia JX, Chien SI. A Dynamic Bus-Arrival Time Prediction Model Based on APC Data. Computer-Aided Civil and Infrastructure Engineering. 2004;19:364-376.

[4] Acevedo-Rodríguez J, Maldonado-Bascón S, Lafuente-Arroyo S, Siegmann P, López-Ferreras F. Computational load reduction in decision functions using support vector machines. Signal Processing. 2009;89(10):2066-2071.

[5] Dong B, Cao C, Lee SE. Applying support vector machines to predict building energy consumption in tropical region. Energy and Buildings. 2005;37:545-553.

[6] Elish KO, Elish MO. Predicting defect-prone software modules using support vector machines. Journal of Systems and Software. 2008;81(5):649-660.

[7] Yu B, Yang ZZ, Yao BZ. Bus Arrival Time Prediction Using Support Vector Machines. Journal of Intelligent Transportation Systems. 2006;10(4):151-158.

[8] Wu CH, Ho JM, Lee DT. Travel-time prediction with support vector regression. IEEE Transactions on Intelligent Transportation Systems. 2004;5(4):276-281.

[9] Hellinga BR, Fu LP. Reducing bias in probe-based arterial link travel times estimates. Transportation Research Part C. 2002;10:257-273.

[10] Cathey FW, Dailey DJ. A prescription for transit arrival/departure prediction using automatic vehicle location data. Transportation Research Part C. 2003;11: 241-264.

[11] Abbas K, Ehsan M, Saeid N, Doug C, Van LJWC. A genetic algorithm-based method for improving quality of travel time prediction intervals. Transportation Research Part C. 2011;19(6):1364-1376.

[12] Ehsan M, Geoff R, Graham C, Sara M. Prediction intervals to account for uncertainties in neural network predictions: Methodology and application in bus travel time prediction. Engineering Applications of Artificial Intelligence. 2011;24(3):534-542.

[13] Ngoduy D. Applicable filtering framework for online multiclass freeway network estimation. Physica A: Statistical Mechanics and its Applications. 2008;387(2-3):599-616.

[14] Vapnik VN. An Overview of Statistical Learning Theory. IEEE Transactions on Neural Networks. 1999;10(5):988-999.

[15] Vapnik VN. The Nature of Statistical Learning Theory. New York: Springer; 2000.

[16] Yuan F, Cheu RL. Incident detection using support vector machines. Transportation Research Part C. 2003;11:309-328.

[17] Ren JT, Ou XL, Zhang Y, Hu DC. Research on network level traffic pattern recognition. Proceedings of the 5th IEEE Conference on Intelligent Transportation Systems; 2002 Sep 3.6; Singapore. IEEE; 2002. doi: 10.1109/ITSC.2002.1041268

[18] Reyna R, Giralt A, Esteve D. Head detection inside vehicles with a modified SVM for safer air bags. Proceedings of the IEEE Conference on Intelligent Transportation Systems; 2001 Aug 25-29; Oakland, CA. IEEE; 2001. doi: 10.1109/ITSC.2001.948667

[19] Yu B, Yang ZZ, Chen K, Yu B. Hybrid model for prediction of bus arrival times at next station. Journal of Advanced Transportation. 2010;44(3):193-204.

[20] Yu B, Lam WHK, Tam M L. Bus Arrival Time Prediction at Bus Stop with Multiple Routes. Transportation Research Part C. 2011;19(6):1157-1170.

[21] Yu B, Yang ZZ, Li S. Real-Time Partway Deadheading Strategy Based on Transit Service Reliability Assessment. Transportation Research Part A. 2012;46(8):1265-1279.

[22] Grubbs FE. Sample Criteria for Testing Outlying Observations. Annals of Mathematical Statistics. 1950;21(1): 27-58.

[23] Grubbs FE. Procedures for Detecting Outlying Observation in Samples. Technometrics. 1969;11(1):1-21.

[24] NIST/SEMATECH e-Handbook of Statistical Methods; 2010. Available from: http://www.itl.nist.gov/div898/handbook

[25] Yang ZZ, Jin LJ, Wang MH. Forecasting Baltic Panamax Index with Support Vector Machine. Journal of Transportation Systems Engineering and Information Technology. 2011;11(3):50-57.

[26] Moller MF. A Scaled conjugate gradient algorithm for fast supervised learning. Neural Networks. 1993;298(6):523-533.

[27] Yao BZ, Yang CY, Yao JB, Sun J. Tunnel Surrounding Rock Displacement Prediction Using Support Vector Machine. International Journal of Computational Intelligence Systems. 2010;3(6): 843-852.

[28] Yao BZ, Hu P, Zhang MH, Jin MQ. A Support Vector Machine with the Tabu Search Algorithm for Freeway Incident Detection. International Journal of Applied Mathematics and Computer Science. 2014;24(2):397-404.

[29] Cao LJ, Tay FEH. Support Vector Machine with Adaptive Parameters in Financial Time Series Forecasting. IEEE Transactions on Neural Networks. 2003;14 (6):1506-1518.

[30] Yao P, Lu Yh. Neighborhood rough set and SVM based hybrid credit scoring classifier. Expert Systems with Applications. 2011;38 (9):11300-11304.

[31] Suykens JAK, Vandewalle J. Least Squares Support Vector Machine Classifiers. Neural Processing Letters. 1999;9(3):293-300.

[32] Chua KS. Efficient computations for large least square support vector machine classifiers. Pattern Recognition Letters 2003;24(1-3):75-80.

[33] Ram BJ. A recursive version of Grubbs' test for detecting multiple outliers in environmental and chemical data. Clinical Biochemistry. 2010;43(12):1030-1033.

[34] Azadeh A, Rouzbahman M, Saberi M, Valianpour F, Keramati A. Improved prediction of mental workload versus HSE and ergonomics factors by an adaptive intelligent algorithm. Signal Processing. 2013;94:350-358.

[35] Aytug H, Koehler GJ, He L. Risk minimization and minimum description for linear discriminant functions. Informs Journal on Computing. 2008;20(2):317-331.