TRANSPORT
ISSN 1648-4142 print / ISSN 1648-3480 online
2012 Volume 27(2): 158–164
doi:10.3846/16484142.2012.692710

IMPROVED ITERATIVE PREDICTION FOR MULTIPLE STOP ARRIVAL TIME USING A SUPPORT VECTOR MACHINE

Chang-Jiang Zheng[1], Yi-Hua Zhang[2], Xue-Jun Feng[3]

College of Civil and Transportation Engineering, Hohai University,
1 Xikang Road, 210098 Nanjing, PR China
E-mails: [1]zheng@hhu.edu.cn (corresponding author); [2]yhzhang@hhu.edu.cn; [3]fxj@hhu.edu.cn

Submitted 26 March 2011; accepted 1 June 2011

Abstract. The paper presents an improved iterative prediction method for bus arrival time at multiple downstream stops. A multiple=stop prediction model includes two stages. At the first stage, an iterative prediction model is developed, which includes a single stop prediction model for arrival time at the immediate downstream stop and an average bus speed prediction model on further segments. The two prediction models are constructed with a support vector machine (SVM). At the second stage, a dynamic algorithm based on the Kalman filter is developed to enhance prediction accuracy. The proposed model is assessed with reference to data collected on transit route No 23 in Dalian city, China. The obtained results show that the improved iterative prediction model seems to be a powerful tool for predicting multiple stop arrival time.

Keywords: multiple stop, arrival time, prediction, support vector machine, Kalman filter.

1. Introduction

Accurate prediction for bus arrival time plays an important role in advanced public transportation systems (APTS). It can be used for transportation planning and operation. Forecasting bus arrival time is especially critical for operators to make a decision or plan schedules, while it can also reduce passenger waiting time and improve transit service level. Arrival time estimates are subject to a number of factors such as traffic congestion and passenger arrivals at stops. Therefore, predicting arrival time in an accurate and timely manner is a challenging task.

There is plenty of literature on predicting arrival time published in the past decade. Dailey *et al.* (2000) presented an algorithm for predicting bus arrival time based on time and location received from an automated vehicle location system. Chien *et al.* (2002) developed a link-based artificial neural network (ANN) and a stop-based ANN to predict bus arrival time where a dynamic algorithm was also presented to dynamically improve outputs. Cathey and Dailey (2003) made up a general prescription to indentify the factors necessary for forecasting bus arrival/departure. The prescription consisted of a tracker, a filter and a predictor and corresponding algorithms were adopted in the predictor component. Chen *et al.* (2004) introduced a dynamic model consisting of an artificial neural network model and the Kalman filter technique to predict bus arrival times based on data collected by a real world APC. Chien *et al.* (2007) designed a probabilistic model for disseminated bus arrival time which took the total waiting time caused by pre-trip passengers as its aim. Yu *et al.* (2006, 2011) proposed a SVM-based model for forecasting bus arrival time, in which the weather, segment and traffic conditions were considered. Wu *et al.* (2004) provided support vector regression (SVR) to predict highway travel time. Hellinga and Fu (2002) developed a methodology of adjusting bias in arrival time. The methodology can effectively use stratified sampling techniques to provide simulation results for a single intersection approach and an arterial corridor. Li and McDonald (2002) adopted an approach to analyzing the speed time profile reflecting the difference between travel time and the mean travel time of the probe vehicle. Vanajakshi *et al.* (2009) used the Kalman filtering technique to predict bus travel time by using global positioning system data.

Most of the methods proposed in previous work have gained successful results to predict bus arrival time at the immediate downstream stop. To the best of our knowledge, the existing literature is rarely found to predict arrival times at multiple downstream stops as well as at the immediate downstream stop. The paper denotes prediction for arrival time at the immediate

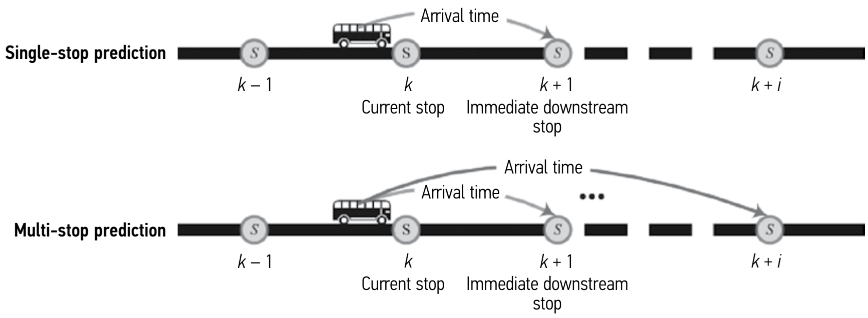

Fig. 1. An example of single-stop and multi-stop predictions

downstream stop (e.g., the stop $k + 1$) as a single-stop prediction while the prediction of arrival times at multiple downstream stops (e.g., the stop $k + 1 \ldots k + i$) is defined as multi-stop prediction.

Single-stop and multi-stop predictions are shown in Fig. 1.

If arrival times at multiple downstream stops were available, it would provide operators with more reliable and timelier information on dispatching and scheduling for transit operation. This can provide passengers with more convenience and greatly improve transit service quality. This paper focuses on predicting bus arrival times at multiple downstream stops.

The support vector machine (SVM) (Vapnik 1999a, 1999b), which is a very specific type of learning algorithms, can map the relationship between the input and output of a complex system. It has been successfully applied to solve some transportation problems (Yu *et al.* 2006; Wu *et al.* 2004; Yuan, Cheu 2003; Ren *et al.* 2002). Furthermore, the studies presented by Yu *et al.* (2006, 2010, 2011) and Wu *et al.* (2004) suggested that the SVM model was suitable for predicting travel time in transportation field. These successful results of time varying applications with SVM prediction motivate our research to use the SVM for modelling prediction for multiple stop arrival time.

The paper first develops the SVM-based model for forecasting bus arrival time at the immediate downstream stop. Then, the SVM-based model for estimating traffic conditions on the road links after the immediate downstream stop is proposed to improve the accuracy of multi-stop prediction. The structure of this paper is organized as follows: Section 2 provides the structure of the iterative method predicting bus arrival times at multiple downstream stops and a dynamic algorithm; Section 3 contains results and analyses including the performance evaluation of the methodology; finally, conclusions are presented in Section 4.

2. Model Development

2.1. Support Vector Machines for Regression

The support vector machine is a kind of a learning machine based on the statistical learning theory developed by Vapnik (1999a, 1999b) and is gaining popularity due to many attractive features. The support vector machine (SVM) implements the principle of structural risk minimization (SRM), which shows to be more superior to a traditional principle of empirical risk minimization (ERM) carried out by conventional neural networks. Thus, the SVM usually achieves optimum network structure. The SVM regression function (Vapnik 1999a, 1999b; Cao, Tay 2003; Dong *et al.* 2005) can be formulated as follows:

$$f(x) = \omega x + b. \tag{1}$$

Coefficients ω and b are estimated by minimizing the following cost function:

$$\frac{1}{2}\|\omega\|^2 + C\frac{1}{l}\sum_{i=1}^{l} L_\varepsilon\left(y_i, f(x_i)\right), \tag{2}$$

where: C is the regularization constant. The first term $\|\omega\|^2$ is called the regularized term and is used for making regular weight sizes and penalizes large weights. The second term $\frac{1}{l}\sum_{i=1}^{l} L_\varepsilon\left(y_i, f(x_i)\right)$ is a penalty function penalizing large errors.

The minimization of Eq. (2) is a standard problem in the optimization theory: minimization with constraints. This can be solved by applying Lagrangian theory and the weight vector. ω equals a linear combination of training data:

$$\omega = \sum_{i=1}^{l}\left(a_i - a_i^*\right) x_i. \tag{3}$$

In this formula, a_i and a_i^* are Lagrange multipliers associated with a specific training point. The asterisks again denote the difference between data above and below the regression line. Putting this formula into Eq. (1), the following solution is obtained with unknown data point x:

$$f(x) = \sum_{i=1}^{l}\left(a_i - a_i^*\right) < x_i \bullet x > + b. \tag{4}$$

By introducing kernel function $K(x_i, x_j)$ Eq. (4) can be rewritten as follows:

$$f(x) = \sum_{i=1}^{l}\left(a_i - a_i^*\right) K\left(x_i, x\right) + b. \tag{5}$$

By using kernels, all necessary computations can be performed directly in input space without computing map x.

2.2. Single Stop Arrival Time Prediction Model

Road traffic condition is one of the most important variables that may contribute to the variation of bus running. However, traffic conditions are complicated and difficult to measure, and thus this research uses bus speed on the road link for estimating traffic conditions of the links. Assume that there are stops numbered from $k - 1$ to $k + i$, as shown in Fig. 2. While vehicle m reaches stop k, arrival time at stop $k + 1$ can be predicted by using a single stop arrival time prediction model.

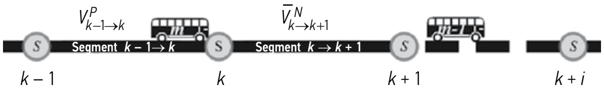

Fig. 2. An example of predicting arrival time at the immediate downstream stop

The speed $\left(V^P_{k-1\to k}\right)$ that the vehicle has just finished on the proceeding segment and the current average speed $\left(\bar{V}^N_{k\to k+1}\right)$ on the predicted segment are used for acting as input variables (Fig. 1). For example, if the current stop is stop k, $V^P_{k-1\to k}$ is speed on the segment between stop $k - 1$ and stop k, and $\bar{V}^N_{k\to k+1}$ is the average speed of several proceeding vehicles that have just passed through stop $k + 1$ on the predicted segment (segment $k \to k + 1$):

$$\bar{V}^N_{k\to k+1} = \frac{1}{n} \times \sum_{j=1}^{n} V^N_{m-j,k\to k+1}, \tag{6}$$

where: $V^N_{m-j,k\to k+1}$ is the observed speeds of proceeding vehicles that have just passed through stop $k + 1$ on segment $k \to k + 1$; n is the number of the considered proceeding vehicles.

The output of the single-stop arrival time prediction model is the forthcoming speed $\left(\hat{V}^P_{k\to k+1}\right)$ of the vehicle on the predicted segment. Then, arrival time $\left(\hat{T}_{k+1}\right)$ at the immediate downstream stop can be indirectly obtained. The structure of the single stop arrival time prediction model is shown in Fig. 3.

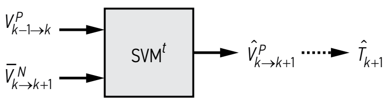

Fig. 3. The structure of the single-stop prediction model

2.3. Average Bus Speed Prediction Model on a Further Segment

Average bus speeds on further segments are predicted by using the observed speeds on the segments in advance. If the single stop method was used to recursively forecast arrival times at the following stops after the immediate downstream stop, prediction errors would greatly increase along the bus route. Even small prediction errors existing at the beginning of the horizon will accumulate, propagate and finally lead to poor prediction accuracy. In fact, there may be some other reasons. One of those is a lack of measurements when predicting arrival times at downstream stops after the immediate downstream stop.

If we assume the current stop is stop k using the single-stop prediction model to forecast arrival times at the stops after stop $k + 1$, traffic conditions on further segments should be estimated. The observed speeds of the vehicles that have just passed through the immediate downstream stop are used as inputs so as to find relation with traffic conditions on further segments.

When forecasting average bus speeds on further segments, there are several inputs: the segment and observed speeds on the immediate downstream segment (segment $k \to k + 1$). The segment is the predicted route section between two stops. The output of the model is the estimated average speed $\left(\hat{\bar{V}}^N_{k+i-1\to k+i}\right)$ on the predicted segment. The proposed average bus speed prediction model on a further segment is displayed in the structure shown in Fig. 4.

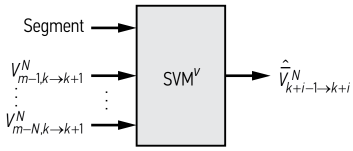

Fig. 4. The structure of the average bus speed prediction model on a further segment

2.4. Dynamic Algorithm

Although speeds in the above models have some 'dynamic' feature, they are still based on a historical data pool of bus trips. To improve prediction accuracy, an adaptive algorithm (Chien *et al.* 2002; Chen *et al.* 2004) should be developed to adjust prediction results in real time.

Therefore, to adjust prediction results, a dynamic algorithm is developed based on the Kalman filtering technique that uses the observed bus information together with the estimated speeds generated by SVM models. The Kalman filter is a minimum mean square error estimator that can estimate an instantaneous 'state' of a linear dynamic system corrupted by white noise. The resulting estimator is statistically optimal with respect to any quadratic function of estimation error. Consider the following state space model:

$$x_m = \Phi_{m-1} x_{m-1} + w_{m-1}; \tag{7}$$

$$y_m = H_{m-1} x_{m-1} + v_{m-1}, \tag{8}$$

where: x_m is the state vector (in this case, it denotes the speed of the current vehicle (m) from the single stop prediction mode or the average speed from the average speed prediction mode on the estimated segment); y_m is the observation vector denoting the observed speed of the current vehicle or the observed average speed on the estimated segment; Φ_{m-1} and H_{m-1} are state transition matrix and observation matrix respectively. Here, x_{m-1} and y_{m-1} denote speed and are one dimensional variables, and thus $\Phi_{m-1} = (1)$ and $H_{m-1} = (1)$. w_{m-1} and v_{m-1} represent the process and measurement noise with zero means and covariance matrices Q_{m-1} and R_{m-1} respectively.

Let $\hat{x}^{-}_{m-1}$ be a prior state estimate at step $m-1$ given knowledge of the prior process, and $\hat{x}_{m-1}$ be a posterior state estimate at step $m-1$ given measurement y_{m-1}. The two state estimates can be computed as follows:

$$\hat{x}^{-}_{m} = \Phi_{m-1}\hat{x}_{m-1}; \tag{9}$$

$$\hat{x}_{m} = \hat{x}^{-}_{m} + K_m\left(y_m - H_m\hat{x}^{-}_{m}\right). \tag{10}$$

Kalman gain K_m reflects the stochastic nature of the process and measurement computed by an optimal linear estimator that minimizes the squared error on the expected value of state estimation $\hat{x}_m$.

We can then define prior and posterior estimate errors as $P^{-}_{m-1} = E\left[\left(x_{m-1} - \hat{x}^{-}_{m-1}\right)^2\right]$ and $P_{m-1} = E\left[\left(x_{m-1} - \hat{x}_{m-1}\right)^2\right]$ respectively. One form of resulting K is given by (11):

$$K_m = P^{-}_{m}H^{T}_{m}\left(H_m P^{-}_{m} H^{T}_{m} + R_m\right)^{-1}. \tag{11}$$

For more details about the Kalman filter, refer to Grewal and Andrews (2001).

2.5. Improved Iterative Prediction Model for Multiple Stop Arrival Times

In the improved iterative prediction model for multiple stop arrival times, the single stop prediction model is first used for forecasting arrival time at the immediate downstream stop. Then, the average speeds on further segments are estimated based on the average bus speed prediction model. Following this, arrival times at further downstream stops are forecasted by using outputs at the previous stops as inputs. For example, forecasted arrival time at stop $k+i$, using the iterative approach, can be computed with the estimated speed on segment $k+I-1 \to k+i$ that can be predicted by two input variables: estimated speeds on segment $k+i-2 \to k+i-1$ and segment $k+i-1 \to k+i$.

$$\hat{T}_{k+i} \to \hat{V}_{k+i-1\to k+i} = f\left(\hat{V}'^{P}_{k+i-2\to k+i-1}, \hat{\bar{V}}'^{N}_{k+i-1\to k+i}\right), \tag{12}$$

where: $\hat{V}'^{P}_{k+i-2\to k+i-1}$ and $\hat{\bar{V}}'^{N}_{k+i-1\to k+i}$ denote the predicted outputs adjusted by the dynamic algorithm.

For example, when the bus arrives at stop k, $\hat{V}^{P}_{k\to k+1}$ can be predicted by the single stop prediction model with $V^{P}_{k-1\to k}$ and $\bar{V}^{N}_{k\to k+1}$. Then, arrival time $\left(\hat{T}_{k+1}\right)$ at the immediate downstream stop $(k+1)$ can be indirectly obtained. To predict arrival time at downstream stop $k+2$, $\hat{V}^{P}_{k\to k+1}$ and $\hat{\bar{V}}^{N}_{k+1\to k+2}$ are acted as two input variables. $\hat{V}^{P}_{k\to k+1}$ can be acquired by the single stop prediction model and $\hat{\bar{V}}^{N}_{k+1\to k+2}$ can be acquired by the average bus speed prediction model. The two variables from the prediction model have some estimation errors. To decrease the estimated errors, $\hat{V}^{P}_{k\to k+1}$ and $\hat{\bar{V}}^{N}_{k+1\to k+2}$ will be adjusted by the dynamic algorithm to $\hat{V}'^{P}_{k\to k+1}$ and $\hat{\bar{V}}'^{N}_{k+1\to k+2}$ respectively. Arrival time at downstream stop $k+2$ can be predicted in Fig. 5. As arrival time at downstream stop $k+2$ is predicted, arrival time from downstream stop $k+3$ to the last one can also be predicted.

Based on the improved iterative prediction model, the dynamic algorithm is introduced to enhance prediction accuracy. The process of the improved iterative prediction model is shown in Fig. 6.

3. Numerical Test

The improved iterative prediction model has been tested using data obtained on transit route No 23 in Dalian city, China, the road that passes from Ligongdongmen to Waiguoyuxueyuan with a total amount of 19 stops at one direction (Yu *et al.* 2012). A part of route segments from stop Tiyuchang (the stop 1) to terminal Waiguoyuxueyuan (the stop 7) is chosen as the study bed (Fig. 7). Assume that prediction action occurs at Stop 1 only, i.e. while a vehicle reaches Stop 1, all arrival times at downstream stops (Stop 2 … Stop 7) can be predicted.

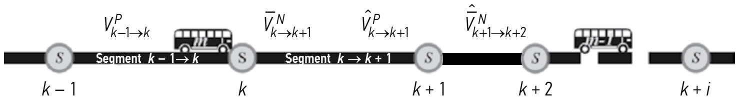

Fig. 5. An example of predicting arrival time at downstream stop $k+2$

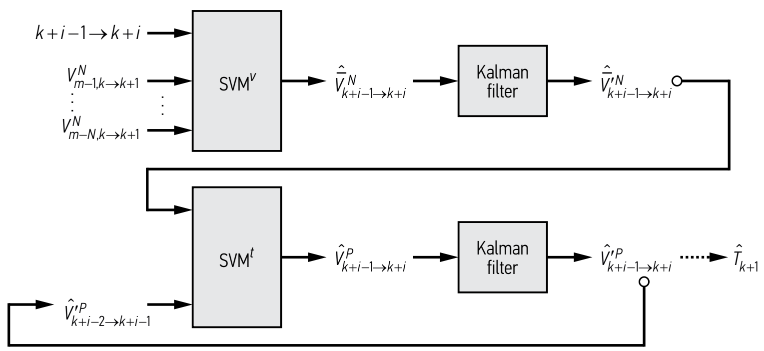

Fig. 6. The structure of the improved iterative prediction model

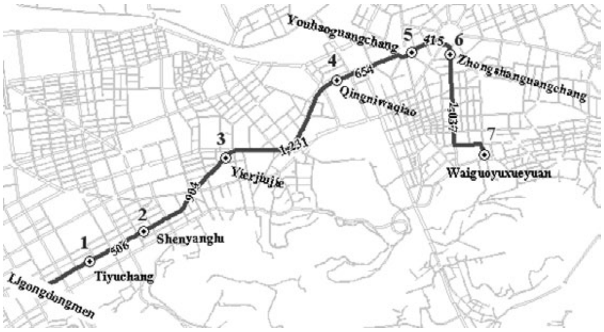

Fig. 7. The configuration of transit route No 23

To gain test data, we have collected arrival times at 7 stops on each individual trip at peak time (7:00 A.M. – 8:00 A.M.) and off-peak time (10:00 A.M. – 11:00 A.M.) during weekdays from September to October, 2006. The total amount of data acquired at peak and off-peak time is 3360. Then, speed on each segment can be computed based on arrival times at the stops.

The basic strategy used for the training process consisted of utilizing three data subsets, including training, cross-validation and testing. First, about 10% (330) of the samples of the obtained data were set aside as testing data. The remaining data are randomly assigned into two groups, the first one makes about 70% (2350) of the samples for training and the second one goes with data containing 680 samples for cross validation.

C is the regularization constant that reflects the importance of empirical risk. ε is the tube size that tunes approximation accuracy on training data points. The values of *C* and ε are determined by the users. The value of γ will affect the performance of the kernel function.

In this paper, the parameters are calibrated by grid-search (Yu *et al.* 2006; Dong *et al.* 2005) where all combinations (*C*, ε and γ) are tried and, finally, the one with the best performance is chosen (2^{-2}, 2^{-5}, 1.58). Furthermore, in single stop prediction and average speed prediction models, the number of the considered proceeding vehicles is set to three by simulation.

To examine the performance of the improved iterative prediction model for multiple stop arrival times (IMATP), we have implemented work compared with a history mean prediction model (HMP), a single stop arrival time prediction model (OATP) and an iterative multiple stop arrival time prediction model (MATP). In the HMP model, the speeds between two stops are directly computed by test data, and therefore arrival times at the stops can be obtained. The structure of the OATP is as shown in Fig. 3 and arrival times at downstream stops are recursively estimated. The difference between the IMATP model and the MATP model is that no dynamic algorithm in the MATP model exists. In order to have the same basis of comparison, the same training and verification sets are used for four models. Prediction accuracy is evaluated by computing the mean absolute percentage error (MAPE) that can be obtained from:

$$\mathrm{MAPE} = \frac{1}{J}\sum_{j=1}^{J}\frac{|T_{k+j} - \hat{T}_{k+j}|}{T_{k+j}} \times 100\%, \tag{13}$$

where: J is the number of test samples; T_{k+j} is the observed arrival time at stop $k + j$.

Fig. 8 shows a comparison of the MAPEs of the four models at all stops. Interestingly, the MAPEs of the four approaches in the peak period are less than the ones in the off-peak period. It can be attributed that the test bed locates in the centre of Dalian city where traffic at peak time slows to a near standstill. This reduces variability in travel times and decreases prediction errors at the peak period. However, there is no heavy traffic congestion at the off-peak period while traffic volume in the area is still large. This brings some uncertainty of predicting and decreasing prediction accuracy. Furthermore, one can observe that predicting Stop 4 using all four approaches have acutest changes in the peak or off-peak period. It can be attributed that the stop locates on the central business district (CBD) of Dalian city. Traffic conditions around Stop 4 are always complex. This can decrease prediction accuracy.

In addition, the MAPEs of each approach at the peak or off-peak period show a trend towards an increase along the bus route. This can be attributed to the propagation of preceding prediction errors since each prediction is based on the prediction output of the preceding stop. As a simple statistical approach, the HMP model only uses the average speeds from test data to predict arrival times at multiple downstream stops and yields the largest MAPE among the four ap-

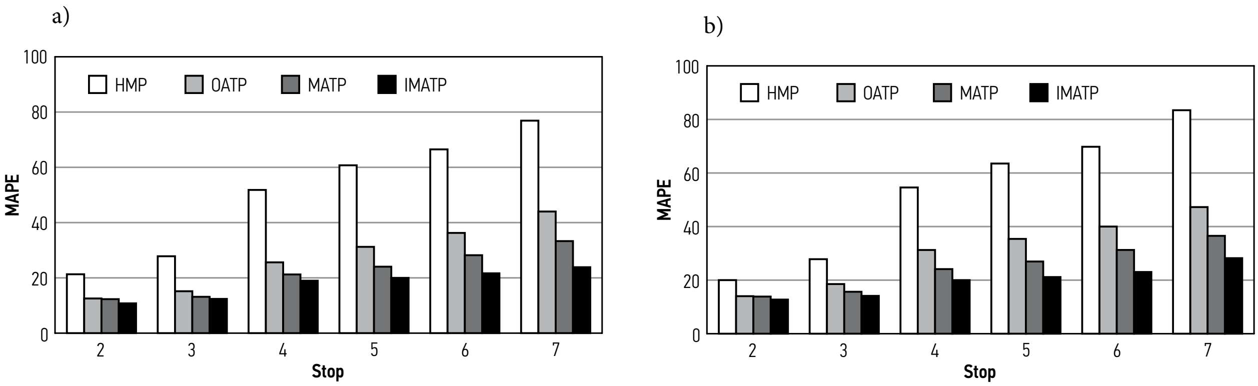

Fig. 8. A comparison of the MAPEs of four methods: a – peak period; b – off-peak period

proaches. The performance of the OATP model is good and slightly inferior to the one of the IMATP model forecasting arrival time at the immediate downstream stop. However, since the OATP model cannot consider traffic conditions on further segments after the immediate downstream stop, MAPEs predicting arrival times at further downstream stops are greatly increased. The MATP can provide better performance than HMP and OATP models. This indicates that the average bus speed prediction model on a further segment is effective. Furthermore, according to test results, adding a dynamic algorithm to MATP can also improve the performance of the model. Meanwhile, using IMATP can yield better solutions than the other stated approaches. So the IMATP model seems to be a powerful tool for predicting multiple stop arrival time.

4. Conclusions

Along with the development of automatic vehicle location or identification systems and automatic passenger counters, there is a growing interest in providing real-time bus arrival information, especially multi-stop prediction. However, the above presented study on predicting arrival time at multiple downstream stops is not receiving increasing attentions to transportation management.

The purpose of this paper is to develop a prediction method for bus arrival times at multiple downstream stops. First, an iterative prediction model that consists of a single stop prediction model for arrival time at the immediate downstream stop and an average bus speed prediction model on a further segment is presented. Then, the two prediction models are used to recursively forecast arrival times at multiple downstream stops. Furthermore, in order to enhance prediction accuracy, a dynamic algorithm based on the Kalman filter is developed. To evaluate the performance of the proposed approaches, transit data obtained from Dalian city are used for testing information. The received results show that the improved iterative prediction model for multiple stop arrival times outperforms the other three methods.

The compatibility of the method with real-world data is essentially an important aspect of research. Moreover, SVM-based models would be retrained at regular schedules to increase its reliability in performance when new data were obtained. In addition, these methods have important implications for practitioners and model developers. Practitioners could use the accurate prediction results of multi-stop arrival time to timely improve transit operation and provide passengers with convenient and good service. Model developers could validate the performance of the methods in the transit systems of other cities or fields. The method can also be explored in other applications such as predicting road traffic on speed/condition/time, predicting financial time series, exponents, etc. To improve the performance of the proposed prediction models, further studies should consider integrating more factors such as weather and incident.

Acknowledgments

This work was supported in Natural Science Foundation of Jiangsu in China (BK2011745) and National Natural Science Foundation of China (51108053).

References

Cathey, F. W.; Dailey, D. J. 2003. A prescription for transit arrival/departure prediction using automatic vehicle location data, *Transportation Research Part C: Emerging Technologies* 11(3–4): 241–264.
http://dx.doi.org/10.1016/S0968-090X(03)00023-8

Cao, L. J.; Tay, F. E. H. 2003. Support vector machine with adaptive parameters in financial time series forecasting, *IEEE Transactions on Neural Networks* 14(6): 1506–1518.
http://dx.doi.org/10.1109/TNN.2003.820556

Chen, M.; Liu, X.-B.; Xia, J.-X.; Chien, S. I. 2004. A dynamic bus-arrival time prediction model based on APC data, *Computer-Aided Civil and Infrastructure Engineering* 19(5): 364–376.
http://dx.doi.org/10.1111/j.1467-8667.2004.00363.x

Chien, S. I.-J.; Daripally, S. K.; Kim, K. 2007. Development of a probabilistic model to optimize disseminated real-time bus arrival information for pre-trip passengers, *Journal of Advanced Transportation* 41(2): 195–215.
http://dx.doi.org/10.1002/atr.5670410205

Chien, S. I.-J.; Ding, Y.; Wei, C. 2002. Dynamic bus arrival time prediction with artificial neural networks, *Journal of Transportation Engineering* 128(5): 429–438.
http://dx.doi.org/10.1061/(ASCE)0733-947X(2002)128:5(429)

Dailey, D. J.; Wall, Z. R.; Maclean, S. D.; Cathey, F. W. 2000. An algorithm and implementation to predict the arrival of transit vehicles, in *2000 IEEE Intelligent Transportation Systems. Proceedings*, 161–166.
http://dx.doi.org/10.1109/ITSC.2000.881043

Dong, B.; Cao, C.; Lee, S. E. 2005. Applying support vector machines to predict building energy consumption in tropical region, *Energy and Buildings* 37(5): 545–553.
http://dx.doi.org/10.1016/j.enbuild.2004.09.009

Grewal, M. S.; Andrews, A. P. 2001. *Kalman Filtering: Theory and Practice Using MALTAB.* 2nd edition. Wiley-Interscience. 416 p.

Hellinga, B. R.; Fu, L. 2002. Reducing bias in probe-based arterial link travel time estimates, *Transportation Research Part C: Emerging Technologies* 10(4): 257–273.
http://dx.doi.org/10.1016/S0968-090X(02)00003-7

Li, Y.; McDonald, M. 2002. Link travel time estimation using single GPS equipped probe vehicle, in *The IEEE 5th International Conference on Intelligent Transportation Systems, 2002. Proceedings*, 932–937.
http://dx.doi.org/10.1109/ITSC.2002.1041345

Ren, J.-T.; Ou, X.-L.; Zhang, Y.; Hu, D.-C. 2002. Research on network-level traffic pattern recognition, in *The IEEE 5th International Conference on Intelligent Transportation Systems, 2002. Proceedings*, 500–504.
http://dx.doi.org/10.1109/ITSC.2002.1041268

Vanajakshi, L.; Subramanian, S. C.; Sivanandan, R. 2009. Travel time prediction under heterogeneous traffic conditions using global positioning system data from buses, *IET Intelligent Transport Systems* 3(1): 1–9.
http://dx.doi.org/10.1049/iet-its:20080013

Vapnik, V. N. 1999a. An overview of statistical learning theory, *IEEE Transactions on Neural Networks* 10(5): 988–999.
http://dx.doi.org/10.1109/72.788640

Vapnik, V. N. 1999b. *The Nature of Statistical Learning Theory*. 2nd edition. Springer. 333 p.

Wu, C.-H.; Ho, J.-M.; Lee, D.-T. 2004. Travel-time prediction with support vector regression, *IEEE Transactions on Intelligent Transportation Systems* 5(4): 276–281. http://dx.doi.org/10.1109/TITS.2004.837813

Yu, B.; Lam, W. H. K.; Tam, M. L. 2011. Bus arrival time prediction at bus stop with multiple routes, *Transportation Research Part C: Emerging Technologies* 19(6): 1157–1170. http://dx.doi.org/10.1016/j.trc.2011.01.003

Yu, B.; Yang, Z.-Z.; Jin, P.-H.; Wu, S.-H.; Yao, B.-Z. 2012. Transit route network design-maximizing direct and transfer demand density, *Transportation Research Part C: Emerging Technologies* 22: 58–75. http://dx.doi.org/10.1016/j.trc.2011.12.003

Yu, B.; Yang, Z.-Z.; Chen, K.; Yu, B. 2010. Hybrid model for prediction of bus arrival times at next station, *Journal of Advanced Transportation* 44(3): 193–204. http://dx.doi.org/10.1002/atr.136

Yu, B.; Yang, Z.-Z.; Yao, B.-Z. 2006. Bus arrival time prediction using support vector machines, *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations* 10(4): 151–158. http://dx.doi.org/10.1080/15472450600981009

Yuan, F.; Cheu, R. L. 2003. Incident detection using support vector machines, *Transportation Research Part C: Emerging Technologies* 11(3–4): 309–328. http://dx.doi.org/10.1016/S0968-090X(03)00020-2