BUS ARRIVAL TIME PREDICTION USING SUPPORT VECTOR MACHINE WITH GENETIC ALGORITHM

M. Yang, *C. Chen,* *L. Wang,* *X. Yan,* *L. Zhou*

Abstract: Accurate prediction of bus arrival time is of great significance to improve passenger satisfaction and bus attraction. This paper presents the prediction model of bus arrival time based on Support Vector Machine with genetic algorithm (GA-SVM). The character of the time period, the length of road, the weather, the bus speed and the rate of road usage are adopted as input vectors in Support Vector Machine (SVM), and the genetic algorithm search algorithm is combined to find the best parameters. Finally, the data from Bus No. 249 in Shenyang, china are used to check the model. The experimental results show that the forecasting model is superior to the traditional SVM model and the Artificial Neural Network (ANN) model in terms of the same data, and is of higher accuracy, which verified the feasibility of the model to predict the bus arrival time.

1.   Introduction

Bus arrival time prediction is the important component to realize the information technology of urban public transport system, and can positively promote the development of urban public transport system. From the view of passengers, the timely information of bus arrival time can not only reduce the waiting time of passengers, but also make them reasonably arrange their travel plans, and select the most convenient way of travelling and interchange. From the perspective of public transport operators, it will greatly enrich service content, be conducive to timely adjusting the bus departure interval and the running time especially under

---

[*]Ming Yang, Nanjing Institute of City Transportation Planning, Transport Mansion, Nanjing, 210008, China

[†]Chao Chen, School of Automotive Engineering, Dalian University of Technology, Dalian, 116024, China

[‡]Lu Wang, China Academy of Civil Aviation Science and Technology, Beijing, 100028, China

[§]Xinxin Yan – Corresponding author, College of Management and Economics, Tianjin University, Tianjin, 300072, China, E-mail: `yanxx_tj@163.com`

[¶]Liping Zhou, Wuxi Mingda Traffic & Technology Consulted Co.,Ltd. Wuxi City , 214125, China

emergency, establishing a good image of the efficient operation, and enhancing the attraction of the public transit.

As the key information technology of public transport system, the prediction of bus arrival time attracts the great attention of scholars both at home and abroad. Patuaik et al. use a multivariate regression prediction model to forecast the bus arrival time, through the vehicle automatic passenger counting system which gets the boarding passenger number and the delayed time, and combined with the factors of station distance, detention time, station number [11]. Bae and Kachroo simulate the dynamic characteristics of a bus driver based on passenger arrival rate and the proportion of boarding passengers, and then use the least squares estimation method to predict the passenger time of getting on a bus and the parameters of departure interval time, and then use the parameter adaptive algorithm to establish the forecast model of bus arrival time [1]. Cathey and Dailiey design the prediction model based on Kalman filter by automatic vehicle location system, and test the model through the actual data of bus lines in Portland and Seattle. Results show that the prediction model is better than the timetable, but line should be equipped with the vehicle positioning system [4]. Bie et al. present the forecast model based on the GPS data to predict the bus arrived time of signal intersection [2]. Chen et al. develop a prediction model of neural network considering the effect of weather and time period to predict the bus arrival time [5]. Yu et al. propose the prediction model of bus arrival time based on SVM and forgetting factor. The real time of bus arriving at each time point is predicted, by taking the time, weather and links from the historical data as input vectors. A k-Nearest Neighbor Model for Multiple-Time-Step Prediction is introduced to predict Short-Term Traffic Condition. And the Grubbs' test method is applied to remove outliers from the input data [26]. Wang et al. propose an approach including two phases. Firstly, Radial Basis Function Neural Networks (RBFNN) model is used to learn and approximate the nonlinear relationship in historical data. Then, an online oriented method is introduced to adjust to the actual situation [18]. Other bus arrival time predictions can be found in these literatures [6, 25, 27, 12, 14, 15]. It can be seen that most scholars analyze and forecast the arrival time based on historical and real-time data, and the model been put forward mainly include the SVM prediction model, Kalman filter model, Artificial Neural Network model, nonparametric regression model, time series model, and so on.

Differed from the traditional research methods, this paper presents the prediction model of bus arrival time based on the GA-SVM method. Support Vector Machine method has obvious advantages mainly in solving the problem of small sample, nonlinear, multivariable classification and regression problems, and is of good generalization ability [17]. However, the parameters in the SVM govern the training process and have a profound effect on the performance of the SVM [22]. GA is well-known as an interesting and widely used variable selection method [10]. Therefore, the genetic algorithm is adopted to optimize the learning parameters of SVM, and a better prediction model based on GA-SVM to predict bus arrival time is constructed. The model can not only predict the bus arrival time dynamically, but also have the advantages of less calculation and higher precision of prediction.

2. The prediction model of bus arrival time based on GA-SVM

2.1 Selection of impact factor set and predictive index

The prediction of bus arrival time should be of reliability, real-time and high precision. Usually the speed of the bus on the links is affected by many random factors, and to accurately predict the vehicle running time is very difficult. Therefore, the appropriate input variables are to be select. Considering the large difference of travel speed in sunny and rainy day, mooring and evening peak and off peak, the time period and weather are consider as input vector. In addition, the rate of road usage (the road flow/the road capacity), which represents the congestion condition of roads, may influence the travel speed and should be regard as one of the input vectors. Furthermore, there are two mainly vectors, that is, bus speed and length of road should be not leave out. Therefore, in this paper, the character of time period, length of road, weather, bus speed and the rate of road usage are adopted as input vectors. By the five input vectors, the time arriving at the link end point t_m is to be predicted in SVM. The time period and the length of road represent the static characteristics of a specific operation, and identify the vehicle's specific pattern. The weather, bus speed, and rate of road usage stand for the dynamic information of the vehicle, which will change according to the link, traffic, passenger flow and so on. After the structure of SVM determined, the SVM also need to be trained to look for the support vector to predict. The SVM trains with database updating, then the trained SVM method is applied to predict the running time of the bus on the links. With the operation of the vehicle, bus time of arrival in the follow-up time is constantly revised, until the end of the vehicle running. The process of arrival time prediction model is as shown in Fig. 1.

2.2 Support Vector Machine

Support Vector Machine is a learning method in small sample situation proposed by Vapnik [16]. It is a kind of learning theory based on statistical, realizing the principle of structural risk minimization. It shows high generalization ability by using a set of high dimensional linear functions, and thus it can capture reliability data patterns more easily than other models [19, 21]. The algorithm of SVM can be transferred into a quadratic optimization problem with linear constraints (a convex optimization problem), which can gain the global optimal solution, and falling into local optimal solution of the problem can be avoided. For this study, the SVM is employed. The principles of the SVM are briefly described below. The details can be found in these literatures [3, 13].

Assume the training input is defined as vectors $x(i) \in R^{In}$ for $i = 1, \ldots, N$, which are independent and identically distributed data with sample size N. The training output is defined as $y(i) \in R^1$ for $i = 1, \ldots, N$. The SVM maps $x(i)$ into a feature space $R^h (h > In)$ with higher dimension using a function $\Phi(x(i))$ to linearize the nonlinear relationship between $x(i)$ and $y(i)$. The estimation function of $y(i)$ is

$$\hat{y} = f(x) = w^T \Phi(x) + b, \tag{1}$$

where $w \in R^h$ and $b \in R^1$ are coefficients which are derived by solving the following optimization problem:

$$\min Z\left(w, \varepsilon, \xi_i, \xi_i^*\right) = \frac{1}{2}w^T w + C\left\{v\varepsilon + \frac{1}{N}\sum_{i=1}^{N}\left(\xi_i + \xi_i^*\right)\right\}. \tag{2}$$

Subject to

$$w^T \Phi\left(x\left(i\right)\right) + b - y\left(i\right) \leq \varepsilon + \xi_i \forall i = 1, \ldots, N, \tag{3}$$

$$y\left(i\right) - w^T \Phi\left(x\left(i\right)\right) - b \leq \varepsilon + \xi_i^* \forall i = 1, \ldots, N, \tag{4}$$

$$\xi_i, \xi_i^* \geq 0 \forall i = 1, \ldots, N, \tag{5}$$

$$\varepsilon \geq 0. \tag{6}$$

The first item in Eq. (2) aims at the flat of fitting function to improve the generalization performance, and the second item represent the prediction error, aiming at improving the accuracy. The whole Eq. (2) represents the minimization of structural risk, where ξ_i, ξ_i^* are slack variables; C is a regularization parameter, and v is a second parameter. For each $x(i)$ the allowable error is ε. Slack variables ξ_i, ξ_i^* capture errors above ε and are penalized in the objective function via a regularization constant C.

The RBF kernel maps samples into a higher dimensional space and, can handle the case when the relation between class labels and attributes is nonlinear. Furthermore, the linear kernel is a special case of RBF as Keerthi and Lin showed: the
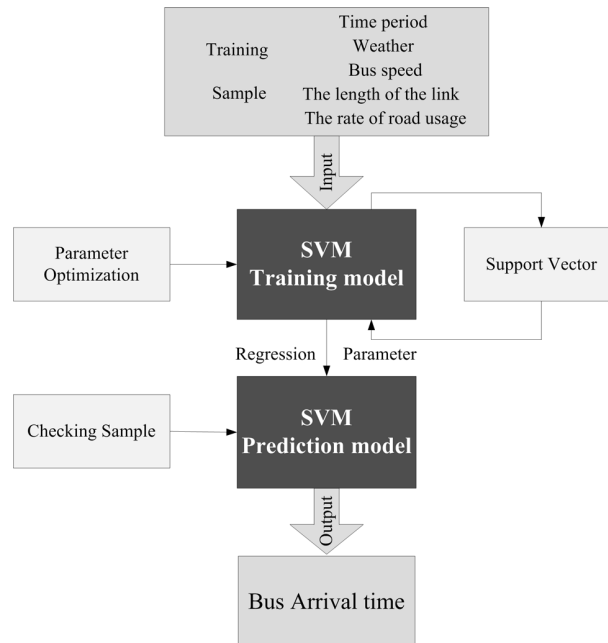

Fig. 1 *Predictable process of SVM.*

linear kernel with a penalty parameter C had the same performance as the RBF kernel with some parameters (C, ε, γ) [7]. In addition, the number of hyperparameters of the RBF kernel is less than the polynomial kernel, which means the RBF kernel has less numerical difficulties in contrast to polynomial kernels whose values may go to infinity or zero. Moreover, it is noted that the sigmoid kernel is not valid under some parameters [16]. Therefore, RBF kernel is selected in this study. Combining with RBF kernel, Eq. (1) can be obtained as the following estimated function of $y(i)$:

$$\hat{y} = f(x) = \sum_{i=1}^{N} (\alpha_i^* - \alpha_i) \times e^{-\gamma \|x(i) - x(j)\|^2} + b, \qquad (7)$$

where α_i and α_i^* are Lagrange multipliers for the constraints in Eqs. (3) and (4) respectively. γ is a parameter. With the radial basis function as the kernel function, the SVM has three parameters (C, ε, γ) that need to be determined. There is always a globally optimal solution to w and b with the input of three parameters (C, ε, γ).

3. Parameter selection of SVM based on genetic algorithm

The prediction accuracy of Support Vector Machines is largely affected by the parameter selection. Traditional parameters are selected according to the experience of trial and error, but the time and optimization results cannot be guaranteed. To increase the speed and optimality of the parameter selection, it needs to adopt a suitable intelligent optimization algorithm to search the optimal combination of the various parameters. GA is a randomized search method evolved from the laws of evolution (survival of the fittest) of the biosphere. The operation of GA is simple, finding the optimal parameter combination quickly. It is applicable to solve the optimization problem [8, 9].

3.1 Coding

In this paper, a multi-dimension code format based on real number is introduced [20]. Each gene value of the individual is in a range of a floating point number. The scope of penalty parameter C is [0–1000], the scope of loss function parameter ε is [0-1], and the scope of kernel parameter γ is [0–50]. The coding length of individuals is equal to the number of decision variables. In this paper, each coding is on behalf of one parameter as the Fig. 2 shows. The values of C, ε, γ generate randomly within the scope of a floating point number.

Fig. 2 *Example of coding.*

3.2 Genetic operations

3.2.1 Selection operation

Selection is the operation to select superior individuals and phase out the inferior individuals from the population. The purpose of selection method is to give opportunity of rapid growth to excellent individual. Roulette selection method is used in this paper. In this method, the selection probability of each individual is proportional to its fitness value. This plays an important role in the process of convergence of the algorithm [23, 24]. Elite strategy is also cooperated with it. So the better individuals in each generation do not participate in the crossover and mutation operation, being saved directly to the next generation. Assume that population size is M and the fitness of the individual i is F_i. The probability of individual to be chosen is P_i, as the following equation shows:

$$P_i = \frac{F_i}{\sum_{i=1}^{M} F_i} \qquad (i = 1, 2, \ldots, M). \tag{8}$$

3.2.2 Crossover operation

Crossover operator imitates genetic recombination process. It refers to the operation which exchanges genetic information. The parents interchange portions of strings to generate offspring. An arithmetic crossover is showed as the Eq. (9):

$$\begin{cases} q_{ix}^{(t)} = q_{ix}^{(t-1)} + (1 - r_i'')q_{iy}^{(t-1)} \\ q_{iy}^{(t)} = q_{iy}^{(t-1)} + (1 - r_i'')q_{ix}^{(t-1)} & r_i > 0.5 \\ q_{ix}^{(t)} = q_{ix}^{(t-1)} \\ q_{iy}^{(t)} = q_{iy}^{(t-1)} & \text{otherwise} \end{cases}, \tag{9}$$

where $q_{ix}^{(t-1)}$, $q_{iy}^{(t-1)}$ represent value of gene i of two parent, x and y, in generation t-1; r_i, r_i'' are the random numbers between (0-1). If $r_i > 0.5$, two parent chromosomes will cross to generate the child chromosomes; otherwise, the value of gene i of child chromosomes is the same as parent chromosomes.

3.3 Mutation operation

The mutation operator means to change the gene locus value of some individual string in the groups. In general, it is first to determine individuals which should mutate. Then randomly select gene locus to conduct mutation of the selected individuals. The non-uniform mutation operation is used in this paper. The parent chromosome is represented by $\left(q_1^{(t-1)}, q_2^{(t-1)}, \ldots, q_m^{(t-1)}\right)$, in which $q_k^{(t-1)}$ is the candidate gene locus that needs to be mutated. The gene value of offspring is shown in Eq. (10):

$$q_k^t = \begin{cases} q_k^{(t-1)} + \Delta(t, q_{k,\max} - q_k^{(t-1)}) & r_i > 0.5 \\ q_k^{(t-1)} - \Delta(t, q_k^{(t-1)} - q_{k,\min}) & \text{otherwise} \end{cases}. \tag{10}$$

The function $\Delta(t, z)$ can generate a value between $[0, z]$. The Eq. (11) is shown below:

$$\Delta(t, z) = z \times \left(1 - r'''^{(1 - \frac{t}{T_{\max}})^\alpha}\right), \tag{11}$$

where r''' is a random number between [0,1]; $T_{\max}$ is the maximum generation, and t is the current generation. $\alpha(2 \rightarrow 5)$ is a parameter determining the degree of dependency on the number of iterations. In early iterations, the range of no-uniform search is big, but with the increase of the number of iterations, it gets smaller and smaller.

3.4 The process of algorithm

Algorithm 1 GA optimizes the parameters of SVM as the following steps.

Set the initial parameters of GA, such as population size and number of iterations, set num = 1;

Determine the encoding interval of C, ε, γ. Real number coding is chosen to generate the chromosomes,

Mean square error (MSE) is chosen as the fitness function

repeat

Roulette selection is used cooperating with elite strategy.

crossover and mutation operators are used to create a child population

set num = num +1

if fitness agrees **then**

       Output the best individual and optimal solution.

  else

       run the operators of selection, crossover and mutation

end if

until the stopping criterion is met

---

4. Case study

4.1 Data

Based on the actual data from the bus No. 249 in Shenyang, Liaoning Province, China, the model is checked. Bus No. 249 is from Vanke Wonderland Garden to Zhongxing St. & shayang Rd., with the length of 15.2 km, via 27 stations. The single-way time of the line is 59 minutes. There are 10 control points, that is, 9 roads are set up. The road segment from Vanke Wonderland Garden to Zhongxing St. & shayang Rd. are denoted as road segment 1, road segment 2, ..., road segment 9, respectively. On weekdays, the bus arrival time at the 9 main points is predicted. The line information is shown in Fig. 3.

In data collection, a database from Shenyang Bus Group on each link at working day is established, and the information about the length can be seen in Tab. I, And the other variable data, that is, the time period, weather, the rate of road usage, bus speed were gained by Investigation.

Fig. 3 *The line information of bus No. 249.*

Route segment	1	2	3	4	5	6	7	8	9
The length of link [km]	1.72	2.61	1.42	2.07	1.16	1.28	1.33	1.43	0.96

Tab. I *The basic data of Shenyang Bus 249.*

4.1.1 Data normalization

As the data units are different and the difference of magnitude orders are big, the data for each variable has to be normalized. If the model calculates with raw data directly, data submerged is likely to happen. After normalizing, the data will fit well, which improves the precision of prediction. The normalization is accomplished using the following equation:

$$M_i^l = \frac{M_i}{\|M\|_2} = \frac{M_i}{\sqrt{M_1^2 + M_2^2 + \cdots + M_i^2}}. \tag{12}$$

4.2 The training of the model

In the process of the model training, the data samples are divided into three parts: training samples (70%), validation samples (20%) and test samples (10%), and the parameters of genetic algorithm are set as follows. The population scale is 20, the iteration number is 200, the crossover rate is 0.7, and mutation probability is 0.03. Using genetic algorithm, the best parameters $C = 645, v = 0.56, \gamma = 7.82$ are gained. Then these parameters are applied into the SVM model. The result

analysis is evaluated with root mean square error (RMSE) and the expression of RMSE is as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(T_{R,i} - T_{M,i})^2}{n}}, \tag{13}$$

where $T_{R,i}$ is the actual bus arrival time, $T_{M,i}$ stands for the predicted arrival time by the model, n is the number of sample data.

The training diagram of GA-SVM prediction model under sunny and peak (SP) condition can be seen in Fig. 4. The figure shows the average RMSE of 9 road segments is around 40 second, and all RMSEs of 9 road segments are below 70 second. Hence, it can be said that they fit well. The bus arrival time prediction model based on GA-SVM has strong identification ability.
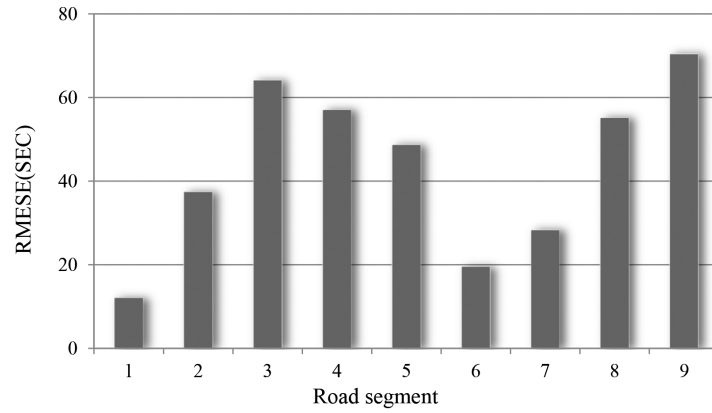
Fig. 4 *Training diagram of SVM prediction model under "SP".*

4.3   The prediction of the model

At the same time, this paper compares the prediction precision with the traditional Support Vector Machine method and Artificial Neural Network in terms of the same data situation. The traditional Support Vector Machine refers to the SVM model whose parameters are determined by experience of trial and error, without any optimized method. For the Artificial Neural Network model with the same input vector, a scaled conjugate gradient algorithm is used to train the ANN model. The number of hidden neurons is attained as five in this study. Thus, the final ANN model in this study is the ANN model with three-layer and five hidden neurons for travel speed prediction. Predicting results are as shown in Fig. 5, in which the contrast is done under four situations of sunny and peak (SP), sunny and off-peak (SO), raining and peak (RP) and raining and off-peak (RO). From Fig. 5 it can be seen that, the RMSE in rainy day is bigger than that in the sunny day, this is because that there are more uncertain factors in the rainy day, and the data collected in sunny day is more than that in rainy day, which can increase the predict accuracy. And the RMSE in peak period is bigger than that in off-peak
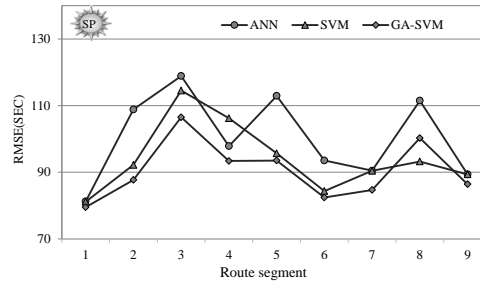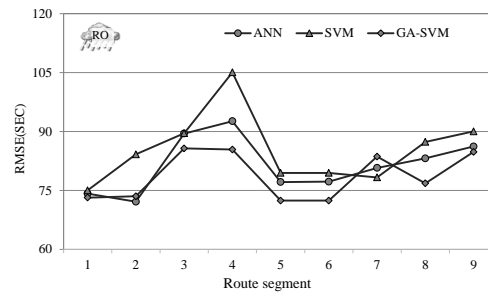
(a) Sunny and peak (SP).
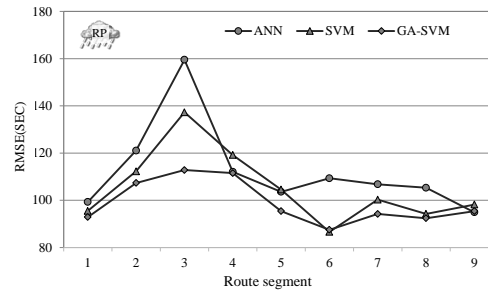
(b) Raining and peak (RP).
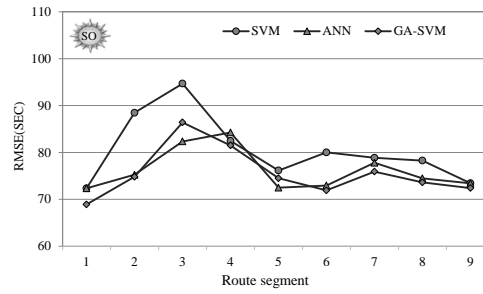
(c) Sunny and off-peak (SO).

(d) Raining and off-peak (RO).

Fig. 5 *The contrast illustration of three methods.*

period. It is because that in peak period the traffic condition is more complex, and the traffic congestion could cause delay. In terms of road segments, we can see the RMSE of road segment 6 is much bigger than others, this is because there are more traffic flows and mixed travels; finally, compared with ANN, the RMSE of GA-SVM is smaller. This can be attributed that the GA-SVM implement the structural risk minimization principle, while ANN models implement the empirical risk minimization principle. The solution of the GA-SVM may be globally optimal, while the ANN models may fall into a local optimal solution. While compared with traditional SVM, the performance of the GA-SVM has higher accuracy. The main reason is that GA optimizes the parameters of the Support Vector Machine. Hence, we can say that the GA-SVM model is superior to the forecasting method of neural network and the traditional SVM method, has the feasibility to predict the bus arrival time.

5. Conclusions

The SVM model has the advantages of strong learning ability in small sample situation, fast learning speed and good generalization ability and so on. The genetic algorithm is simple in program implementation, less in setting parameters and fast in calculating convergence speed. The prediction model of bus arrival time based on GA-SVM, which uses genetic algorithm to optimize the parameters of SVM, is the optimal SVM prediction model. The character of the time period, the length of road, the weather, the bus speed and the rate of road usage are adopt as input vectors to predict the bus arrival time, which is less used in other references. And the numerical tests were conducted under four cases: sunny and peak (SP), sunny and off-peak (SO), raining and peak (RP) and raining and off-peak (RO). The example analysis results show that the forecasting method based on GA-SVM model is superior to the forecasting method of neural network and the traditional SVM method in terms of the same data, and it overcomes the problem of "over learning" phenomenon in neural network training progress, avoids the local optimal solution, and has extremely good generalization ability. Therefore, the prediction model based on GA-SVM is better than that of general prediction model of bus arrival time and has better prediction accuracy.

References

[1] BAE S., KACHROO P. Proactive travel time predictions under interrupted flow condition. In: *Proceedings of the 6th International Vehicle Navigation and Information Systems Conference (VNIS), 1995*, Seattle, WA. A Ride Into the Future. In Conjunction with the Pacific Rim Transtech Conference. IEEE, 1995, pp. 179−186, doi: 10.1109/VNIS.1995.518836.

[2] BIE Y.M., WANG D.H., QI H.S. Prediction model of bus arrival time at signalized intersection using GPS data. *Journal of Transportation Engineering.* 2012, 138(1), pp. 12−20.

[3] BOSER B.E., GUYON I.M., VAPNIK V.N. A training algorithm for optimal margin classifiers. In: *Proceedings of the 5th annual workshop on Computational learning theory (COLT'92)*. ACM, 1992, pp. 144−152, doi: 10.1145/130385.130401.

[4] CATHEY F.W., DAILEY D.J. A prescription for transit arrival/departure prediction using automatic vehicle location data. *Transportation Research Part C: Emerging Technologies.* 2003, 11(3), pp. 241−264, doi: 10.1016/S0968−090X(03)00023−8.

[5] CHEN M., LIU X., XIA J., CHIEN S.I. A Dynamic BusArrival Time Prediction Model Based on APC Data. *Computer-Aided Civil and Infrastructure Engineering*. 2004, 19(5), pp. 364−376.

[6] JEONG R., RILETT L.R. Bus arrival time prediction using artificial neural network mode. In: *Proceedings of the 7th International IEEE Conference on Intelligent Transportation Systems, 2004*. IEEE, 2004, pp. 988.−993, doi: 10.1109/ITSC.2004.1399041.

[7] KEERTHI S.S., LIN C.J. Asymptotic behaviors of support vector machines with Gaussian kernel. *Neural computation*. 2003, 15(7), pp. 1667−1689, doi: 10.1162/089976603321891855.

[8] KIM D.S., NGUYEN H.N., PARK J.S. Genetic algorithm to improve SVM based network intrusion detection system. In: *Proceedings of the 19th International Conference on Advanced Information Networking and Applications, 2005 (AINA 2005)*. IEEE, 2005, 2, pp. 155−158, doi: 10.1109/AINA.2005.191.

[9] LIU H.B., JIAO Y.B. Application of genetic algorithm-support vector machine (GA-SVM) for damage identification of bridge. *International Journal of Computational Intelligence and Applications*. 2011, 10(4), pp. 383−397.

[10] PAI P.F. System reliability forecasting by support vector machines with genetic algorithms. *Mathematical and Computer Modelling*. 2006, 43(3-4), pp. 262−274, doi: 10.1016/j.mcm.2005.02.008.

[11] PATNAIK J., CHIEN S., BLADIKAS A. Estimation of Bus Arrival Times Using APC Data. *Journal of Public Transportation*. 2004, 7(1), pp. 1−20.

[12] RAJBHANDARI R. *Bus arrival time prediction using stochastic time series and Markov chains* [online]. New Jersey, 2005. PhD thesis, New Jersey Institute of Technology [viewed 2008-4-24]. Available from: http://www.library.northwestern.edu/libraries-collections/evanston-campus/transportation-library/services/document-delivery-ill-services.

[13] SCHÖLKOPF B., SMOLA A.J., WILLIAMSON R.C., BARLETT P.L. New support vector algorithms. *Neural computation*. 2000, 12(5), pp. 1207−1245.

[14] SHALABY A., FARHAN A. Bus travel time prediction model for dynamic operations control and passenger information systems [online]. Transportation Research Board 82nd Annual Meeting, Washington D.C., 2003 [viewed 2016-06-15]. Available from: http://www.civ.utoronto.ca/sect/traeng/its/downloads/civ1507_pubtran/lecture8_traveltimepred.pdf.

[15] SUN D., LUO H. FU L., LIU W., LIAO X., ZHAO M. Predicting bus arrival time on the basis of global positioning system data. *Transportation Research Record: Journal of the Transportation Research Board*. 2007, 2034(1), pp. 62−72, doi: 10.3141/2034-08.

[16] VAPNIK V.N. *The nature of statistical learning theory*. New York: Springer-Verlag, 1995.

[17] VAPNIK V.N. An overview of statistical learning theory. *IEEE Transactions on Neural Networks*. 1999, 10(5), pp. 988−999, doi: 10.1109/72.788640.

[18] WANG L., ZUO Z., FU J. Bus Arrival Time Prediction Using RBF Neural Networks Adjusted by Online Data. *Procedia-Social and Behavioral Sciences*. 2014, 138, pp. 67−75, doi: 10.1016/j.sbspro.2014.07.182.

[19] YAO B.Z., YAO J.B., ZHANG M.H., YU L. Improved support vector machine regression in multi-step-ahead prediction for rock displacement surrounding a tunnel. *Scientia iranica*. 2014, 21(4), pp. 1309−1316.

[20] YU B., KONG L., SUN Y., YAO B.Z., GAO Z.Y. A bi-level programming for bus lane network design. *Transportation Research Part C*. 2015, 55, pp. 310−327, doi: 10.1016/j.trc.2015.02.014.

[21] YAO B.Z., YANG C.Y., YAO J.B. Tunnel Surrounding Rock Displacement Prediction Using Support Vector Machine. *International Journal of Computational Intelligence Systems*. 2010, 3(6), pp. 843−852, doi: 10.1080/18756891.2010.9727746.

[22] YAO B.Z., HU P., ZHANG M.H., JIN M.Q. A Support Vector Machine with the Tabu Search Algorithm For Freeway Incident Detection. *International Journal of Applied Mathematics and Computer Science*. 2014, 24(2), pp. 397−404, doi: 10.2478/amcs-2014-0030.

[23] YAO B.Z., HU P., ZHANG M.H., WANG S. Artificial Bee Colony Algorithm with Scanning Strategy for Periodic Vehicle Routing Problem. *SIMULATION: Transactions of the Society for Modeling and Simulation International.* 2013, 89(6), pp. 762−770, doi: 10.1177/0037549713481503.

[24] YAO B.Z., YU B., GAO J.J., ZHANG M.H. An improved particle swarm optimization for carton heterogeneous vehicle routing problem with a collection depot. *Annals of Operations Research.* 2015, pp. 1−18, doi: 10.1007/s10479-015-1792-x.

[25] YU B., LAM W., TAM M. Bus arrival time prediction at bus stop with multiple routes. *Transportation Research Part C.* 2011, 19(6), pp. 1157−1170, doi: 10.1016/j.trc.2011.01.003.

[26] YU B., SONG X.L., YANG Z.M., YAO B.Z. k-Nearest Neighbor Model for Multiple-Time-Step Prediction of Short-Term Traffic Condition. *Journal of Transportation Engineering-ASCE.* 2015, 142(6), 04016018, doi: 10.1061/(ASCE)TE.1943-5436.0000816.

[27] YU B., YANG Z.Z., YAO B.Z. Bus arrival time prediction using support vector machines. *Journal of Intelligent Transportation Systems.* 2006, 10(4), pp. 151−158, doi: 10.1080/15472450600981009.