DE GRUYTER OPEN

Int. J. Appl. Math. Comput. Sci., 2014, Vol. 24, No. 2, 397–404
DOI: 10.2478/amcs-2014-0030

amcs

A SUPPORT VECTOR MACHINE WITH THE TABU SEARCH ALGORITHM FOR FREEWAY INCIDENT DETECTION

BAOZHEN YAO *, PING HU *, MINGHENG ZHANG *, MAOQING JIN **

* School of Automotive Engineering
Dalian University of Technology, Dalian 116024, PR China
e-mail: yaobaozhen@yahoo.cn

**High Technology Research and Development Center
Ministry of Science and Technology, Beijing, PR China
e-mail: Jin@htrdc.com

Automated Incident Detection (AID) is an important part of Advanced Traffic Management and Information Systems (ATMISs). An automated incident detection system can effectively provide information on an incident, which can help initiate the required measure to reduce the influence of the incident. To accurately detect incidents in expressways, a Support Vector Machine (SVM) is used in this paper. Since the selection of optimal parameters for the SVM can improve prediction accuracy, the tabu search algorithm is employed to optimize the SVM parameters. The proposed model is evaluated with data for two freeways in China. The results show that the tabu search algorithm can effectively provide better parameter values for the SVM, and SVM models outperform Artificial Neural Networks (ANNs) in freeway incident detection.

Keywords: automated incident detection, support vector machine, tabu search, freeway.

1. Introduction

In freeway traffic monitoring and control, the detection and verification of incidents is very important. When incidents happen, they can destroy the normal traffic flow and lead to a traffic jam. Rapid and reliable detection is an effective method to reduce the impacts of incidents and guarantee the safety of the freeway as much possible. Thus, the freeway incident detection problem is in freeway and arterial traffic management systems, and has been the subject of research in the past several decades. Automated Incident Detection (AID) is an important component in many Advanced Transportation Management and Information Systems (ATMISs). Using AID is attempted to detect freeway incidents such as accidents, stalled vehicles, spilled loads, temporary maintenance and construction activities. Many scholars have done a lot of work on freeway incident detection methods for several decades.

In the incident detection study, some technologies like time-series (Ahmed and Cook, 1982) and the decision tree for pattern recognition (Chen and Wang, 2009) have been used. In the literature, Ahmed and Cook (1982) presented time-series analysis techniques for automatic incidents detection. The model attempted to provide short-term forecasts of traffic occupancies based on the associated 95% confidence limits and an incident was detected if the observed occupancy value was out of the confidence limits. Chen and Wang (2009) applied a decision tree learning for freeway automatic incident detection. In this research, the traffic data of volume, speed, time headway and occupancy at both upstream and downstream detectors are used for incident detection.

Recently, neural networks have been widely applied in incident detection. Srinivasan *et al.* (2005) presented three neural network models for freeway incident detection, which consisted of a multi-layer feed-forward network, a basic probabilistic network and a constructive probabilistic network. Jin *et al.* (2002) attempted to propose a constructive probabilistic neural network for detecting freeway incidents, in which there is a mixture Gaussian model and a dynamic decay adjustment algorithm used to construct the model.

Artificial Neural Networks (ANNs) are a mathematical model or a computational model motivated

by emulating the intelligent data processing ability of human brains. A neural network consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation. That is, the synaptic weights can be adjusted in a learning process to reflect the input–output relationship for the analyzed system automatically (Hagan *et al.*, 1996; Wei and Wu 1997). ANNs appear to be a promising approach to describe complex systems due to their versatile parallel distributed structures and adaptive learning processes. However, it has been commonly reported that ANN models require a relatively large amount of training data to analyze the distribution of an input pattern, while the performance of the ANN can to be deteriorated for a small data pool. Moreover, it is difficult for it to generalize the results due to its overfitting nature.

The SVM is a very specific type of learning algorithm characterized by capacity control of the decision function, the use of kernel functions, and the sparse solution (Cristianini and Shawe-Taylor, 2000; Vapnik 1999; 2000; Jelen *et al.*, 2008; Mahmoud, 2011; Sumi *et al.*, 2012; Peter, 2013). Like ANNs, SVMs also depend on the similarity between historic and real-time traffic patterns. The SVMs can generally produce better prediction results due to the over-fitting avoidance ability and the high generalization performance. One disadvantage of SVMs is that it requires a large amount of computation time for a large training database (Cao *et al.*, 2003). However, it has provided some breakthroughs and plausible performances, such as traffic-pattern recognition (Ren *et al.*, 2002), head recognition (Reyna *et al.*, 2001), travel time prediction (Wu *et al.*, 2004; Yu *et al.*, 2006; 2010; 2012) and incident detection (Yuan and Cheu, 2003). These successful applications motivate us to apply SVMs for solving the incident detection problem.

The parameters in the SVM govern the training process and the values have a profound effect on the performance of the SVM. Therefore, there are many works on parameter determination for the SVMs. Lin *et al.* (2006) attempted to optimize appropriate parameters for an SVM prediction model by using the structural risk minimization principle. Zhang *et al.* (2010) developed Ant Colony Optimization (ACO) to select optimal parameters for the SVM. Lin *et al.* (2008) introduced particle swarm optimization to optimize the parameters in the SVM. Hou and Li (2009) attempted to determine the values for parameters in the SVM by using an evolution strategy with covariance matrix adaptation. Yao *et al.* (2010) tried to use a shuffled complex evolution algorithm to optimize the parameters for SVMs. Lorena *et al.* (2008) proposed genetic algorithms to optimize a set of parameter values for SVMs. Lebrun *et al.* (2008) proposed tabu search to build a selection of the hyperparameters for SVMs. The tabu search algorithm is a higher level heuristic algorithm for solving combinatorial optimization problems by allowing the search to explore solutions. These solutions do not decrease the objective function value only where these solutions are not forbidden. It has been successfully applied to solving some classic compounding optimization problems (Augugliaro *et al.*, 2002; Bortfeldt *et al.*, 2003; Falco *et al.*, 1994; Ho and Haugland, 2004; Talbi *et al.*, 1998; Yao *et al.*, 2013). The present paper attempts to find the appropriate parameters in SVMs by using the tabu search algorithm.

Most of the AID methods indirectly judge traffic event existence through identifying abnormal changes of traffic flow parameters. Induction coil detector AID system are most widely used to get traffic flow parameters due to the lowest cost. However, the induction coil detector also has its disadvantages. For example, incidents involve some time to get a nearest detector. Moreover, the induction coil detector is not suitable for the low traffic flow. So the traffic management department is not highly concerned about them. The incident detection system based on the induction coil usually uses parameters such as flow, market share, etc., and in some cases it also needs time to detect all kinds of accidents between the upstream and downstream speed parameters.

This paper presents a prediction model based on the SVM for freeway incident detection, and the tabu search algorithm is used for parameters optimization for the SVM. The structure of this paper is as follows. Section 2 provides a brief introduction to a prediction model of the SVM, and parameter optimization for the tabu search algorithm is presented. Section 3 contains a case study and some computational results; and lastly, the conclusions are provided in Section 4.

2. Support vector machine for the incident detection problem

2.1. Support vector machine for regression. The SVM is a non-probabilistic binary linear classifier, which can be adjusted to map the input–output relationship for a non-linear system. By applying a set of high dimensional linear functions, the SVM shows strong resistance to the over-fitting problem and high generalization performance.

Given the training data set

$$\{x_k, y_k\}, \quad k = 1, 2, \ldots, s,$$

$x_k \in \mathbb{R}^m$ is the input data and $y_k \in \mathbb{R}^n$ is the actual value. The SVM estimates the relationships between the input and the output by the following function:

$$f(x) = \langle w, x \rangle + b, \quad w, x \in \mathbb{R}^m, \quad b \in \mathbb{R}^n. \quad (1)$$

Here, $\langle w, x \rangle$ is the feature of the inputs. The coefficients w and b are estimated by the so-called regularized risk

functional:

$$\min J = \tfrac{1}{2}\|w\|^2 + C\sum_{i=1}^{s}(\xi_i^* + \xi_i)$$

subject to

$$\begin{cases} y_i - \langle w, x\rangle - b \le \varepsilon + \xi_i, \\ \langle w, x\rangle + b - y_i \le \varepsilon + \xi_i, \\ \xi_i^*, \xi_i \ge 0. \end{cases} \tag{2}$$

The first term, $\frac{1}{2}\|w\|^2$, is called the regularized term and is used as a measurement of function flatness. C is a regularization constant to determine the trade-off between the training error and the generalization performance. Two positive slack variables ξ, ξ^* are used to cope with infeasible constraints of the optimization problem. The parameter is used to reflect that the loss equals zero if the forecast value is within the "ε-tube".

This constrained optimization problem is solved using the following primal Lagrangian form:

$$\begin{aligned} L =& \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{s}(\xi_i^* + \xi_i) - \sum_{i=1}^{s}(\eta_i\xi_i + \eta_i^*\xi_i) \\ & - \sum_{i=1}^{s}\alpha_i(\varepsilon + \xi_i - y_i + \langle w, x_i\rangle + b) \\ & - \sum_{i=1}^{s}\alpha_i^*(\varepsilon + \xi^* - y_i + \langle w, x_i\rangle + b). \end{aligned} \tag{3}$$

Here, L is the Lagrangian and η_i,η_i^*, α_i, α_i^* are Lagrange multipliers. Hence the dual variables in (3) have to satisfy the following nonnegativity constraints:

$$\eta_i, \eta_i^*, \alpha_i, \alpha_i^* \ge 0. \tag{4}$$

The above problem can be converted into a dual one where the task is to optimize the Lagrangian multipliers, α_i and α_i^*. The dual problem contains a quadratic objective function of α_i and α_i^* with one linear constraint:

$$\begin{aligned} \max J7 =& -\frac{1}{2}\sum_{i=1}^{s}(\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)\langle x_i, x_j\rangle \\ & + \sum_{i=1}^{s}\alpha_i^*(y_i - \varepsilon) - \sum_{i=1}^{s}\alpha_i^*(y_i + \varepsilon) \end{aligned} \tag{5}$$

subject to

$$\begin{cases} \sum_{i=1}^{s}\alpha_i = \sum_{i=1}^{s}\alpha_i^*, \\ 0 \le \alpha_i \le C, \\ 0 \le \alpha_i^* \le C. \end{cases}$$

Let

$$\omega = \sum_{i=1}^{s}(\alpha_i - \alpha_i^*)x_i. \tag{6}$$

Thus

$$f(x) = \sum_{i=1}^{s}(\alpha_i - \alpha_i^*)\langle x_i, x_j\rangle + b. \tag{7}$$

By introducing kernel function $K(x_ix_j)$, Eqn. (8) can be rewritten as follows:

$$f(x) = \sum_{i=1}^{s}(\alpha_i - \alpha_i^*)K(x_i, x_j) + b. \tag{8}$$

The kernel function is proven to simplify the use of a mapping. The value of $K(x_ix_j)$ is equal to the inner product of two vectors x_i and x_j in the feature space $\phi(x_i)$ and $\phi(x_j)$, that is, $K(x_i, x_j)$= $\phi(x_i)$ $\phi(x_j)$. By the use of kernels, all necessary computations can be performed directly in the input space, without having to compute the map $\phi(x)$. More details on SVMs are given Vapnik (1999) and Cao *et al.* (2003).

2.2. Tabu search algorithm for parameter optimization of SVMs. The kernel function is the core of the SVM which helps it to get an optimal solution. In general, the RBF kernel, as a nonlinear kernel function, is a reasonable first choice (Dong *et al.*, 2005). The parameters C, ε and σ are key elements of the RBF kernel and directly exert considerable influence on the generalization ability of the SVM. The parameter C controls the trade-off between the margin and the size of the slack variables (Shawe-Taylor and Cristianini, 2004). If the value of C is too large, the classification accuracy rate is very high in the training phase, but very low in the testing phase. Otherwise, if the value of C is too small, the classification accuracy rate is unsatisfactory. The parameter σ has an effect on the partitioning outcome in the feature space. If the value of σ is too large, it will lead to over-fitting. If the value of σ is too small, it will lead to under-fitting (Pardo and Sberveglieri, 2005). The parameter ε reflects the range of the prediction deviation. If the value of ε is too large, it will results in the solutions to be more independent of the existing data. If the value of ε is too small, it will result in overfitting. So parameter optimization is an important factor for improving the prediction accuracy of the SVM. The tabu search algorithm is applied to optimize the parameters in the SVM.

Initial solution. In this study, the determination of the initial solution in the tabu search algorithm is to optimize the parameters (C, ε and σ) for the current SVM model. Firstly, an initialization solution is presented by random. To reduce the search space referring to previous literature using the SVM (Yu *et al.*, 2006; 2011), it is recommended to introduce the constraints of the three parameters which respectively attribute to the range C $\in[2^{-5}, 2^5]$, $\varepsilon\in[2^{-5}, 2^5]$, and $\sigma\in[2^{-5}, 2^5]$. The solution is computed by appending the nearest unused

neighbor values of the three parameters with respect to the minimum sum of the Root Mean Squared Error (RMSE) which is adopted in this paper,

$$\text{RMSE} = \left[\frac{1}{n-p}\sum_{i=1}^{n}(y_i - \hat{y})\right]^{1/2}, \qquad (9)$$

where $\hat{y}$ is the prediction value for the model, y_i is the observed value n is the number of observations and p is the number of model parameters. The initial solution will be evaluated by appending the nearest unvisited neighbor with respect to the minimum sum of the negative root squared error. The process repeats until all the neighbors are visited.

Neighborhood. Like the nearest neighbor search, the neighborhood of our tabu search algorithm is also based on a 2-opt exchange. In this paper, the 2-opt operation is adopted to expand the search space by exchanging information between two solutions.

Tabu criteria. The tabu criteria, if applied unconditionally, sometimes reject worthwhile candidates. To avoid this situation, an aspiration criterion is used to override the tabu restriction. When a tabu neighbor has a cost lower than the lowest status, a tabu move is permitted. If all candidate solutions are part of the tabu list and they all fail to meet the criterion, the candidate with higher permanency time in the Tabu list is selected for the move.

Termination. In this paper, the search continues until $\text{RMSE}_n - \text{RMSE}_{n-1} < 0.0001$ or the number of generation reaches the maximum number of generations $T_{\max}$.

2.3. Applying the SVM for freeway incident detection. The input data are to be fed to the SVM to identify the relation between the input and the output. Thus, it is important to normalize the input data. In freeway incident detection, the traffic conditions like weather or time-of-day have more influence on the traffic flow. Furthermore, the upstream and the downstream, the upstream occupancy, upstream volume, downstream occupancy, downstream volume and segment can describe the state of traffic along the freeway. Therefore, weather, time-of-day, occupancy and volume of the upstream and downstream are selected as the input data which are easy to reflect the variation of the incidents. The number of the segment selected for prediction is assumed as t. The prediction window m is used to describe the influence of the m-th segment far from the current segment t (Fig. 1). Thus, the proposed SVM model is the structure shown in Fig. 2. The input vector (x) consists of weather (x^1), time-of-day (x^2), segment (x^3), the upstream occupancies (x^1) on the segments $t, t-1, t-2, \ldots, t-m$, the downstream occupancies on the segments $t+1, t+2, \ldots, t+m$, the upstream volumes on the segments $t, t-1, t-2, \ldots, t-m$, the downstream volumes on the segments $t+1, t+2, \ldots, t+m$. The outputs (y) of the models are the incident detection results of the target segment.

3. Case study

The SVM model for freeway incident detection is tested with the data of the Shenda freeway and the Heda freeway in the Liaoning province. The Shenda freeway leads from Shenyang city to Dalian and is the first freeway in China mainland. The total length is 400 km. The Heda freeway leads from Jixi to Mudanjiang and its total length is 1390 km. In the numerical test, one section of the Shenda freeway (from Dalian city to Pulandian city) and one section of the Heda freeway (from Dalian city to Zhuanghe city) are taken as the test bed. The lengths of the two sections are 88 km and 174 km, respectively. The routes of the two sections are shown in Fig 3.

3.1. Data collection and processing. To acquire the loop detector data at the two sections, we conducted an experiment on the two sections simultaneously, from April 16 to 20, 2012. There are double lanes in one direction, the vehicle speed is assumed as 60 km/h, and the vehicle flow is about from 1500 to 2000 vehicles/h. The interval between two test points is from 500 to 700 m. The loop detector data about the time-of-day, weather, segments, occupancy about all the segments

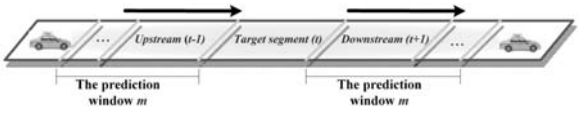

Fig. 1. Prediction configuration of the routes.

Fig. 3. Information on the two freeways.

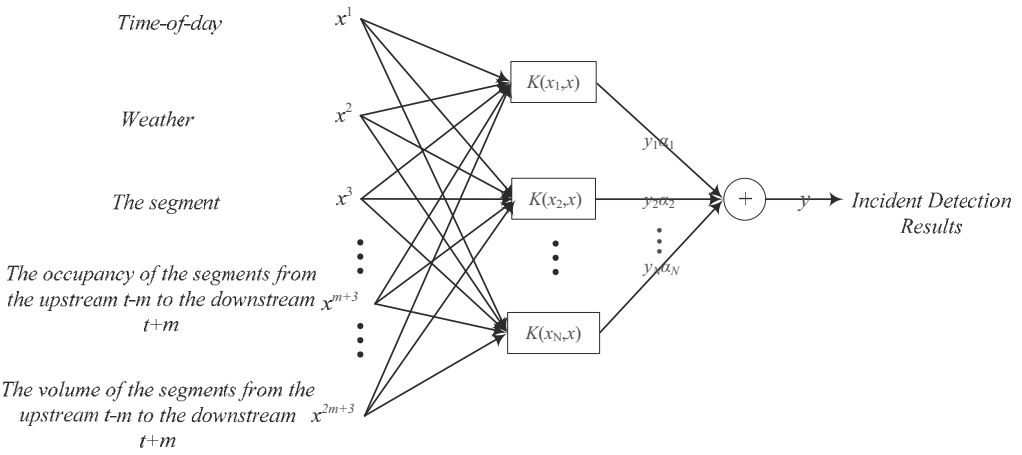

Fig. 2. Framework of the proposed SVM model.

and all the volumes from all the tested segments were collected at 30-second intervals. The data collected from all the lanes at a station were fed into the SVM models. There are 140 incidents and 164 incidents from the two sections in the whole process of the experience. Easy data are divided into three sub-sets, which represent training samples, testing samples and inspection samples, respectively. There are about 70% samples for training, 30% samples for testing and the remaining samples are for inspection.

3.2. Determination of the prediction window m. The prediction window m is selected to reflect the influence from the m-th segment far to the current one. The choice of the prediction window m is typically a compromise between the ability to track changes in the input data. A long horizon m is used when the learning is in the steady state and there is no obvious model variation, while a small one is used which will lead to large errors. Thus, a too large and a too small prediction window m will affect the prediction accuracy. To determine the value of the prediction window m based on the data of the two sections, the prediction errors from the SVM with different m are shown in Figs. 4 (a) and (b).

From Fig. 4, it can be found that the horizon m is 5–8, and the RMSE of the SVM model is almost the same. This indicates that lengthening the horizon m again cannot improve the prediction accuracy. Thus, the horizon $m = 5$ is selected in this paper.

3.3. Tabu search algorithm for parameter identification. In this paper, the parameters of the tabu search algorithm used for freeway incident detection are as follows. The length of tabu is 5. The tabu search algorithm was coded in Visual C++. Net 2003 and executed on a PC equipped with 512 MB of RAM and a Pentium processor running at 1000 MHz. The tabu search algorithm continues running 10 times under the same condition. The computation results for 10 times are shown in Fig. 5. From Fig. 5, it can be seen that the prediction error decreases fast before the 640-th generation, and then it changes smoothly. The least prediction error appears at about the 720-th generation, and then it remains almost unchanged. It can be also found that the calculation results for the ten times are almost equal. This suggests that the tabu search algorithm has good convergence and the results with the least prediction error correspond to the optimization for three parameters, that is, $C = 3.1042$, $\varepsilon = 0.0032$ and $\sigma = 1.1732$, for the freeway incident detection model.

3.4. Computational results. The performance of an AID model is often evaluated by the following three indices (Srinivasan *et al.*, 2005):

1. Detection Rate (DR), defined as the ratio of the number of incidents correctly detected to the total number of incidents known: $DR =$ (no. of incidents detected/total no. of incident cases) 100%.

2. False Alarm Rate (FAR), defined as the ratio of the number of false alarm cases to the total number of applications or decisions made by the algorithm. $FAR =$ (number of false alarms/total number of incident-free input patterns) 100%.

3. Mean Time-To-Detect (MTTD) is the average time an algorithm takes to detect incidents,

$$\text{MTTD} = \frac{1}{n}\sum_{i=1}^{n}(t_{id} - t_{io}), \qquad (10)$$

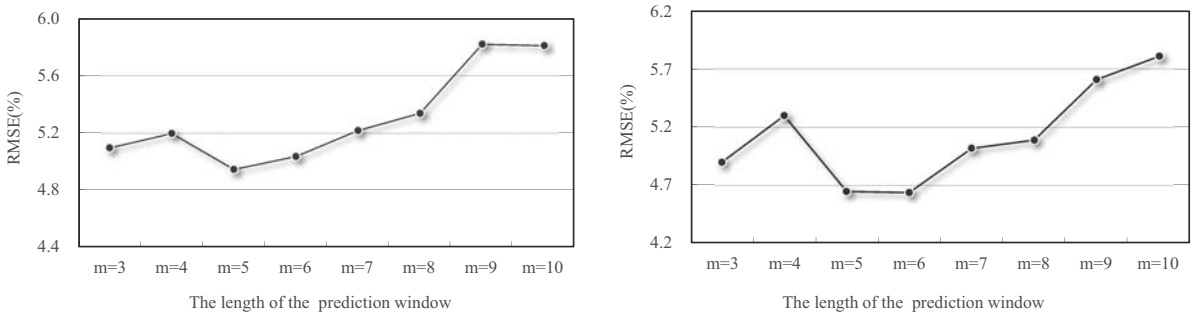

Fig. 4. Comparison of the prediction window m: section from Dalian to Pulandian (a), section from Dalian to Zhuanghe (b).

where t_{id} is the time when the incident was detected, t_{io} is the time when the incident occurred and n is the number of correctly detected incident cases.

4. MisClassification Rate (MCR) is the percentage of misclassified patterns out of the total number of input patterns. MCR = (number of wrongly classified input patterns/total number of input patterns) 100%.

To test the effect of the support vector machine, a standard three-layer artificial neural network is used for predicting the freeway incidents. Before applying ANN models, the input parameters of the ANN model need to be determined using an experiment as SVM identification. The combinations of the input parameters of ANN models are the same as for the SVM experiment. The test results of the ANN and SVM based on the two sections from two freeways are shown in Tables 1 and 2.

From Tables 2 and 3 it can be found that the detection speed of the neural network is slightly faster than that for the support vector machine, but the detection rate, the false positive rate and the misclassification rate of the SVM are superior to those the artificial neural network. A higher detection rate can effectively reduce negative efficiency which is brought by a transportation incident. A low false positive rate also saves the system operation cost. Thus, the results suggest that our SVM model is an effective method for freeway incident detection.

In addition, to examine the effectiveness of the tabu search algorithm proposed in this paper, the computational results of the SVM with the tabu search algorithm (SVM-T) and the original SVM are compared. Furthermore, the mean time-to-detect is used to be computed and compared for the SVM model with the tabu search algorithm and without it. The comparison results are shown in Fig. 6.

From Fig. 6 it can be observed that, among the three methods, the detection performances of SVM-T and the SVM are better than that of MTTD. This can be

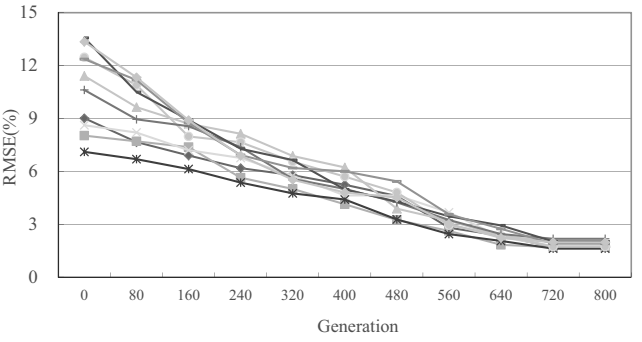

Fig. 5. Fitness of each calculation by the tabu search algorithm.

Table 1. Performance of the SVM and the ANN on the section from Dalian to Pulandian.

Method	DR	FAR	MTTD	MCR
SVM	96.2	4.24	73.4	3.52
ANN	92.9	5.08	71.2	4.63

Table 2. Performance of the SVM and the ANN on the section from Dalian to Zhuanghe.

Method	DR	FAR	MTTD	MCR
SVM	95.7	4.82	72.6	3.50
ANN	93.6	4.96	70.7	4.71

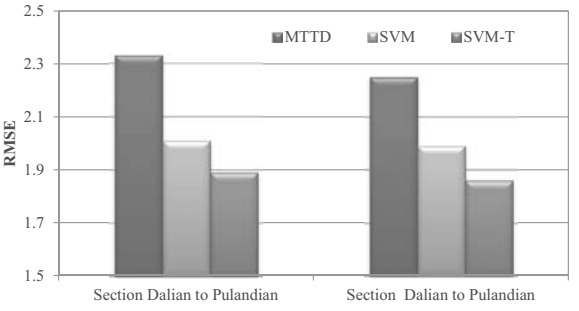

Fig. 6. Comparison results for MTTD, the SVM and SVM-T.

explained by the fact that the SVM uses the structural risk minimization principle to minimize the generalization error, while MTTD uses the average time to detect incidents, which is not suitable for the input–output non-linear relationship, especially for data of a relatively small size. Furthermore, compared with the SVM and SVM-T, it can be seen that the tabu search algorithm can greatly improve the detection performance of SVM. This is because the parameters optimized by the tabu search algorithm are more suitable for the freeway incident detection. Based on the results of Fig. 5, it can be attained that the tabu search algorithm can improve the detection performance of the SVM and the proposed SVM-T is effective for freeway incident detection.

4. Conclusions

The paper attempted to use support vector machines for freeway incident detection. To improve the prediction performance of the SVM, a tabu search algorithm was used to optimize the parameters for the SVM. The prediction model was tested on the data of two freeways in China to validate the feasibility and efficiency of the model. The results show that the tabu search algorithm has good convergence and relative stable performance. Furthermore, to evaluate the prediction performance of the proposed SVM, an ANN is used on the same data as the SVM. The comparison results with the proposed SVM and ANN suggest that the former provides a lower prediction error and a longer time. This indicates that the proposed SVM is a powerful tool for freeway incident detection.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (51208079 and 11272075), the National Science Foundation for Post-doctoral Scientists of China (2013M530924) and the National Science & Technology Pillar Program (2013BAG05B00).

Maoqing Jin is the corresponding author for this paper.

References

Ahmed, S.R. and Cook, A.R.(1982). Application of time-series analysis techniques to freeway incident detection, *Transportation Research Record* **841**: 19–21.

Augugliaro, A., Dusonchet, L. and Sanseverino, E.R. (2002). An evolutionary parallel Tabu search approach for distribution systems reinforcement planning *Advanced Engineering Informatics* **16**(3): 205–215.

Bortfeldt, A., Gehring, H. and Mack, D. (2003). A parallel tabu search algorithm for solving the container loading problem, *Parallel Computing* **29**(5): 641–662.

Cao, L.J. and 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.

Chen, S. and Wang, W. (2009). Decision tree learning for freeway automatic incident detection, *Expert Systems with Applications* **36**(2): 4101–4105.

Cristianini, N. and Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, Cambridge University Press, New York, NY.

Dong, B., Cao, C. and Lee, S.E. (2005). Applying support vector machines to predict building energy consumption in tropical region, *Energy and Buildings* **37**(5): 545–553.

Falco, D., Del Balio, R., Tarantino, E. and Vaccaro, R. (1994). Improving search by incorporating evolution principles in parallel tabu search, *IEEE Conference on Evolutionary Computation, Orlando, FL, USA,* Vol. 2, pp. 823–828.

Hagan, M.T., Demuth, H.B., and Beale, M. (1996). *Neural Network Design*, PWS, Boston, MA.

Ho, S.C. and Haugland, D. (2004). A tabu search heuristic for the vehicle routing problem with time windows and split deliveries, *Computers & Operations Research* **31**(12): 1947–1964.

Hou, S.M. and Li, Y.R. (2009). Short-term fault prediction based on support vector machines with parameter optimization by evolution strategy, *Expert Systems with Applications* **36**(10): 12383–12391.

Jeleń, L., Fevens, T. and Krzyżak, A. (2008). Classification of breast cancer malignancy using cytological images of fine needle aspiration biopsies, *International Journal of Applied Mathematics and Computer Science* **18**(1): 75–83, DOI: 10.2478/v10006-008-0007-x.

Jin, X., Cheu, R.L. and Srinivasan, D.(2002). Development and adaptation of constructive probabilistic neural network in freeway incident detection, *Transportation Research* **10**(2): 121–147.

Lebrun, G., Charrier, C., Lezoray, O. and Cardot, H. (2008). Tabu search model selection for SVM Lebrun, *International Journal of Neural Systems* **18**(1): 19–31.

Lin, J.Y., Cheng, C.T. and Chau, K.W. (2006). Using support vector machines for long-term discharge prediction, *Hydrological Sciences Journal* **51**(4): 599–612.

Lin, S.W., Ying, K.C., Chen, S.H. and Lee, Z.J. (2008). Particle swarm optimization for parameter determination and feature selection of support vector machines, *Expert Systems with Applications* **35**(4): 1817–1824.

Lorena, A.C. and de Carvalho, A.C.P.L.F. (2008). Evolutionary tuning of SVM parameter values in multiclass problems, *Neurocomputing* **71**(16–18): 3326–3334.

Mahmoud, T.A. (2011). Adaptive control scheme based on the least squares support vector machine network, *International Journal of Applied Mathematics and Computer Science* **21**(4): 685–696, DOI: 10.2478/v10006-011-0054-6.

Pardo, M. and Sberveglieri, G. (2005). Classification of electronic nose data with support vector machines, *Sensors and Actuators B* **107**(2): 730–737.

Peter, T. (2013). Modeling nonlinear road traffic networks for junction control, *International Journal of Applied Mathematics and Computer Science* **22**(3): 723–732, DOI: 10.2478/v10006-012-0054-1.

Ren, J.T., Ou, X.L., Zhang, Y., and Hu, D.C. (2002). Research on network level traffic pattern recognition, *IEEE 5th International Conference on Intelligent Transportation Systems, Singapore*, pp. 500–504.

Reyna, R., Giralt, A., and Esteve, D. (2001). Head detection inside vehicles with a modified SVM for safer airbags, *IEEE 4th International Conference on Intelligent Transportation Systems, Oakland, MN, USA*, pp. 500–504.

Shawe-Taylor, J. and Cristianini, N. (2004). *Kernel Methods for Pattern Analysis*, Cambridge University Press, New York, NY.

Srinivasana, D., Jin, X. and Cheu, R.L. (2005). Adaptive neural network models for automatic incident detection on freeways, *Neurocomputing* **64**: 473–496.

Sumi, S.M., Zaman, M.F. and Hirose, H. (2012). A rainfall forecasting method using machine learning models and its application to the Fukuoka city case, *International Journal of Applied Mathematics and Computer Science* **22**(4): 841–854, DOI: 10.2478/v10006-012-0062-1.

Talbi, E.G., Hafidi, Z. and Geib, J.M. (1998). A parallel adaptive tabu search approach, *Parallel Computing* **24**(14): 2003–2019.

Vapnik, V.N. (1999). An overview of statistical learning theory, *IEEE Transactions on Neural Networks* **10**(5): 988–999.

Vapnik, V.N. (2000). *The Nature of Statistical Learning Theory*, Springer, New York, NY.

Wei, C. and Wu, K. (1997). Developing intelligent freeway ramp metering control systems, *National Science Council in Taiwan* **7C**(3): 371–389.

Wu, C.H., Ho, J.M. and Lee, D.T. (2004). Travel-time prediction with support vector regression, *IEEE Transactions on Intelligent Transportation Systems* **5**(4): 276–281.

Yao, B.Z., Hu, P., Lu X.H., Gao, J.J. and Zhang, M.H. (2013). Transit network design based on travel time reliability, *Transportation Research C*, DOI:10.1016/j.trc.2013.12.005, (in press).

Yao B.Z., Yang, C.Y., Yao, J.B. and Sun, J. (2010). Tunnel surrounding rock displacement prediction using support vector machine, *International Journal of Computational Intelligence Systems* **3**(6): 843–852.

Yu, B., William, H.K.L. and Mei, L.T. (2011). Bus arrival time prediction at bus stop with multiple routes, *Transportation Research C* **19**(6): 1157–1170.

Yu, B., Yang, Z.Z., Chen, K. and Yu, B. (2010). Hybrid model for prediction of bus arrival times at next station, *Journal of Advanced Transportation* **44**(3):193-204.

Yu, B., Yang, Z.Z. and Li, S. (2012). Real-time partway deadheading strategy based on transit service reliability assessment, *Transportation Research A* **46**(8): 1265–1279.

Yu, B., Yang, Z.Z. and Yao, B.Z. (2006). Bus arrival time prediction using support vector machines, *Journal of Intelligent Transportation Systems* **10**(4): 151–158.

Yuan F. and Cheu R.L. (2003). Incident detection using support vector machines, *Transportation Research C* **11**(3–4): 309–328.

Zhang, X.L., Chen, X.F. and He, Z.J. (2010). An ACO-based algorithm for parameter optimization of support vector machines, *Expert Systems with Applications* **37**(9): 6618–6628.

Baozhen Yao received her Ph.D. degree from Beijing Jiaotong University, China, in 2011. Currently she is a postdoctoral researcher in the School of Automotive Engineering, Dalian University of Technology, China. Her present research interests include public transportation, transportation planning, as well as management and logistic systems.

Ping Hu is a professor in the School of Automotive Engineering, Dalian University of Technology, China. The scope of his research interests is very broad, covering intelligent algorithms, automotive technology and computer software, as well as network technology.

Mingheng Zhang received his Ph.D. degree from Jilin University, Changchun, China, in 2007. Currently, he is a lecturer in the School of Automotive Engineering, Dalian University of Technology, China. His present research interests include artificial intelligence, intelligent vehicle navigation and image processing.

Maoqing Jin is the director of the Department of Transportation of the High Technology Research and Development Center, Ministry of Science and Technology, Beijing, China. Her current research interests include public transportation, transportation planning and management, communication and transportation engineering, urban logistics, ITS and other areas.

Received: 18 July 2013
Revised: 30 October 2013