CHEMICAL ENGINEERING TRANSACTIONS
VOL. 51, 2016

A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.aidic.it/cet

Guest Editors: Tichun Wang, Hongyang Zhang, Lei Tian
Copyright © 2016, AIDIC Servizi S.r.l.,
ISBN 978-88-95608-43-3; **ISSN** 2283-9216

DOI: 10.3303/CET1651101

An Improved Support Vector Machine Algorithm and its Application in Intelligent Transportation System

Ronghui Fu

College of Computer Science, Neijiang Normal University, Neijiang, SiChuan, China, 641110
furonghui@163.com

Intelligent transportation system as the combination of computer technology, information technology, communication technology, electronic control technology and transportation has been a hot point has been widely used to save many existing problems in transportation. It is widely accepted that traffic incident has strong randomness and unpredictable destructiveness Support vector machine proposed by Vapnik et al. is introduced in this paper to solve the existing problems in traffic incidents in order to help improve the efficiency and effect in intelligent transportation system. Here, support vector machine is improved by introducing particle swarm optimization (PSO) which is a powerful and easy way to implement. This improved method can optimize both the optimal feature subset and parameters in SVM, which can further to solve time in computation. Finally, an experiment is demonstrated to show the application of the proposed method in intelligent transportation system.

1. Introduction

With the rapid development of highway construction, our country has acquired significant economy benefits and social benefits. Highway as the result of economy development has been an important symbol of modernization of a country (Levin and Krause, 1978; Li and McDonald, 2004). However, recently many dangerous events have happened and resulted in losing lives. Then, intelligent transportation system, as the combination of computer technology, information technology, communication technology, electronic control technology and transportation has been a hot point in the academy and industry fields (Janusova and Cicmancova, 2016). Intelligent transportation system has been widely used to save many existing problems in transportation such as highway. Traffic incident has strong randomness and unpredictable destructiveness (Yuan and Chen, 2003). So, forecast related to traffic incidents in transportation especially in highway is also a difficult problem.

Up to now, many researchers have proposed many algorithms and models to deal with this problem (Aburomman and Ibne, 2016; Dae, et al., 2015). As early as 1978, Levin and Krause developed Bayesian algorithm to distinguish crowded events by calculating the change of occupancy of traffic emergency. Based on floating car, Li and McDonald proposed a new detection algorithm to detect traffic incidents in highway. Here, support vector machine is a popular algorithm to handle prediction and detection problems in intelligent transportation system (Le, 2014).

Support vector machine proposed by Vapnik et al. is a set of related supervised learning methods used for prediction and classification. To build a SVM based prediction and classification model, feature subset selection is an important issue. There is advantage to limit the number of input features in a classifier to produce a good predictive and less computationally intensive model. With a small and appropriate subset, the rational classification decision can be generated easier. In general, SVM uses a function to map the data into a different feature space. This transformation often comes in the form of mapping to a high-dimensional space. A function used to perform this transformation is called kernel function which plays a critical role both in the theory and application of SVM.

In this paper, based on particle swarm optimization (PSO) (Malik et al., 2011; Kenedy and Eberhart, 1995), an improved SVM alogrithm is proposed to be applied in intelligent transportation system so as to detect or predict traffic incidents in time. POS is a population-based search algorithm that is initialized with a population of random particles. Each particle in the POS flies through the search space at a velocity that is dynamically adjusted according to its own and its companion's historical behavior. This improved method can optimize both the optimal feature subset and parameters in SVM, which can further to solve time in computation. Finally, an experiment is demonstrated to show the application of the proposed method in intelligent transportation system.

This pape is organized as follows. Section 2 describes the basic concepts of kernel function and support vector machine. Section 3 demonstrates the concepts of particle swarm optimization and the process of the improved SVM algorithm. Section 4 shows the appliction of the proposed method in intelligent transportation system. Stection 5 concludes this paper. Section 6 gives the references in this paper.

2. Basic concepts

In this section, we will introduce the basic concepts about kernel function and SVM so as to help understand the proposed method in section 3.

2.1 Kernel function

In general, kernel function as an important function in support degree machine can provide a way to transform from the linear learning machine to nonlinear learning machine. Then, we will demonstrate some relative concepts of kernel function in the following.

Definition 1. (Feature Space) Given an initial data set $s=\{(x_1, y_1),(x_2, y_2),(x_i, y_i)\}$, where $x_i \subseteq R^n$ and $y_i \subseteq R^m$. A feature mapping is

$$\phi: x \in R^n \to H \subseteq R^n, \tag{1}$$

where H is feature space, denoting space $Hilbert$.

Definition 2. (Kernel function) Let binary function $K(x,x')$ belonging to $R^n \times R^n$ be a kernel function of $R^n \times R^n$. If there is a mapping from R^n to any feature space H, then it can be obtained that

$$K(x,x') = \langle \phi(x_i), \phi(x_j) \rangle, \tag{2}$$

where $<\phi(x_i), \phi(x_j)>$ denotes inner product of H.

Definition 3. (Positive semi-definite matrix) Let binary function $K(x_i, x_j):R^n \times R^n \to R$ be positive definite. If it is symmetric such as $K(x_i, x_j)=K(x_j, x_i)$, and satisfies

$$\sum_{i,j=1}^{m} \alpha_i \alpha_j K(x_i, x_j) \geq 0, \tag{3}$$

where $m \in I$ (positive integer), $x_i, x_{2\ldots}, x_m \in R^n$ and $a_i, a_{2\ldots}, a_m \in R^n$.

Definition 4. (*Gram* matrix) Given $K(x,x'):R^n \times R^n \to R$ and $x_i, x_{2\ldots}, x_i \in R^n$. Then, the element in the *i* line of the *j* column constructs a $l \times l$ matrix of $K_{ij}=K(x_i, x_j)$. Therefore, K is a $Gram$ matrix of $K(x,x')$ related to $x_i, x_{2\ldots}, x_l$.

The common core function includes the following types:

(1) polynomial core function,

$$k(x_i, x) = (xx_i + 1)^d;$$

(2)Gauss core function,

$$k(x_i, x) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right);$$

(3) Sigmoid core function,

$$k(x_i, x) = \tanh(v < x, x_i > c).$$

2.2 Modelling of SVR

Based on VC dimension theory of statistical learning and structural risk minimization, Vapnik et al. proposed support vector machine as a new general learning method. Through using finite sample information, SVM can search an optimal compromise between complexity and learning capability of model to obtain better predictive ability and solve some practical problems such as small sample, non-linearity, high dimension and partial

minimal point. Here, SVM only need to determine insensitive coefficient, penalty factor and core parameter where the modelling is simpler than BP network.
Firstly, given the training set T= s={(x_1, y_1),…,(x_N, y_N)}, where $x_i \in R^n$ and $y_i \in$(-1,1), i=1, 2, …, N. Then, we can obtain that

$$\min \frac{1}{2}\|\omega\|^2 \tag{4}$$

s.t. $y_i(\omega^T x_i + b) \geq 1$ (5)

where b is a threshold value used to classify.
In order to deal with this binary optimization problem, Lagrange equation where optimal solution is under constraint condition is constructed as follows:

$$L(\omega, b, \alpha) = \frac{1}{2}\|\omega\|^2 - \sum_{i=1}^{N} a_i \cdot (y_i(\omega^T x_i + b) - 1) \tag{6}$$

where α_i is Lagrange multiplier. Based on optimality principle, it can be acquired that

$$\frac{\partial L(\omega, b, \alpha)}{\partial \omega} = 0 \tag{7}$$

$$\frac{\partial L(\omega, b, \alpha)}{\partial b} = 0 \tag{8}$$

Solving the Eqs. (7)-(8), it can be further obtained that

$$w = \sum_{i=1}^{N} y_i \alpha_i x_i \tag{9}$$

$$\sum_{i=1}^{N} y_i \alpha_i = 0 \tag{10}$$

Then, using Eq. (10), we can obtain that

$$\min \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} y_i y_j \alpha_i \alpha_j \left(x_i \cdot x_j\right) - \sum_{j=1}^{N} \alpha_j \tag{11}$$

S.t. $\sum_{i=1}^{N} y_i \alpha_i = 0, \alpha_i \geq 0 \quad i = 1, 2, \cdots, N$. (12)

Using optimization algorithm, we can acquire that

$$b^* = y_i - \sum_{i=1}^{N} y_i \alpha_i^* (x_i \cdot x_k) . \tag{13}$$

After that, it can be derived that

$$y(x) = \mathrm{sgn}\left[\sum_{i=1}^{N} y_i \alpha_i^* (x_i \cdot x_k) + b^*\right], \tag{14}$$

where sgn represents sign function.
Parameter optimization and feature selection are two important aspects in SVM. Now, many researchers have devoted to study these two points such as grid searching method, gradient descent and meta-heuristic algorithm (Cordeiro and Pappa, 2011; Garcia Nieto et al., 2016). However, there are some drawbacks in these methods. So, it is difficult to obtain optimal result. Then, in this paper, we propose an improved SVM method based on POS algorithm.

3. Improved support vector machine

In this section, the new SVM algorithm is proposed and the process is demonstrated below.

Firstly, we will introduce the concepts of particle swarm optimization which is an evolutionary computation technique. Inspired by social behaviour among individuals, particles can represent a potential problem solution moving through n-dimensional search space. The PSO starts out with a set of agents called particles in random velocity. Each particle shows a record of the position of its previous best performance in a vector called *pbest*. The *nbest* is another best value and can be tracked by the particle swarm optimizer, which can be also obtained so far by any particle in that particle's neighbourhood. If a particle takes the entire population as its topological neighbours, the best value is a global best and is called *gbest*. It is important step to define fitness in POS. Classification accuracy and the number of feature are two criteria used to define a fitness function.

Let $p_{i,j}$ denote the best previous position encounter by the *i*th particle. $p_{g,j}$ denotes the global best position so far and t denotes the iteration counter. The current velocity of *d*th dimension of the *i*th particle at time t is defined as

$$v_{i,j}(t)=v_{i,j}(t-1)+c_1\times r_1\times\left(p^t_{i,j}-x^t_{i,j}\right)+c_2\times r_2\times\left(p^t_{g,j}-x^t_{g,j}\right)$$

where r_1 and r_2 are random function in the range [0,1], position constant c_1 and c_2 are personal and social learning factors, $X = \{x_1, x_2, \ldots, x_n\}$ is population consisted by n particles, velocity is restricted to the $[-v_{max},v_{max}]$ range in which v_{max} denotes the predefined boundary value.

Compared with genetic algorithm, POS has no evolution operators such as crossover which makes it easy to implement with great success to several problems.

In order to obtain better performance of SVM, POS algorithm is introduced to acquire balance between global level and local level.

Step 1: Select n+2 dimensional particles to code individual.

Step 2: Initialize population randomly and assign upper and lower bound, initial speed, size of population and the number of iterations.

Step 3: Train SVM model using the feature sets selected in Step 2.

Step 4: Construct objective function considering ACC, SVs and the number of features.

$$\begin{cases} f_1 = avgacc = \dfrac{\sum_{i=1}^{K} Test\ Accuracy}{K} \\ f_2 = \left(1-\dfrac{nsv}{m}\right) \\ f_3 = avgacc = \left(1-\dfrac{\sum_{i=1}^{n} ft_i}{n}\right) \\ f = \alpha\times f_1+\beta\times f_2+\lambda\times f_3 \end{cases} \tag{15}$$

α, β, γ represents the weights of classification accuracy, the number of support vector and selected feature sets. Weight can adjust the contributions of different objectives on classification.

Here, α, β, γ can be obtained from $\alpha=(\alpha_1\text{-}\alpha_2)\dfrac{t}{t_{\max}}+\alpha_2$, $\beta=(\beta_1\text{-}\beta_2)\dfrac{t}{t_{\max}}+\beta_2$, $\lambda=(\lambda_1\text{-}\lambda_2)\dfrac{t}{t_{\max}}+\lambda_2$ which satisfy $\alpha_1+\beta_1+\lambda_1=1$ and $\alpha_2+\beta_2+\lambda_2=1$.

Step 5: Increase the number of iterations, iter = iter +1.

Step 6: Increase the number of population and renew the position and speed of each particle.

Step 7: Introduce mutation strategy as follow:

$$x_k^{'}=\begin{cases} x_k+\Delta(t,UB-x_k) & if\ rand=0 \\ x_k-\Delta(t,x_k-LB) & if\ rand=1 \end{cases}, \tag{16}$$

where x_k demonstrates a variation in particles, $x_k^{'}$ demonstrates the value after mutation and UB and LB shows upper limit and lower limit of x_k.

Step 8: Train SVM model using the feature sets selected in Step 6 and calculate fitness values of each particle based on Eq. (15).

Step 9: Compare the existing fitness value and *pfit* in internal storage to renew optimal fitness value of individual *pfit* and optimal position of individual *pbest*.
Step 10: If the order of particles reaches the maxi size of population, go to Step 11. Otherwise, go to Step 6.
Step 11: Compare the *gfit* value and the optimal *pfit* to renew optimal fitness value of global level *gfit* and optimal position of global level *gbest*.
Step 12: If condition is satisfied, then go to Step 13. Otherwise, go to Step 5.
Step 13: Obtain optimal parameters from *gbest*.

4. Simulated Experiment

In this section, we will demonstrate the application of the proposed method in intelligent transportation system. The simulated data set is collected from traffic flow of road system in main road, which can signify the classificatory accuracy and correct features selection of the improve SVM method. Generally, the drawback of evolutionary-based such as the GA is the long training time when deal with a large scale data set. But in this case, by using POS, the improved method can implement very well. This experiment is performed by using MATLAB 2011. In the method, the particle movement update is performed in the application server site. A new SVM classification is constructed and the feature subset is prepared. The application server can find the global and personal best of each particle according to particle's fitness and update the new position for each particle using the particle's velocity and update formula. Some parameters are set as α_1, β_1, λ_1, α_2, β_2 and λ_2 are equal to 0.3, 0.7, 0, 0.6, 0.3, and 0.1. c_1 and c_2 are set as 2. The fitness value of selected two data sets from 9.00-10.00 and 18.00-17.00 in Figure 1. The accuracy of data sets is demonstrated in the Figure 2.

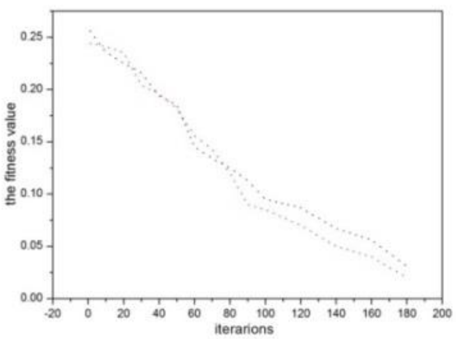

Figure 1: The fitness value of the training set

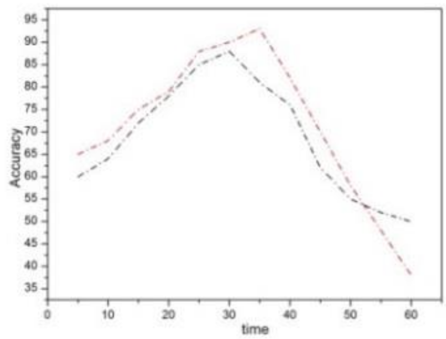

Figure 2: The accuracy of data sets during different period

Then, based on the mentioned two figures, it is easy to know that the accuracy of the proposed method range from 50%-90%. Different data set in different time has different accuracy. With the number of iteration, the fitness is decreased from the global view. To sum up, the improved method in this paper have better accuracy of data sets and feature selection which can help us to predict the data flow in intelligent transportation system and further control traffic incidents.

5. Conclusion

With the rapid development of highway construction, our country has acquired significant economy benefits and social benefits. However, recently many dangerous transportation events have happened and resulted in losing lives. Then, intelligent transportation system, as the combination of computer technology, information technology, communication technology, electronic control technology and transportation has been a hot point in the academy and industry fields. It is widely accepted that traffic incident has strong randomness and unpredictable destructiveness Support vector machine proposed by Vapnik et al. is introduced in this paper to solve the existing problems in traffic incidents in order to help improve the efficiency and effect in intelligent transportation system. Here, support vector machine is improved by introducing particle swarm optimization (PSO) which is a powerful and easy way to implement. This improved method can optimize both the optimal feature subset and parameters in SVM, which can further to solve time in computation. Finally, an experiment is demonstrated to show the application of the proposed method in intelligent transportation system.

Acknowledgment

Project Name: Provincial Excellent Engineer Programs-The Software Engineering(the file of Neijiang Normal University in 2014 with number 19).

Reference

Aburomman A. A., Ibne Reaz M. B., 2016, A novel SVM-KNN-POS ensemble method for intrusion detection system, Applied Soft Computing, 38, 360-372.

Cordeiro Z., Pappa C. L., 2011, A POS algorithm for improving multi-view classification, in: 2011 IEEE Congress on Evolutionary Computation (CEC), IEEE, 925-932.

Dae Ko Y., Jae Jang Y., Seok Lee M., 2015, The optimal economic design of the wireless powered intelligent transportation system using genetic algorithm considering nonlinear cost function, 89, 67-79.

Garcia Nieto P. J., Garcia-Gonzalo E., Alonso Fernandez J. R., Diaz Muniz C., 2016, A hybrid PSO optimized SVM-based model for predicting a successful growth cycle of the spirulina platrnsis from raceway experiments data, Journal of Computational and Applied Mathematics, 291, 293-303.

Janusova L., Cicmancova S., 2016, Improving safety of transportation by using intelligent transport systems, Procedia Engineering, 134, 14-22.

Kenedy J., Eberhart R. C., 1995, Particle swarm optimization, in: Proceedings of IEEE International Conference on Neural Networks, Piscataway, NJ, 1942-1948.

Le Thi H. A., Thanh Vo X., Pham Dinh T., 2014, Feature selection for linear SVMs under uncertain data: Robust optimization based on difference of convex functions algorithms, Pattern Recognition Letters, 42, 40-46.

Levin M., Krause G. M., 1978, Incident detection: A Bayesian approach.

Li Y., McDonald M. 2004, Motorway incident detection using probe vehicles. Transport, 158, 11-15.

Yuan F., Cheu R. L., 2003, Incident detection using support vector machines. Transportation Research Part C, 11, 309-328.

Malik A. J., Shahzad W., Kahn F. A., 2011, Binary POS and random forests algorithm for probe attacks detection in a network, in: 2011 IEEE Congress on Evolutionary Computation (CEC), IE