ARCHIVES OF TRANSPORT *ISSN (print):* *0866-9546*
Volume 39, Issue 3, 2016 ***e-ISSN (online):*** *2300-8830*
DOI: *10.5604/08669546.1225447*

TRAFFIC FATALITIES PREDICTION BASED ON SUPPORT VECTOR MACHINE

Ting Li[1], Yunong Yang[1], Yonghui Wang[1], Chao Chen[2], Jinbao Yao[3]

[1]Dalian Maritime University, Transportation Management College, Dalian, PR China

[2]Dalian University of Technology, Automotive Engineering College, Dalian, PR China

[3]Beijing Jiaotong University, School of Civil Engineering and Architecture, Beijing, PR China

[3]e-mail: bao_yaojin@163.com

To effectively predict traffic fatalities and promote the friendly development of transportation, a prediction model of traffic fatalities is established based on support vector machine (SVM). As the prediction accuracy of SVM largely depends on the selection of parameters, Particle Swarm Optimization (PSO) is introduced to find the optimal parameters. In this paper, small sample and nonlinear data are used to predict fatalities of traffic accident. Traffic accident statistics data of China from 1981 to 2012 are chosen as experimental data. The input variables for predicting accident are highway mileage, vehicle number and population size while the output variables are traffic fatality. To verify the validity of the proposed prediction method, the back-propagation neural network (BPNN) prediction model and SVM prediction model are also used to predict the traffic fatalities. The results show that compared with BPNN prediction model and SVM model, the prediction model of traffic fatalities based on PSO-SVM has higher prediction precision and smaller errors. The model can be more effective to forecast the traffic fatalities. And the method using particle swarm optimization algorithm for parameter optimization of SVM is feasible and effective. In addition, this method avoids overcomes the problem of "over learning" in neural network training progress.

Key words: *traffic accident; support vector machine (SVM); Particle Swarm Optimization (PSO); prediction model; optimal parameters.*

1. Introduction

1.1. Background

In recent years, road infrastructure construction of China has improved constantly. Vehicle number and highway mileage increase accordingly. This drives the development of economy, but also produces some negative effects, i.e. road traffic accidents occur frequently. In traffic accidents, the traffic fatality is the most harmful, and become a menace to our life. It has become a severity societal problem we should pay full attention to. However, the traffic fatality has randomness which is affected by the factors such as driver and passenger characteristics, vehicles types, traffic conditions, as well as geometric design characteristics, etc. However, the complex relationship between traffic fatalities and various influencing factors is nonlinear. As some factors generate influence on one another, it is hard to use only one factor to fully explain traffic fatalities. Therefore, it is necessary to summarize and analyze the data about traffic safety. The development trend of traffic fatalities under the existing road traffic conditions can be predicted by finding out the inherent law of accident. For making plan and decision of road traffic safety, prediction of traffic fatalities has practical significance.

1.2. Literature review

Currently, many methods in traffic accident prediction are used, the application conditions and modeling mechanism of which are different. Binomial regression, Bayesian approach, back-propagation neural network models and some new methods are used to fit the accident data. Pamuła (2012) presented a method of classification of time series of traffic flow. Nowakowska (2012) pointed highlight road traffic accident patterns in the context of interrelations between road characteristics and a traffic safety threat. Ghasemlou et al. (2015) aimed to predict the crash severity with the traffic injury data by implementing the Artificial Neural Networks (ANN), Regression Trees (RT) and Multiple Linear Regression modelling (MLRM) method. Mitas et al. (2013) researched the traffic security level. Poch et al. (1996) estimated a negative binomial regression of the frequency of accidents at intersection approaches. Clarke et al. (1998) employed a machine learning method to create decision trees. The characteristics of accidents that resulted in injury or in damage only are distinguished. Abdel-Aty et al. (2011) explored to combine multivariate adaptive regression splines (MARS) with another machine learning technique (random forest). Xu et al. (2013) aimed to build the genetic programming (GP) model for real-time crash prediction on freeways. The application of the model was evaluated. Ramani and Selvaraj (2014) optimized the Aggregated Feature Selection (VAAFS) with Voting Algorithm. An optimal number of significant features with majority votes were selected. Other similar traffic accident prediction can be found in these literatures (Tesema et al., 2005; Lee and Wei, 2010; Zhang et al., 2014; Nassiri et al., 2014; Zong et al., 2013a; Yao et al., 2014).

Some research has proposed innovative models to predict traffic accident. Yasdi (1991) and Quek et al. (2006) employed artificial neural networks for traffic forecasting which was applied on a road section. Xie et al. (2007) evaluated the application of Bayesian neural network models for predicting motor vehicle crashes. Kunt et al. (2011) employed twelve accident-related parameters in a genetic algorithm (GA), pattern search and artificial neural network (ANN) modelling methods. Stelmach (2012) dealed with mathematical modeling of the aircraft landing phase using artificial neural networks. The severity of freeway traffic accidents was predicted by these models. Similar researches had been studied by these literatures (Zong et al., 2013b, Abdelwahab & Abdel-Aty, 2001; Deublein et al., 2013).Though artificial neural network has characteristic of identifying complex non-linear system, there are problems, such as slow convergence speed, over-learning and local extreme. All these problems affect the prediction precision of it.

In recent years, support vector machine is used in traffic accident prediction. SVM can study and optimize itself and adjust according to the variation of data (Yao et al., 2014). Problems of small sample, non-linear and local extreme can be solved by it. Li et al. (2008) predicted motor vehicle crashes applying Support Vector Machine (SVM) models. The study showed that SVM models predict crash data more effectively and the accuracy is higher than traditional Negative Binomial (NB) models. Li et al. (2012) developed a SVM model for predicting the injury severity associated with individual crashes. The performance of the SVM model and the ordered probit (OP) model were compared. Yang and Zhao (2013) introduced accident rate per 10000 cars and accident rate per 10000 capita in the paper. Based on the theory of support vector machine (SVM), the improved SVM models were proposed. SVM also has some disadvantages. For example, the performance of support vector machine depends on the parameters. Before the training phase, there are three parameters (C, v, γ) need to be determined. Many literatures suggested that heuristic algorithms have been successfully used in many complex problems (Yu et al., 2010, 2011, 2013), their algorithms are tested to be effective by the results. Furthermore, in majority of cases, the best results found are obtained by these algorithms, particularly on real-life optimization problems. To select the parameter values of SVM automatically, heuristic algorithms are also used in this paper.

1.3. Contributions

There are two main contributions in this paper: firstly, a prediction model of traffic accident based on PSO-SVM is proposed. Particle swarm optimization (PSO) algorithm is introduced to find the optimal parameter combination of SVM. Highway mileage, vehicle number and population size are put into the model to get the number of traffic fatalities, which is the most comparable indicator in traffic accident; secondly, the performance of the PSO-SVM, SVM and neural network prediction model are compared. Both fitting and predicting abilities of the models are evaluated through computing error values.

The rest of the paper is organized as follows: section 2 introduces the principle of support vector machine. The model and process of traffic accident prediction based on PSO-SVM are described respectively in section 3 and 4; Test results and error value comparison of different model are presented in section 5; finally, the conclusions and direction for future research are presented in section 6.

2. The principle of support vector machine (SVM)

The theory study of Support Vector Machine (SVM) has been fairly mature. This method is a learning method in small sample situation proposed by Vapnik (1999), based on the theory of statistical learning law. Support Vector Machine (SVM) algorithm maps the sample space to a high-dimensional feature space by nonlinear mapping, transferring the search for the optimal linear regression hyperplane algorithm into solving convex programming problem under convex constraint, so as to get the global optimal solution (Gan et al., 2010; Guan et al., 2008).At the same time, the Support Vector Machine (SVM) method changes the product calculation in the high-dimensional space into the kernel function calculation in the original space by defining a kernel function (Dong et al., 2005), which greatly simplifies the calculation.

In case that the training sample is nonlinear, the fitting function can be obtained through the method as follows. Through a nonlinear function, each sample point is mapped to a high-dimensional feature space, the linear regression in the high-dimensional feature space is performed, and then the nonlinear regression of the original space is got. The fitting function can be expressed as the following equation (Cao and Francis, 2003).

$$y = \omega\varphi(x) + \varepsilon \tag{1}$$

Where, ω is for the weight vector, x is input vector, and ε represents the offset value

To minimize the following two values through training,

$$P(f) = c\frac{1}{l}\sum_{i}^{l} L_{\varepsilon}(y_i - f(x_i)) + \frac{1}{2}\|\omega\|^2 \tag{2}$$

$$L_{\varepsilon}(y_i - f(x_i)) = \begin{cases} |y_i - f(x_i)| - \varepsilon & |y_i - f(x_i)| \geq \varepsilon \\ 0 & |y_i - f(x_i)| < \varepsilon \end{cases} \tag{3}$$

Where, l is the total number of training samples, $c\frac{1}{l}\sum_{i}^{l} L_{\varepsilon}(y_i - f(x_i))$ is for the experienced error term, $\frac{1}{2}\|\omega\|^2$ is a regular item. $L_{\varepsilon}(y_i - f(x_i))$ represents the loss function, balancing the weighting function of training error term and the complex term. c is the penalty factor, ε stands for loss function parameter, whose value affects the number of support vector. Here introduces the slack variables ξ_i and ξ_i^*, and then the optimization problem can be converted into:

$$\min \frac{1}{2}|\omega|^2 + c\sum_{i}(\xi_i + \xi_i^*) \tag{4}$$

$$s.t. \begin{cases} y_i - \omega\varphi(x) - \varepsilon \leq \varepsilon + \xi_i \\ \omega\varphi(x) + \varepsilon - y_i \leq \varepsilon + \xi_i^* \end{cases} \tag{5}$$

Where, the Lagrange multiplier a_i and a_i^* are introduced, and the problem is transferred further into a simple optimization problem of the dual problem,

$$\begin{aligned} \max \sum_{i} y_i(a_i - a_i^*) - \theta\sum_{i}(a_i + a_i^*) \\ -\frac{1}{2}\sum_{i}\sum_{j}(a_i - a_i^*)(a_j - a_j^*)k(x_i, x_j) \\ -\sum_{i}(a_i - a_i^*) = 0 \end{aligned} \tag{6}$$

$$s.t. \quad 0 \leq a_i \leq C, 0 \leq a_i^* \leq C \tag{7}$$

The final prediction function finished is as follows,

$$y = \sum_{i}(a_i - a_i^*)k(x_i, x_j) + \varepsilon \tag{8}$$

Where, $k(x_i, x_j)$ represents the kernel function, which can complete the product operation of the input samples in low dimension of aunknown nonlinear mapping function in a high dimensional feature space. The kernel function.is the core of SVM, and different kernel functions have different structure. The detailed deduced process can refer to literature (Chang, 2005; Cao and Xu, 2007)

3. The prediction model of traffic fatalities based on PSO-SVM

To achieve comprehensive measure of traffic accident, choosing the index of traffic accident should follow three principles: measurability, representativeness and comparability. Traffic system is consisted of three basic factors, which are people, vehicle and road. Traffic accidents have great randomness, affected by many factors which are quantitative factors and qualitative factors. In related literature of traffic accident prediction, highway mileage, vehicle number, Lane width, average daily traffic and population size are selected as impact factors (Cao and Xu, 2007). In this paper, from the point of person, vehicle and road factors, highway mileage, vehicle number, population size are selected to be the set of impact factors.

Traffic accident prediction index currently widely used are number of traffic fatalities, number of injury, number of road accidents and economic loss. Because there are major differences on the definition of injuries and statistics about road accidents is incomplete. The number of traffic fatalities is the predictor index which is the most comparable one. Therefore, the structure of traffic fatalities prediction model can be seen in Fig. 1.

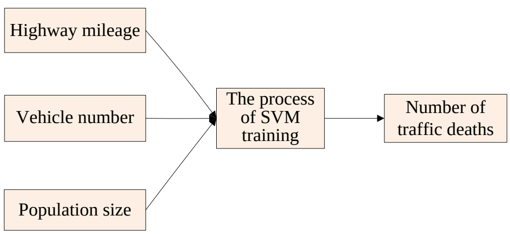

Fig. 1. The structure diagram of traffic fatalities prediction model

4. The process of traffic fatalities prediction based on PSO-SVM

Support Vector Machine (SVM) is a theory of machine learning law in small sample situation, and it has the very advantage in prediction, but there is no specific theory for parameter selection to depend on in the learning process for support vector machine, which seriously restrain the prediction accuracy and effect of the Support Vector Machine (SVM) method. The value of penalty factor c and kernel parameter σ affects the prediction accuracy of SVM, and finding the optimal c and σ is the priority. At present parameter is usually defined artificially based on the specific issues, and the optimal parameter combination is determined by choosing the parameters for many times and comparing with each other. Parameters that are manually set, are blind and of low efficiency, so it is needed to adopt swarm intelligence optimization algorithm to improve the parameter choosing of the Support Vector Machine (SVM).At the same time, the design and implementation of Particle Swarm Optimization algorithm (PSO) is relatively simple. Not only the convergence speed is fast, but the parameters required to be set are less (Cao and Xu, 2007).

Particle Swarm Optimization (PSO) algorithm is inspired by birds foraging behavior proposed by Eberhart and Kennedy (1995), which is a random search optimization algorithm generated by swarm intelligence based on group cooperation and competition. Compared with evolutionary computation, particle swarm optimization algorithm adopts a global search strategy, uses v-s model with simple operation, and abandons the complex genetic manipulation. Its special memory mechanism can adjust the search strategy by keeping track of the current search based on real-time, which makes PSO a kind of efficient parallel search algorithm. As a result of the fast convergence speed of particle swarm optimization algorithm and few requirements on parameter setting, PSO has drawn extensive concern in the academic field.

In the process of particle swarm optimization algorithm solving the problem, each particle are representative for a solution to the asked problem. Through the preset fitness function, each particle has its corresponding fitness value. Particle velocity determines the direction and distance moved, at the same time, by taking examples from motion inertia of the particle itself and the surrounding particle, the velocity can be dynamically updated timely, so as to achieve the search process of the solution. In every optimization search process, the particle is updated by two values. One value is the optimal solution obtained by the particle itself, known as the individual extremum, and the other is global optimal solution, called the global extremum. The specific steps are as follows:

Step1: Initialization. Initialize the number of the population and iterations, and speed.

Step2: Choose fitness function. The mean square error is to be set as fitness function

Step3: Update the particle position. Compare the current particle fitness value and the best location to the original. If the original fitness value is better, then the best fitness value remains unchanged. But if the current particle fitness value is better, the current fitness value is set to the best location.

Similarly is the comparison of the current fitness value and the global optimal value.

Step4: Adjust the particle velocity and position dynamically.

$$V_{i+1} = \omega V_i + c_1\lambda_1\left(P_i - X_i\right) + c_2\lambda_2\left(P_g - X_i\right),$$
$$X_{i+1} = X_i + V_{i+1}.$$

Where, $i = 1, 2, ..., n$, ω is the inertia weight to control the influence degree of the current speed by the former, c_1 and c_2 are accelerated factors, and λ_1 and λ_2 are random numbers within the range of $[0,1]$.

Step5: Judgment result. If the result has been the optimal, then output current optimal parameter values. If the optimal value is not achieved, then transfer to Step2.

Thus, the prediction model combined PSO algorithm with SVM (referred to as PSO-SVM method) is proposed to predict the traffic fatalities. The concrete steps of PSO-SVM method are as shown in Fig2.

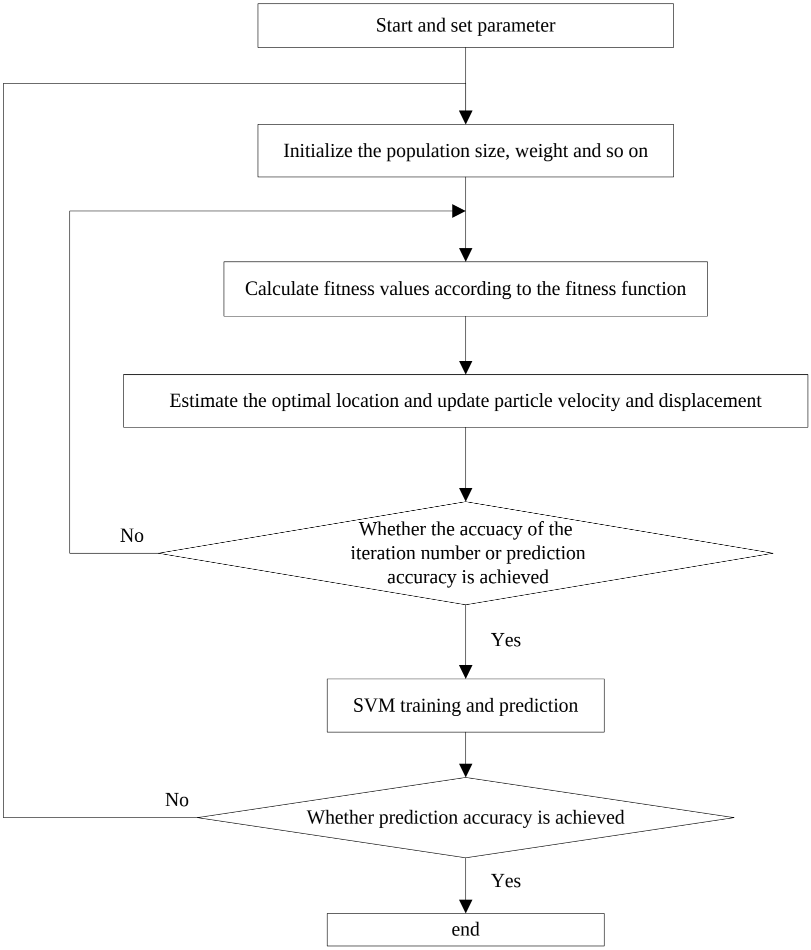

Fig. 2. The concrete steps of PSO - SVM method

5. Numerical test

5.1. Data

The occurrence of traffic accidents is a result of the combined action of many influence factors, for example, road traffic environment, vehicle number, etc. In this paper, highway mileage, vehicle number and population size are input variables for predicting accident. Traffic fatality is one of the results which are most harmful in traffic accident. And traffic accidents involving the fatalities have been highly valued and statistics have few omissions. Therefore, traffic fatality is used as the output variables. We collect the data from the website of National Bureau of Statistics of China, and the related data is shown in Table 1.

Sample data from 1981 to 2012 are chosen as experimental data. Samples of 1981-2006 are training data, while 2007-2012 are test data. In the process of training samples, parameters of PSO are set as follows: The population scale is of 20, the iteration number is 200, the cross rate is 0.7. The initial values of accelerating factor c_1 and c_2 are 1.5 and 1.7 respectively.

Data normalization

As the units of data are different and orders of magnitude difference are big, the data for each variable have to be normalized. If the model calculate 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:

$$A_i^l = \frac{A_i}{\|A\|_2} = \frac{A_i}{\sqrt{A_1^2 + A_2^2 + \ldots + A_i^2}} \tag{9}$$

Table 1. Traffic accident statistics of China

Year	Traffic fatalities people	Highway mileage ten thousand km	Vehicle number ten thousand	Population size ten thousand
1981	22499	89.75	234.73	100072
1982	22164	90.7	259.77	101654
1983	23944	91.51	283.88	103008
1984	25251	92.67	587.37	104357
1985	40906	94.24	655.65	105851
1986	50063	96.28	819.07	107507
1987	53439	98.22	1061.03	109300
1988	54814	99.96	1190.2	111026
1989	50441	101.43	1318.53	112704
1990	49243	102.83	1476.26	114333
1991	53204	104.11	1657.66	115823
1992	58723	105.67	1945.03	117171
1993	63551	108.35	2331.64	118517
1994	66362	111.78	2735.6	119850
1995	71494	115.7	3179.78	121121
1996	73655	118.58	3609.65	122389
1997	73861	122.64	4209.32	123626
1998	78067	127.85	4861.3	124761
1999	83529	135.17	5404.73	125786
2000	93853	167.98	6016.22	126743
2001	105930	169.8	6851.88	127627
2002	109381	176.52	7975.68	128453
2003	104372	180.98	9649.96	129227
2004	107077	187.07	10783.44	129988
2005	98738	334.52	13039.45	130756
2006	89455	345.7	14528.90	131448
2007	81649	358.37	15977.76	132129
2008	73484	373.02	16988.77	132802
2009	67759	386.08	18658.07	133450
2010	65225	400.82	20706.13	134091
2011	62387	410.64	22512.08	134735
2012	59997	423.75	24102.03	135404

Where, A_i is the i^{th} original value of variables that needs to be normalized. In this paper, it refers to the i^{th} original value of highway mileage, vehicle number and population size. And A_i^l is the i^{th} value of highway mileage, vehicle number and population size after normalization.

Performance index

Two evaluation criteria are adopted for this study to compare the performance of models.
The specific formula is as follows:
Mean absolute percentage error

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{10}$$

Coefficient of determination

$$(R^2) = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{11}$$

Where, n is the size of fitting or predicting sample; $\hat{y}_i$ is the estimated traffic fatalities at year i; and y_i is the observed number of traffic fatalities, $\bar{y}$ is the average value of traffic fatalities. The model performance is better if the value of MAPE is smaller and $R^2$is larger.

5.2. The training of the model

The training curve of traffic fatalities based on SVM prediction model is shown in Fig. 3. The black ones are actual output, while yellow ones are fitting output. It can be seen that they fit well. The mean absolute percentage error5.0133% and coefficient of determination (R^2) is 0.9475. The traffic accident prediction model based on PSO-SVM has strong identification ability, and the fitting is stable.

5.3. The prediction of the model

The traffic fatalities of 2007-2012 can be predicted by the model that has been trained. Fig. 4 shows the absolute percentage error of the traffic accident prediction. Mean absolute percentage error is 4.311% and coefficient of determination (R^2) is 0.947.

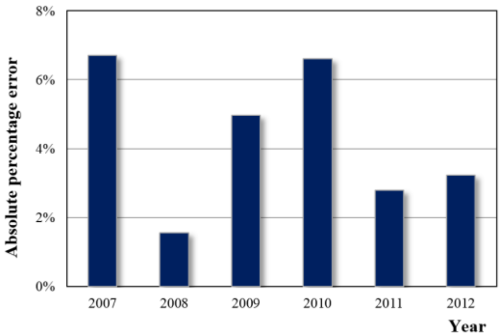

Fig. 4. Absolute percentage error of the traffic fatalities prediction

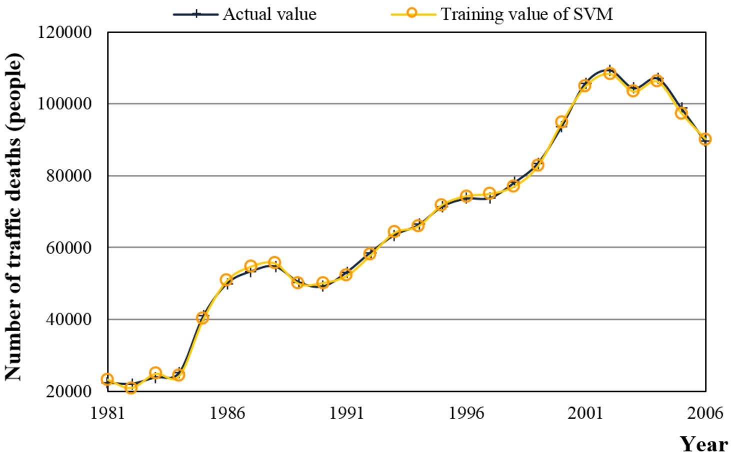

Fig. 3.Training diagram of SVM prediction model

In this paper, prediction results of PSO-SVM, SVM and BPNN are compared as the Table 2.Mean absolute percentage errors are 4.311%, 6.429% and 7.388% respectively. R^2 are 0.947, 0.8782 and 0.8234. Fig. 5 shows the predict results of the several method. It is obvious that the prediction model of SVM based on PSO is better than SVM and BPNN model. And the SVM model is slightly better than neural network model. This is because the SVM method with global optimality will not get into local minimum point. This avoids the defects of the neural network method and improves the prediction precision. SVM based on PSO searches the value of important parameters C, σ consecutively. The artificial selection for parameters can be avoided, so as to improve the prediction accuracy.

6. 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 PSO model is simple in program implementation, less in setting parameters and fast in calculating convergence speed. The prediction model of traffic fatalities based on PSO-SVM, which uses PSO to optimize the parameters of SVM, is the optimal SVM prediction model.

Table 2. Predict results comparison of PSO-SVM, SVM and BPNN

Year	Actual value (people)	PSO-SVM		SVM		BPNN	
		Predicted value	APE	Predicted value	APE	Predicted value	APE
2007	81649	76169	6.71%	73169	10.39%	86829	6.34%
2008	73484	74625	1.55%	75205	2.34%	76755	4.45%
2009	67759	71129	4.97%	71239	5.14%	74239	9.56%
2010	65225	69531	6.60%	71981	10.36%	58117	10.90%
2011	62387	64132	2.80%	65792	5.46%	66862	7.17%
2012	59997	61935	3.23%	62935	4.90%	63535	5.90%

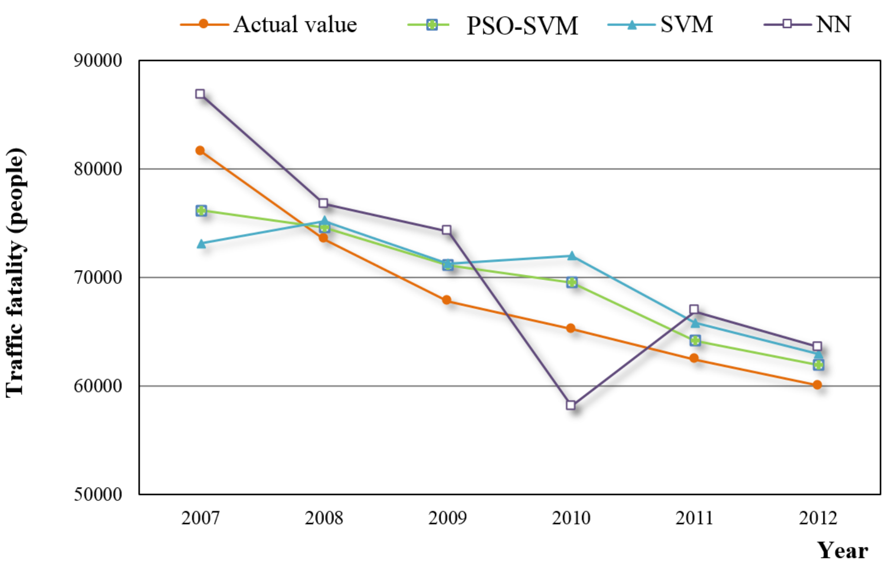

Fig. 5.Comparison of predict results

The example analysis results show that the forecasting method based on PSO-SVM model is superior to the forecasting method of neural network and the ANN 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 PSO-SVM is better than that of general prediction model of traffic fatalities and with better prediction accuracy.

Acknowledgments
This research was supported in National Natural Science Foundation of China 71571026 and 51578112, Liaoning Excellent Talents in University LR2015008, and the central universities DUT16YQ104.

References

[1] ABDEL-ATY, M., HALEEM, K., 2011. Analyzing angle crashes at unsignalized intersections using machine learning techniques. *Accident Analysis & Prevention*, 43(1), pp. 461-470.
[2] ABDELWAHAB, H. T., ABDEL-ATY, M. A., 2001. Development of artificial neural network models to predict driver injury severity in traffic accidents at signalized intersections. *Transportation Research Record: Journal of the Transportation Research Board*, 1746(1), pp. 6-13.
[3] CAO, C., XU, J., 2007. Short-term traffic flow predication based on PSO-SVM. *Proceedings of the first international conference on transportation engineering*, 246, p. 28.
[4] CAO, L. J., FRANCIS, E. H., 2003. Support vector machine with adaptive parameters in financial time series forecasting. *IEEE Transactions on Neural Networks*, 14(6).
[5] CHANG, L. Y., 2005. Analysis of freeway accident frequencies: negative binomial regression versus artificial neural network. *Safety science*, 43(8), pp. 541-557.
[6] CLARKE, D. D., FORSYTH, R., WRIGHT, R., 1998. Machine learning in road accident research: decision trees describing road accidents during cross-flow turns. *Ergonomics*, 41(7), pp. 1060-1079.
[7] DEUBLEIN, M., SCHUBERT, M., ADEY, B. T., KÖHLER, J., FABER M. H., 2013. Prediction of road accidents: A Bayesian hierarchical approach. *Accident Analysis & Prevention*, 51, pp. 274-291.
[8] DONG, B., CAO, C., LEE, S. E., 2005. Applying support vector machines to predict building energy consumption in tropical region. *Energy and Buildings*, 37, pp. 545-553.
[9] GAN, X. S.,DUANMU, J., CONG, W. 2010. Fatalness assessment of flight safety hidden danger based on support vector machine. *Journal of Safety Science and Technology*, 6(3), pp. 206-210.
[10] GHASEMLOU, K., AYDI, M. M., YILDIRIM, M. S., 2015. Prediction of pedal cyclists and pedestrian fatalities from total monthly accidents and registered private car numbers. *Archives of Transport*, 2015, 34(2), pp. 29-35.
[11] GUAN, Z., SONG, T, QI, L., 2008. An Application of Support Vector Machine in Foundation Settlement Prediction [J]. *Transactions of Shenyang Ligong University*, 2, p. 024.
[12] KENNEDY, J., EBERHART, R. C., 1995 Particle swarm optimization, *Proc. of IEEE International Conference on Neural Networks*, Piscataway, NJ, pp. 1942-1948.
[13] KUNT, M. M., AGHAYAN, I., NOII, N., 2011. Prediction for traffic accident severity: comparing the artificial neural network, genetic algorithm, combined genetic algorithm and pattern search methods. *Transport*, 26(4), pp. 353-366.
[14] LEE, Y., WEI, C. H., 2010. A computerized feature selection method using genetic algorithms to forecast freeway accident duration times. *Computer-Aided Civil and Infrastructure Engineering*, 25(2), pp. 132-148.
[15] LI, X., LORD, D., ZHANG, Y.,XIE, Y. 2008. Predicting motor vehicle crashes using support vector machine models. *Accident Analysis & Prevention*, 40(4), pp. 1611-1618.
[16] LI, Z., LIU, P., WANG, W., XIE, Y., 2012. Using support vector machine models for crash injury severity analysis. *Accident Analysis & Prevention*, 45, pp. 478-486.
[17] MITAS, A. W., KONIOR, W., KONIOR A., 2013. VMS parameters impact on safety and

reliability in Road traffic management. *Archives of Transport*, 25-26(1-2), pp. 69-83.

[18] NASSIRI, H., NAJAF, P., AMIRI, A. M., 2014. Prediction of Roadway Accident Frequencies: Count Regressions versus Machine *Learning Models. Scientia Iranica*, 21(2), pp. 263-275.

[19] NOWAKOWSKA, M., 2012. Road Traffic Accident Patterns: A Conceptual Grouping Approach to Evaluate Crash Clusters. *Archives of Transport*, 2012, 24(1), pp. 73-98.

[20] PAMUŁA, T., 2012. Classification and Prediction of Traffic Flow Based on Real Data Using Neural Networks. *Archives of Transport*, 2012, 24(4), pp. 519-529.

[21] POCH, M., 1996. Mannering, F. Negative binomial analysis of intersection-accident frequencies. *Journal of Transportation Engineering*, 122(2), pp. 105-113.

[22] QUEK, C., PASQUIER, M., LIM, B. B. S., 2006. POP-TRAFFIC: A novel fuzzy neural approach to road traffic analysis and prediction. *IEEE Transactions on Intelligent Transportation Systems*, 7(2), pp. 133-146.

[23] RAMANI, R. G., SELVARAJ, S. A., 2014. Pragmatic Approach for Refined Feature Selection for the Prediction of Road Accident Severity. *Studies in Informatics and Control*, 23(1), pp. 41-52.

[24] STELMACH, A., 2012. Neural Model of the Aircraft Landing Phase. *Archives of Transport*, 24(2), pp. 249-258.

[25] TESEMA, T. B., ABRAHAM, A., GROSAN, C., 2005. Rule mining and classification of road traffic accidents using adaptive regression trees. *International Journal of Simulation*, 6(10), pp. 80-94.

[26] VAPNIK, V. N., 1999. An overview of statistical learning theory. *IEEE Transactions on Neural Networks*, 10(5), pp. 988-999.

[27] XIE, Y., LORD, D., ZHANG, Y., 2007. Predicting motor vehicle collisions using Bayesian neural network models: An empirical analysis. *Accident Analysis & Prevention*, 39(5), pp. 922-933.

[28] XU, C., WANG, W., LIU, P., 2013. A genetic programming model for real-time crash prediction on freeways. *IEEE Transactions on Intelligent Transportation Systems*, 14(2), pp. 574-586.

[29] YANG, J., ZHAO, J., 2013. Road Traffic Safety Prediction Based on Improved SVM. *ICTE 2013*, pp. 107-114.

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

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

[32] YASDI, R., 1991. Prediction of road traffic using a neural network approach. *Neural computing & applications*, 8(2), pp. 135-142.

[33] YU, B., YANG, Z. Z., SUN, X. S., YAO, B. Z., ZENG, Q. C., JEPPESEN, E., 2011. Parallel Genetic Algorithm in Bus Route Headway Optimization. *Applied Soft Computing*, 11(8), pp. 5081-5091.

[34] YU, B., YANG, Z. Z., YAO, J. B., 2010. Genetic Algorithm For Bus Frequency Optimization. *Journal of Transportation Engineering*, 136(6), pp. 576-583.

[35] YU, B., ZHU, H. B.,CAI, W. J., MA, N., YAO, B. Z., 2013. Two-phase Optimization Approach to Transit Hub Location--the Case of Dalian. *Journal of Transport Geography*, 33, pp. 62-71.

[36] ZHANG, C., YAN, X., MA, L.,AN, M., 2014. Crash Prediction and Risk Evaluation Based on Traffic Analysis Zones. *Mathematical Problems in Engineering*, 2014(2014), pp. 9.

[37] ZONG, F., JIA, H. F., PAN, X., WU, Y., 2013b. Prediction of Commuter's Daily Time Allocation. *PROMET - Traffic&Transportation*, 25(5), pp. 445-455.

[38] ZONG, F., XU, H. G., ZHANG, H. Y., 2013a. Prediction for Traffic Accident Severity: Comparing the Bayesian Network and Regression Models. *Mathematical Problems in Engineering*, 2013(2013), pp. 9.