

# Research on the Prediction Model for the Security Situation of Metro Station Based on PSO/SVM

Yong Qin<sup>1</sup>, Zhenyu Zhang<sup>1</sup>, Bo Chen<sup>1,2</sup>, Zongyi Xing<sup>3</sup>, Jing Liu<sup>4</sup>, Jun Li<sup>4</sup>

<sup>1</sup>State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China; <sup>2</sup>Tianjin Metro Operation Co., Ltd., Tianjin, China; <sup>3</sup>School of Automation, Nanjing University of Science and Technology, Nanjing, China; <sup>4</sup>Guangzhou Metro Corporation, Guangzhou, China.  
Email: [yqin@bjtu.edu.cn](mailto:yqin@bjtu.edu.cn)

Received May 25<sup>th</sup>, 2013; revised June 25<sup>th</sup>, 2013; accepted July 2<sup>nd</sup>, 2013

Copyright © 2013 Yong Qin *et al.* This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## ABSTRACT

Security situation awareness is a new technology about security. This paper brings it to the assessment of security situation of metro station which serves as a new way to secure the security of passengers as well as the operation of the metro station. This paper sets up an index system for assessing the security situation awareness and makes a prediction model for the security situation of metro station based on PSO/SVM after doing lots of researches and analyses. Furthermore, through case studies, we find that the model has high accuracy and ability to accurately predict the security situation of metro station in the future and a certain practical value.

**Keywords:** Situation Awareness; Situation Prediction; Index System; PSO; SVM

## 1. Introduction

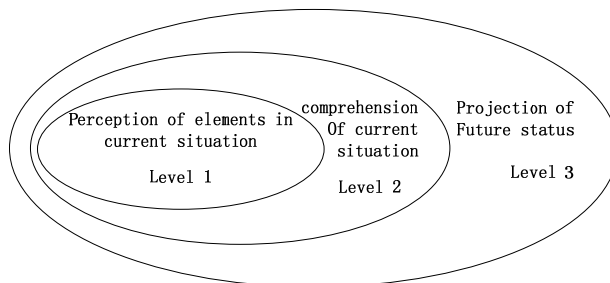
Metro station is a place for passengers to wait for the train which is of many characters such as profession on operation, complication on technologies and equipments that it applies for, and the mass on the number of passengers. Once security incident occurs, it will end with great loss both in life and personal injury and property. Therefore, it is of great importance to assure the security of metro station operation [1]. This paper brings security situation awareness to the assessment of security situation of metro station to get a comprehensive assessment about the security situation of metro station so as to reflect the situation objectively and take measures in time to minimize security risks. It can be served as a new way to secure the security of passengers as well as the operation of the metro station.

The concept of situation awareness originates from the research of ergonomics in the space field reflecting the process in which the pilots observe, analyze and make decision. Now it is widely used in many fields such as network security, military and space [2]. Endsley defines situation awareness as the perception of elements in current situation, comprehension of current situation and projection of future status based on certain condition of time and space [3,4]. The three levels of situation aware-

ness can be analyzed in **Figure 1**.

The security situation awareness of metro station can be analyzed as an extendedness of situation awareness of space in the metro system. At present, the research on security situation awareness of metro station is still in infancy, so it does not get a unified and comprehensive definition. Therefore, this paper tries to give an appropriate definition of it.

The security situation awareness of metro station is the perception, comprehension and projection of future status of the variants which affect the security of passengers and the metro station operation from the perspective of the affective elements to the metro station operation. We can get the security situation value of the past and the



**Figure 1. The three levels of situation awareness.**

present through the security situation comprehension model, then forecast the security situation value of future so as to help the manager of metro station master the status of the development of security situation.

The security situation value of metro station which can reflect the operation situation of station security is the quantitative expression of metro station security situation. The security situation value changes as the rate, number and places of the security accidents change. We can take appropriate actions and make protective plan after we can accurately master the security situation of metro station.

The prediction of security situation of metro station, on the basis of security situation value in the known period in which we supervise it, forecasts the security situation value of future status so as to assure the security situation of metro station in the future. Therefore, it is where the research of security situation awareness can go to forecast the security situation value in the future by setting up an appropriate model for security situation awareness.

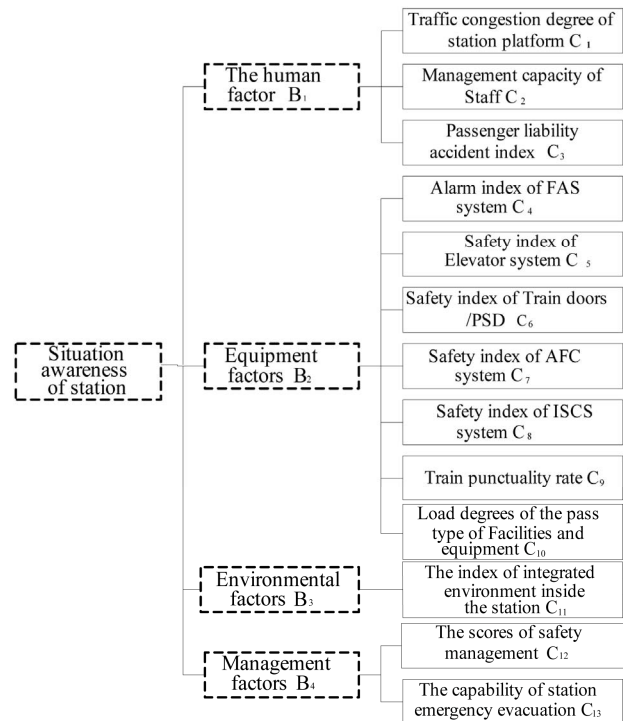
The key point of the research of security situation prediction lies on how to use artificial Intelligence accounting to set up a prediction model with high accuracy. Typical prediction methods include time order prediction, neural network prediction, grey model and SVM [5,6]. SVM is the hot pot among the researches in security situation prediction because it is of many advantages such as being fast in convergence rate, small in absolute error, strong in the ability against over fitting and accurate in the prediction of developing tendency in security situation [4]. However, a problem that we are blind in selecting the training parameters will appear if we only use SVM as prediction model [6].

Therefore, this paper comes up with a prediction model of security situation composed of PSO/SVM. It can optimize the training parameters to use the method of PSO, so we can get developed SVM prediction model. Furthermore it can be used to predict the security situation with high accuracy.

## 2. The Index System of Assessing Security Situation Awareness

The change of Security situation of metro station is the result of interaction and interrelationship of the four elements-humans, equipment, environment and management. Basing on the survey of Guangzhou metro, analyzing the statistics of the injured passengers and analyzing of the elements which affect the security, this paper sets up index system of assessing security situation awareness of metro station from the basic connotation of metro station security.

**Figure 2** is the index system of subway station security situation awareness.



**Figure 2. The evaluation indexes system.**

The index system of assessing security situation awareness of metro station makes an assessment about the insecure situation of human’s behavior, environment and management. It can reflect and describe accurately the security situation and its developing tendency via assessing the security of the traveling of passengers so as to take action to keep off the origin of danger and ensure the security of the station operation and passengers.

The paper analyzes influence of indicators on the station’s security situation, and on the basis of field investigation and literature analysis, deduces the formula of each index, as shown in **Table 1**. Quantitative values of each index calculated can objectively reflect the station’s security situation; increase the subway station security situation assessment result is scientific and accurate.

## 3. Prediction Model for Security Situation of Metro Station

### 3.1. Basic Principles of SVM

SVM [7,8] was come up with by Vapnik and Cortes in 1995. SVM is originally used to do some classifying. In order to extend to regression estimation, we can introduce insensitive loss function  $\epsilon$ . The basic idea is through a nonlinear mapping  $\Phi$  training set data  $x$  mapped into a high dimensional linear feature space  $G$ , and construct regression estimates function in this space.

SVM similar to a neural network, the output is a linear combination of intermediate nodes, each intermediate

**Table 1. The calculation formula of each index.**

The meaning of indicators	Calculation formula & parameters
<p><math>C_1</math> is the ratio of the actual assembling on the station platform and the actual area of the station platform during the peak times of passenger flow</p>	$C_1 = \frac{(Q_1 \cdot T_1 + Q_2 \cdot T_2)}{S} \cdot \phi$ <p><math>Q_1</math>—The pitted people of per second; <math>Q_2</math>—The Outbound people of per second;  <math>S</math>—Actual area of the station platform, <math>m^2</math>; <math>T_1</math>—Train arrived in time interval, <math>s</math>;  <math>T_2</math>—The longest travel time from the platform to the station hall, <math>s</math>;  <math>\Phi</math>—The platform uneven coefficient of passenger flow density</p>
<p><math>C_2</math> is the comprehensive value of staff's performance evaluation results during statistical period</p>	$\left\{ \begin{array}{l} C_2 = \sum_{i=1}^4 \alpha_i d_i \\ d_i = \frac{\sum_{j=1}^m d_j}{m} \end{array} \right.$ <p><math>d_i</math>—The average performance evaluation results of the <math>i</math> class staff;  <math>\alpha_i</math>—Importance degree of the <math>i</math> class staff; <math>m</math>—The number of the <math>i</math> class staff;  <math>d</math>—The performance evaluation results score of the <math>i</math>-th class <math>j</math>-th staff</p>
<p><math>C_3</math> is the number of passenger injury accidents caused by passengers of their responsibilities during statistical period</p>	$C_3 = N$ <p><math>N</math>—The number of passenger injury accidents caused by passengers of their responsibilities</p>
<p><math>C_4</math> is the number of FAS system false alarm(omission) during statistical period</p>	$C_4 = M + N$ <p><math>M</math>—The number of FAS system false alarm; <math>N</math>—The number of FAS system omission alarm</p>
<p><math>C_5</math> is the ratio of the equipment failure maintenance time of elevator system and the program run-time during statistical period</p>	$C_5 = \frac{\sum_{i=1}^n \frac{T}{T_{plan}}}{n} \times 100\%$ <p><math>n</math>—The number of vertical elevators and escalators;  <math>T</math>—The equipment failure repair time of <math>i</math>-th vertical elevator or escalator;  <math>T_{plan}</math>—The program run-time of <math>i</math>-th vertical elevator or escalator</p>
<p><math>C_6</math> is the ratio of the train doors/PSD non-normal switching number of times and the total number of normal switch during statistical period</p>	$C_6 = \frac{\sum_{i=1}^n \frac{N_{doors/PSD}}{Alt_{doors/PSD}}}{n} \times 100\%$ <p><math>N_{doors/PSD}</math>—The <math>i</math>-th train doors/PSD's non-normal switching number of times;  <math>Alt_{doors/PSD}</math>—The <math>i</math>-th train doors/PSD's total number of normal switch;  <math>n</math>—The number of train doors/PSD</p>
<p><math>C_7</math> is the ratio of the equipment failure maintenance time of AFC system and the program run-time during statistical period</p>	$C_7 = \frac{\sum_{i=1}^n \frac{T}{T_{plan}}}{n} \times 100\%$ <p><math>n</math>—The number of floodgate machine and ticket machines;  <math>T</math>—The equipment failure repair time of <math>i</math>-th machines;  <math>T_{plan}</math>—The program run-time of <math>i</math>-th machines</p>
<p><math>C_8</math> is the total equipment alarms number of ISCS and integrated subsystems during statistical period</p>	$C_8 = \sum_{i=1}^n M_i$ <p><math>n</math>—The number of ISCS and integrated subsystems;  <math>M_i</math>—The <math>i</math>-th system's equipment alarms number</p>
<p><math>C_9</math> is the ratio of the number of times of train punctuality arrival and all number of times of the opening line</p>	$C_9 = \frac{N_{on\ time}}{N} \times 100\%$ <p><math>N_{on\ time}</math>—The number of times of train punctuality arrival;  <math>N</math>—All number of times of the opening line</p>
<p><math>C_{10}</math> is the load degrees of the pass type of facilities and equipment during the peak times of passenger flow</p>	$\left\{ \begin{array}{l} C_{10} = \sum_{i=1}^n \lambda_i \cdot S_i \\ S_i = \frac{Q_i}{C_i} \end{array} \right.$ <p><math>S_i</math>—The load degrees of the <math>i</math>-th equipment; <math>\lambda_i</math>—The weight of the <math>i</math>-th equipment;  <math>Q_i</math>—The actual passenger traffic of the <math>i</math>-th equipment;  <math>C_i</math>—The traffic capacity of the <math>i</math>-th equipment;  <math>n</math>—The number of the pass type of facilities and equipment</p>

Continued

$C_{11}$  is the comprehensive evaluation for the environment inside the station during statistical period

$$C_{11} = \sqrt{I_{\max} \cdot I_y} = \sqrt{\max \left[ \frac{C_1}{S_1}, \frac{C_2}{S_2}, \frac{C_3}{S_3} \right] \cdot \frac{1}{3} \sum_{i=1}^3 \frac{C_i}{S_i}}$$

$I_{\max}$ —Maximum mass sub-indices;  $I_y$ —Average quality sub-indices;  $c_i$ —The  $i$ -th factor's measurement values;  $S_i$ —The  $i$ -th factor's standard values;  $i$  are the temperature, humidity and  $CO_2$

$C_{12}$  is the comprehensive evaluation for the safety management institutions, safety production responsibility system, safety education, accident rescue etc.

$$C_{12} = \frac{\sum_{i=1}^n F_i - F_{\max} - F_{\min}}{n - 2} \times 100\%$$

$F_i$ —The  $i$ -th expert's evaluation scores of the safety management capability;  $F_{\max}$ —The maximum evaluation score;  $F_{\min}$ —The minimum evaluation score;  $n$ —The number of expert

$C_{13}$  is the accident evacuation time of the platform layer during the peak times of passenger flow.

$$C_{13} = 1 + \frac{Q_1 + Q_2}{0.9[A_1(N-1) + A_2B]}$$

1—The response time of the people, min;  $Q_1$ —Train passenger numbers;  $Q_2$ —The total number of waiting passengers and staff on the station platform;  $A_1$ —Through capacity of escalator, people/min·m;  $A_2$ —Through capacity of stairway, people/min·m;  $N$ —The number of escalator;  $B$ —The total width of the stairway, m; 0.9 is reduction factor

node corresponds to a support vector, and its weight value is the corresponding Lagrange multiplier, shown in **Figure 3**.

For a group of contiguous sequence  $x(i)$ , the

$$x(i), x(i+1), x(i+2), \dots, x(k+i-1)$$

as an input vector,  $x(k+i)$  as the output value, in order to establish the training sample, and SVM regression model is described as:

$$x(k+i) = f(x(i), x(i+1), \dots, x(k+i-1)) \quad (1)$$

where,  $k$  is the dimension of input vector. Training samples shown in **Table 2**.

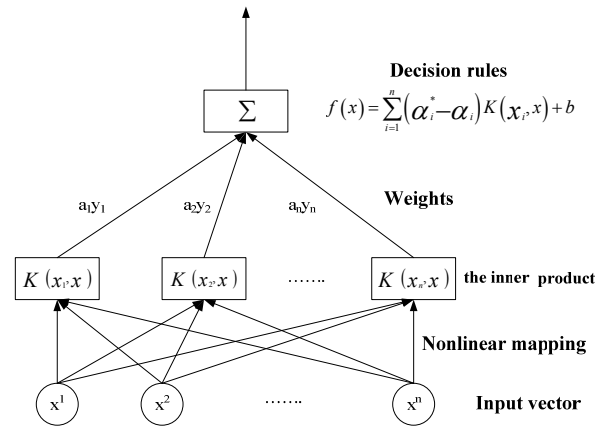
**3.2. Basic Principles of PSO**

PSO was proposed by Dr. Eberhart and Dr. Kennedy, originated from the study on the behavior of birds foraging [9,10]. Assuming a D-dimensional search space, there are  $N$  particles forming a population in which the position of each particle is represented as a D-dimensional vector with  $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})$  representation; particles' flight speed recorded as  $V_i = (v_{i1}, v_{i2}, \dots, v_{iD})$ ; particle  $i$  have been found as the optimal position so far, denoted  $P_{\text{best}}$ ; the best position of whole swarm have been searched so far, denoted by  $g_{\text{best}}$ ; after finding the two positions, the particles update their speed and position according to the following formula:

$$v_{id}^{k+1} = w * v_{id}^k + c_1 r_1 (p_{id}^k - x_{id}^k) + c_2 r_2 (p_{gd}^k - x_{id}^k) \quad (2)$$

$$x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1} \quad (3)$$

where,  $v_{id}^{k+1}$  is the flight speed of  $i$ -th particle in  $(k+1)$  generations;  $x_{id}^{k+1}$  is the position of  $i$ -th particle in  $(k+1)$  generations;  $p_{id}^k$  is the best position of  $i$ -th particle to



**Figure 3. The schematic diagram of SVM.**

**Table 2. Training samples.**

Input	Output
$x(1), x(2), x(3), \dots, x(k)$	$x(k+1)$
$x(2), x(3), x(4), \dots, x(k+1)$	$x(k+2)$
...	...
$x(n-k), x(n-k+1), x(n-k+2), \dots, x(n-1)$	$x(k+n)$

$k$ -generation;  $p_{gd}^k$  is the best position populations to  $k$ -generation;  $p_{gd}^k - x_{id}^k$  is individual cognition;  $p_{gd}^k - x_{id}^k$  is the population cognition;  $w$  is the inertia weight;  $v_{id}^k$  is the velocity of the particle;  $c_1$  and  $c_2$  are learning factors;  $r_1$  and  $r_2$  is the random number of uniformly distribution in  $[0,1]$ ;  $i = 1, 2, \dots, N$ .

**3.3. Security Situation Prediction Model**

Based on PSO/SVM metro station security situation prediction model is divided into three modules: 1) data pre-processing module is that experimental data are normal-

ized to obtain the required model for training and testing data sets; 2) PSO parameter optimization module is the use of the PSO algorithm to optimize the SVM parameters, then the optimal parameters pass to the SVM; 3) SVM prediction module using the training data set has been training to obtain prediction model, and use the test data set for prediction. The structure of prediction model has been shown in **Figure 4**.

In PSO parameter optimization module, we need to determine the fitness function. PSO algorithm used fitness value to evaluate the merits of an individual or population in the search process of evolution, and as the basis of the particle velocity and position changes, it gradually evolved to the optimal solution. In this paper, the mean square error (MSE) is the fitness function:

$$MSE(y_i, y_i^*) = \frac{1}{N} \sum_{i=1}^N (y_i - y_i^*)^2 \quad (4)$$

where  $y$  is the actual value of the training sample,  $y_i^*$  is the predicted value of the model,  $N$  is the number of data samples. The MSE is smaller, which means the higher prediction accuracy, the predicted value closer to the target value.

In SVM prediction module, we need to determine the kernel function. Introducing of kernel function enables inner product operation of high-dimensional space implement by input space, to avoid the curse of dimensionality. Since the radial basis function only needs to determine a variable, compared with other kernel functions easy to determine, therefore, the use of Gaussian radial basis kernel function, SVM can make optimal performance. RBF kernel function is described as:

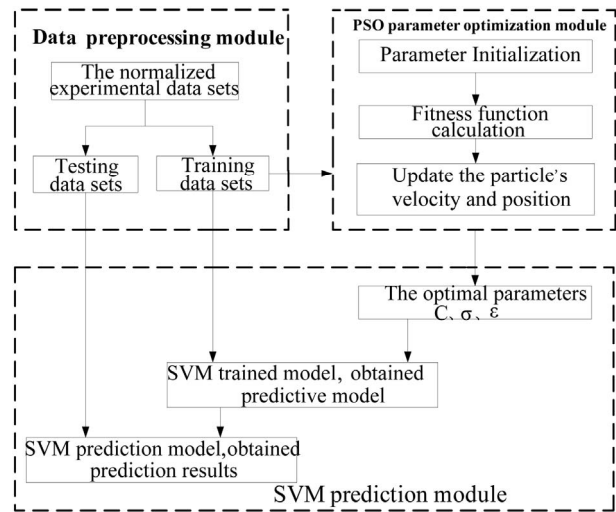
$$K(x_i, x) = \exp\left(-\frac{|x_i - x|}{\sigma^2}\right) \quad (5)$$

where,  $\sigma$  is RBF kernel parameters.

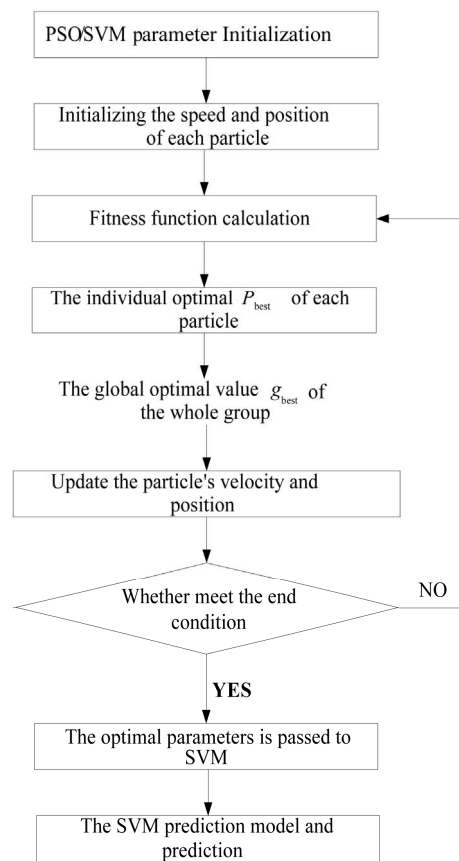
Based on the above prediction model, security situation prediction algorithm steps based on PSO/SVM will be shown in **Figure 5**.

#### 4. Examples and Analysis

To verify the validity of PSO/SVM security situation prediction model, Guangzhou Chen clan academy station, for example, according to the data requirements for each index in the evaluation index system, we got year 2011 data and quantified each index as the original experimental data sample. Using the literature [11] multidimensional contact number method to establish the security situation understanding model to calculate the value of Chen clan academy station 365 days' security situation, using the formula (6) on the security situation value data are normalized, normalized security trend values shown in **Figure 6**. Taking the dimension input vector is 7, building a experimental sample set whose capacity is 358



**Figure 4. The structure of prediction model.**



**Figure 5. Security situation prediction algorithm steps based on PSO/SVM.**

groups, where the data of the first 280 groups as the training sample set, the data of the last 78 groups as the test sample set.

$$x_i = \frac{y_i - \min(y_i)}{\max(y_i) - \min(y_i)} \quad (6)$$

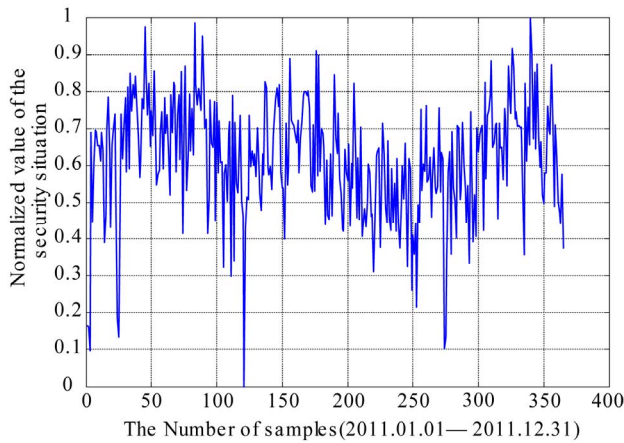


Figure 6. Normalized security trend values.

In order to measure the accuracy of the prediction model, this paper selected mean square error as predictive accuracy of fitness function evaluation model. However, due to the small mean square error value, it is difficult to visually give the difference between the actual value of the security situation and the predicted value of PSO/SVM security situation prediction model, and analysis the correlation through linear regression analysis. Therefore, using the correlation coefficient  $R$  to further measure the curve fitting and linear regression analysis between the predicted value and the actual value of the prediction model security situation:

$$R(y_i, y_i^*) = \frac{\sum_{i=1}^N (y_i - \bar{y}_i)(y_i^* - \bar{y}_i^*)}{\sqrt{\sum_{i=1}^N (y_i - \bar{y}_i)^2 \sum_{i=1}^N (y_i^* - \bar{y}_i^*)^2}} \quad (7)$$

where,  $\bar{y}_i$  is the average of the test sample,  $\bar{y}_i^*$  is the average of the predicted value.  $R$ -value is closer to 1, which means that the model prediction has higher accuracy, the closer to the actual system.

Set PSO parameters are: population size is 40; inertia weight initial  $w_{max} = 0.9$ ; inertia weight final value  $w_{min} = 0.4$ ; learning factor  $c_1 = c_2 = 2$ ; maximum velocity  $V_{max} = 5$ ; maximum speed and position limits scaling factor  $k = 0.6$ ; maximum evolution generation  $T_{max} = 200$ . Training process adaptation curve shown in Figure 7.

Optimal parameters through training  $C, \sigma, \varepsilon$ , shown in Table 3.

Analyzed follows: PSO/SVM security situation prediction model training MSE was  $2.428 \times 10^{-3}$ , testing MSE was  $6.172 \times 10^{-4}$ . Figure 8 shows the comparison between the predicted and actual value of the security situation model. We can see, it has good relevance between the predicted and actual value. Both in the training data or in the test data can basically fit on the actual value. Figure 9 shows us the correlation curve of security situation prediction model predicted value and the

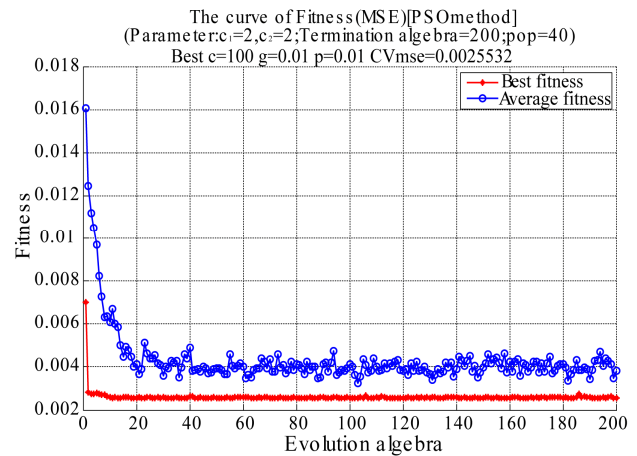
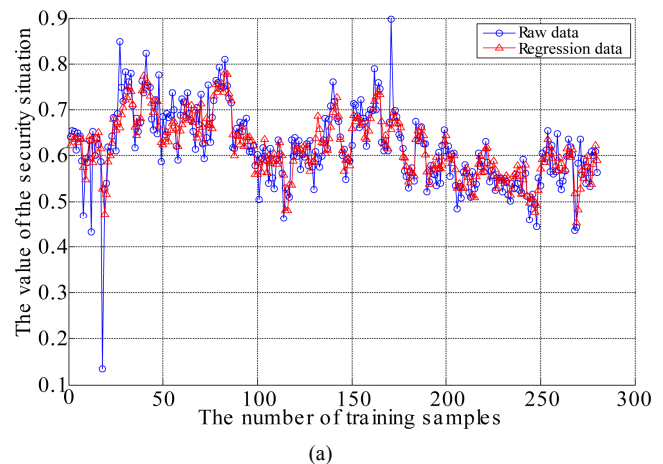
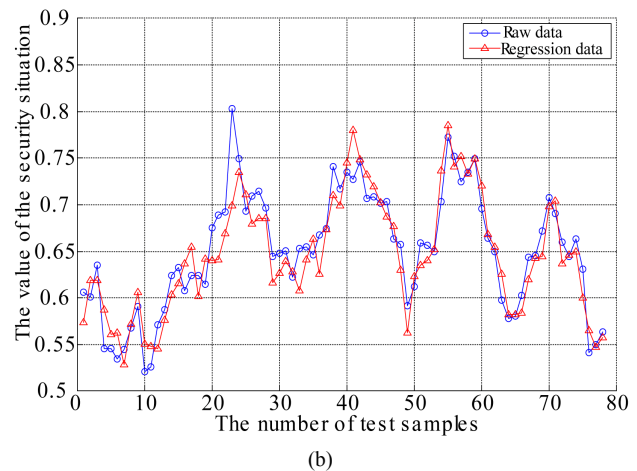


Figure 7. Training process adaptation curve.



(a)

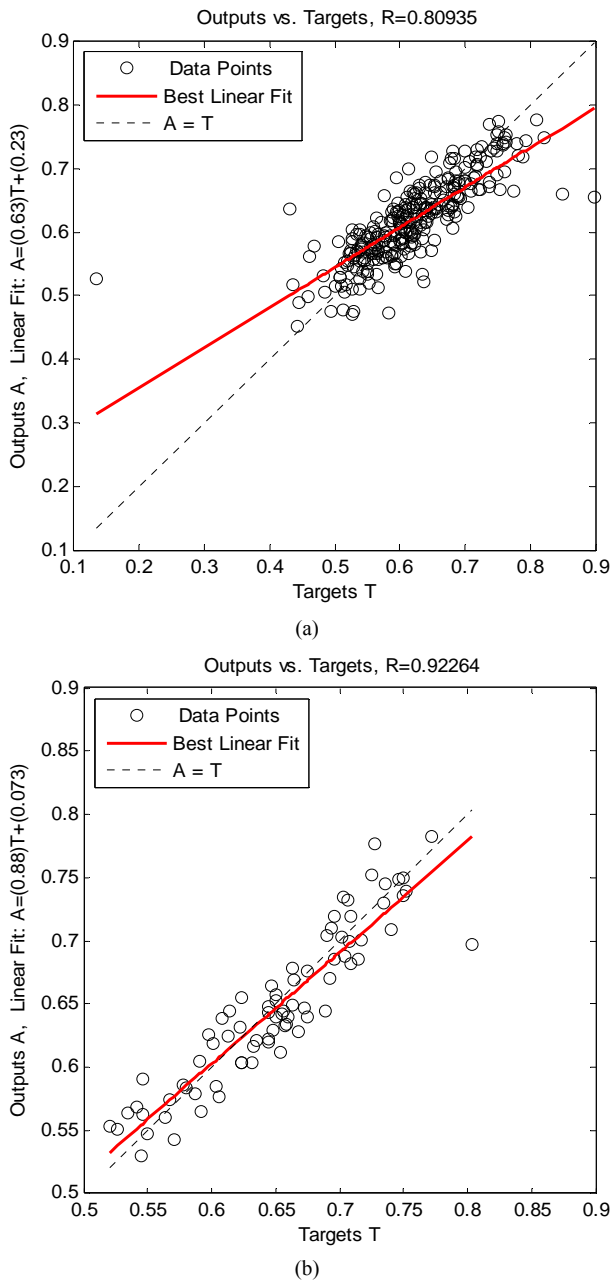


(b)

Figure 8. The comparison between the predicted and actual value. (a) The training set predictive data comparison; (b) The testing set predictive data comparison.

Table 3. Optimal parameters.

	$C$	$\sigma$	$\varepsilon$
Optimal parameters	100	0.01	0.01



**Figure 9. The correlation curve of predicted value and the actual value. (a) Correlation of the training set; (b) Correlation of the testing set.**

actual value, the correlation coefficient  $R$  in the training is 0.80935, the test is 0.92264, showing that the model predicted values had a higher degree of curve fitting on the actual values. **Table 4** gives us the results of prediction model training and test data.

### 5. Conclusions

1) Based on the Guangzhou Metro research analysis, statistical analysis of passengers' injury and security factors, this paper established the evaluation system of

**Table 4. The results of prediction model training and test data.**

	MSE	R
Training data	$2.428 \times 10^{-3}$	0.80935
Test data	$6.172 \times 10^{-4}$	0.92264

metro station's security situation awareness.

2) This paper established metro station security situation prediction model based on PSO/SVM, and through case analysis, it verified the model having high prediction accuracy, which has some practical value. The next step should be to enhance comparative study of prediction algorithm.

Security situation awareness in the field of urban rail transportation applications is still in infancy, and it still needs further study in order to better solve practical problems.

### 6. Acknowledgements

The authors would like to express their thanks to the editor and anonymous reviewers for their help in revising the manuscript. This research is sponsored by National Key Technology R&D Program of China (2011 BAG-01B02).

### REFERENCES

- [1] Z. H. Wang, "Research Oil Operation Security Risk Assessment of Metro Station," Master Thesis, Beijing Jiaotong University, Beijing, 2009.
- [2] H. Q. Wang, J. B. Lai, *et al.*, "Survey of Network Situation Awareness System," *Computer Science*, Vol. 10, No. 2, 2006, pp. 5-10.
- [3] M. R. Endsley, "Design and Evaluation for Situation Awareness Enhancement," *Human Factors Society 32nd Annual Meeting*, Santa Monica, 1988, pp. 97-101.
- [4] M. R. Endsley, R. Sollenberger and E. Stein, "Situation Awareness: A Comparison of Measures," *Proceedings of the Human Performance, Situation Awareness and Automation: User Centered Design for the New Millennium*, Savannah, 2000.
- [5] R. R. Xi, X. C. Yun, *et al.*, "Research Survey of Network Security Situation Awareness," *Journal of Computer Applications*, Vol. 32, No. 1, 2012, pp. 1-4.
- [6] K. L. Gao, J. M. Liu, *et al.*, "A Hybrid Security Situation Prediction Model for Information Network Based on Support Vector Machine and Particle Swarm Optimization," *Power System Technology*, Vol. 35, No. 4, 2011, pp. 176-182.
- [7] Z. H. Han and X. X. Zhu, "Selection of Training Sample Length in Support Vector Regression Based on Information Entropy," *Proceedings of the CSEE*, Vol. 30, No. 20, 2010, pp. 112-116.
- [8] J. Kennedy and R. C. Eberhart, "Particle Swarm Optimization," *IEEE International Conference on Neural Net-*

*works, IV*, Piscataway, 1995, pp. 1942-1948.

- [9] W. B. Langdon and R. Poli, "Evolving Problems to Learn about Particle Swarm and Other Optimizers," *IEEE Transactions on Evolutionary Computation*, Vol. 11, No. 5, 2005, pp. 81-88.
- [10] S. Garnier, J. Gautrais and G. Theraulaz, "The Biological Principles of Swarm Intelligence," *Swarm Intelligence*, Vol. 30, No. 1, 2007, pp. 3-31.  
<http://dx.doi.org/10.1007/s11721-007-0004-y>
- [11] Q.-Z. Hu and J. Wu, "A Monitoring and Controlling Model for Urban Traffic Security State Based on Multi-Dimension Connection Number," *China Security Science Journal*, Vol. 21, No. 10, 2011, pp. 16-22.