Safety Risk Evaluation of Cross-Railway Highway Project Based on SVM

Zhang Yong[1], Zhang Fang[2], Yin Xiang[3]

[1,,2,3] *School of management*
Xian University of Architecture and Technology, China

Abstract **— It is inevitable to make the new-built highways under-pass the existing railroad by virtue of high-speed development of the transportation. Since the relevant surveys on the safety risk evaluation of this kind of the project are lacked at this current stage as well as the limited referential sample data, the survey is to research the factors of safety risk and the self-characteristics as well as to build up the safety risk evaluation model by means of support vector machine plus the analysis with the help of the project example. It turns out that the analyzed result by means of the model has the propinquity with the expected result by experts, as means the model with the limited samples is characterized of improving the objective correctness of the evaluation result in order to supply a scientific and Quantitative method for the safety evaluation of this kind of projects.**

Keywords - Cross-railway highway project;Support vector machine; Safety risk evaluation

I. Introduction

Since implementing the reform and opening-up policy, China's economic has developed rapidly. To the end of 2010, China's economy aggregate surpassed Japan and has become the world's second economic giant. The traffic construction made a significant contribution to such a fast development for the economic. By 2011, with a total mileage of highway in our country has reached 4.1046 million kilometers, which occupies the world's second place (LiuChong and ZhouLiAn,2014). In the face of the rapid development of highway, it is common that more and more newly built highways will cross the existing railways. Cross-railway highway project not only needs to ensure the safety or the operation of existing railways, but also need to complete constructing the new highway smoothly. To ensure the safety of constructing new highway and to ensure the safety of operating existing railway, it is more risky and complicated to implement the project, which has made it is particularly important to do the sufficient risk assessments and preventive measures at the early stage of the construction.

At the present stage, the security issues of Cross-railway highway project are still in the research and summarizing experience stage (A.M. Hefny, H.C. Chua, J.Zhao., 2004; QianQiHu and RongXiaoLi, 2008.) Most projects are taking in-process controls or remedial measures for risks which can not have sufficient prevention and control effect on reducing the risks of safety accidents in the construction process. In order to accurately estimate the probability of accidents and emphasis which are needed to be focused on, and to take the most reasonable and effective precautionary measures during the construction of Cross-railway highway project, this article evaluates the safety of the project risk by the Support Vector Machine (SVM) model. By identifying safety risk factors existing in the engineering, and based on the principles of the construction of index system and the actual situation of the project, the safety risk assessment index system of Cross-railway highway project will be set up. By using Support Vector Machine (SVM) model to evaluate the risk factors so as to achieve the safety accident risk for the purpose of a beforehand control. This paper builds the safety risk assessment model by using the support vector machine (SVM) in a highway engineering example, which as well provides scientific basis for the safety risk assessment of such projects

II. PROJECT SUMMARY

A. Project summary

This article selected engineering for chow Tai second phase of highway engineering as the project. The project began in chow tai highway and even Huo highway cross reserved hub interchange in Xiaoshilou, to the north it crosses Longhai railway and to the south it crosses Beijing-Kowloon railway, which is in total 26.991 km. This is a very important line in the "11th five-year plan" of Henan province highway network planning. Due to the project crosses two railway lines, the complexity and the risk in the process of construction is much higher than others. The example of Cross-railway highway project all using Vertical and horizontal lifting beam overhead reinforcement technique, existing subgrade strengthen using grouting technology. In the process of construction uses jacking construction method underpass existing railway, in the construction site prefabricates two-hole continuous box culvert which scale is 2-14.6m.

B. Analysis of risk factors

Compared with other highway projects, roads cross project has significant characteristics that are strong elusive, process-complex, dynamic, risky and so on (ChenAiRong and RuanXin，2005.). 《Dangerous and Harmful Factors Classification and Code in the Production Process》 classifies the risky factors of engineering construction into four aspects which are people, objects, environment and management. Human factor refers to the potential safety hazard comes from human itself or man-made hidden danger ; Object factor refers to the safety hazard from construction materials, machinery and equipment used in the process of engineering; Environmental factors refers to the engineering project is affected by environment; Management factors refers to the safety hazard caused by neglectful management. In view of the particularity and complexity of this project, this article will build up the evaluation index system of the risk of the project from the people, objects, environment, management and technology after considering the effect of technology in the construction process.

C. The Establishment of Evaluation Index System

On the basis of the safety accident cause theory and according to the engineering actual situation, the safety risk assessment index system of the new-built highways under-pass the existing railroad project sets up following to the principle of the establishment of evaluation index, referring to the relevant standards and norms and combined with previous similar engineering examples and expert experience and the risk index system construction principles, which is shown in Figure 1.

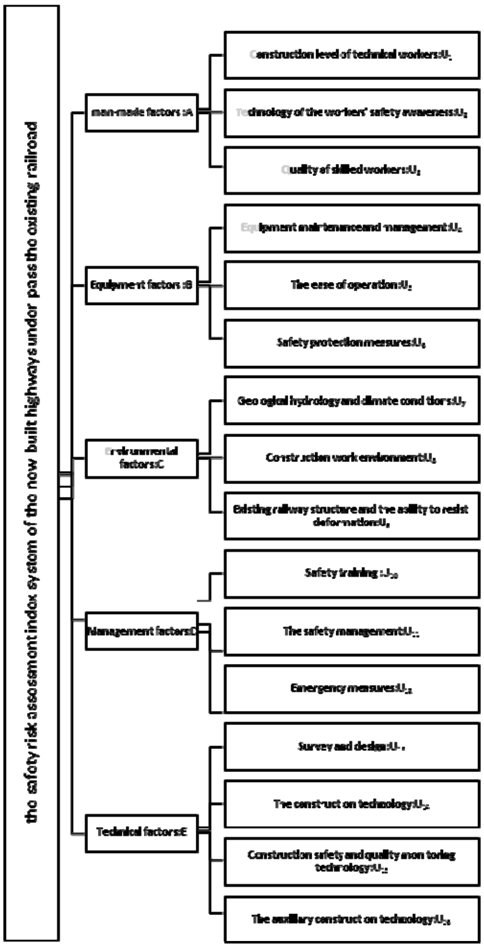

Figure.1:The Safety Risk Assessment Index System.

III. SUPPORT VECTOR MACHINE (SVM) MODELL FOR EVALUATION

A. The Basic Principle of Support Vector Machine (SVM)

SVM (Support Vector Machine, SVM) is a new type of structural Machine learning algorithms based on statistics. In solving the problem of classification and regression problems, the classification and regression problems will be turned into quadratic programming problems. Problems such as small sample, nonlinear and local minimum point can be better solved by solving a linear constrained quadratic programming problem to get the overall optimal solution. This method makes the classification of study objects more accurate and overcomes a lot of shortcomings of neural network, such as the uncertainty of structure and weights; and limited sample (Hou Yan Juan and Zhang Ding Li, 2011). The advantages of the model is not only to minimize the risk of experience, but also minimize the incredible scope, thus the principle of structural risk minimization can be better realized. Support vector machine (SVM) regression method uses nonlinear transformation to convert the original samples in low dimensional space into high dimension space, and then to do the regression analysis(FanXinWei,2003.).

B. Building the Safety Risk Assessment Model

Since the safety risk assessment of cross-railway highway project construction is a complex nonlinear regression problem, this paper puts forward a kind of security risk assessment model based on support vector machine (SVM) particularly for this kind of project. According to the risk assessment index system set up in the front and the collected related data of engineering examples, the data are processed and conducted as input vector. Next, select appropriate parameters and kernel function training and use the established SVM regression model to evaluate the index system. Afterwards, determine whether the model is effective for the safety risk assessment of cross-railway highway project according to the evaluation results or not. The evaluation results can be used to explore the accuracy of risk assessment for the similar projects.

N training samples are given, and try to find a function to make it can infer any input x by y. On the application of support vector machine (SVM) on regression issue, and usually make a less sensitive function defined as the loss function L, namely the deviation of a given function centralized functions and up to epsilon. To the sample not in epsilon range， then by introducing slack variable delta with constraints. At this time the solving of regression issue is transformed into:

$$\min_{w,\delta_1,\delta_2,b} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\delta_1,\delta_2)$$

$$s.t.\begin{cases} w\cdot\emptyset(x_i)+b-y_i\leq\varepsilon+\delta_1 \\ y_i-w\cdot\emptyset(x_i)-b\leq\varepsilon+\delta_2 \\ \delta_1,\delta_2\geq 0, i=1,2,3\cdots,n \end{cases} \quad (1)$$

Quoting the dual form of Lagrange function, the solution is:

$$\alpha^* = \left(\overline{\alpha_1},\overline{\alpha_1^*}\cdots\overline{\alpha_n},\overline{\alpha_n^*}\right)^T$$

$$b^* = y_j - \sum_{i=1}^{n}\left(\overline{\alpha_i^*} - \overline{\alpha_1}\right)K\left(x_i \cdot x_j\right)+\varepsilon$$

When income is not 0, its corresponding sample is support vector.

By learning and training the regression decision function is:

$$f(x)=\sum_{i=1}^{n}\left(\overline{\alpha_i^*} - \overline{\alpha_i}\right)K\left(x_i,\ \ x\right)+\bar{b} \quad (2)$$

C. Select Kernel Function

Kernel function is the generalization pattern which makes low generalization of vector to be mapped to high-dimensional space mode in the nonlinear support vector machine (SVM) model calculation process. The selection of kernel function determines the performance of support vector machine (SVM), which means that different kernel are in different performances for support vector machine (SVM) model and are in different results for the solutions of regression issues at the same time. Currently there are four main kinds of kernel function:

1.polynomial kernel：Kx,xi=[x,xi+1]q

2.adial basis function kernel：Kx,xi=exp-x-xi2σ2

3.sigmoid kernel：Kx,xi=tanh(kx,xi+c)

4. Fourier Kernel: Kx,xi = 1-q22(1-2qcosx-xi+q2

A lot of experiments on the kernel function proved that the radial basis kernel function is superior to all other study methods of support vector machine (SVM) based on other kernel function no matter in the aspect of regression or of classification and recognition. Since the sample data does not have prior knowledge in this case, and the radial basis kernel function are in strong interpolation ability and involved less parameters, the radial basis kernel function is selected as evaluation model (Deng Nai Yang and Tian Ying Jie, 2009; Wang Wei and GuoXiaoMing,2008).

IV. DATA COLLECTION AND PROCESS

A. Data Collection and Processing

The research of this article is conducted in the second phase of Chow Tai highway engineering. The author has researched ten blocks of this engineering, and has investigated each unit of this highway project with some relevant experts to know the cooperation and coordination between them. At the same time, the author has investigated engineering technology department and quality supervision department of Henan Dexing expressway co., LTD., which is responsible for the construction of this project. The author has investigated various factors in the evaluation system of the present situation and has carried on the questionnaire survey. According to the information gathered from the survey, following 10 sets of data come up by asking expert scoring for the scores of the factors of construction risk, the last line U shows the comprehensive safety risk value given by experts according to the engineering situation. As shown in TABLE I .

TABLE I ALL THE FACTORS IMPACT PROFILE

Sample data / influencing factor		1	2	3	4	5	6	7	8	9	10
A	U1	75	75	76.5	50	75	75	76	75	75	75
	U2	50	75	75	50	75	62.5	87.5	62	60	50
	U3	85	82	78	80	85	85	95	90	95	95
B	U4	85	90	75	50	86	75	75	50	86	86
	U5	80	85	90	90	85	85	100	100	95	95
	U6	85	85	75	75	75	62	86.5	75	82.5	75
C	U7	74	90	55	25	75	50	75	25	75	80
	U8	70	75	70	86	75	85	70	75	86	90
	U9	85	100	50	25	86	75	85	50	100	100
D	U10	100	100	87	75	86	80	100	75	100	90
	U11	80	87	75	50	50	75	85	50	100	75
	U12	75	75	72	50	75	62	60	50	85	75
E	U13	90	100	85	62	80	87	100	62	100	75
	U14	85	75	75	60	75	75	75	75	85	82
	U15	75	75	62	50	50	75	75	75	85	82
	U16	80	100	75	60	75	75	75	50	100	75
U		0.25	0.15	0.5	0.75	0.35	0.4	0.2	0.7	0.1	0.3

Due to the differentiation of various index as well as the dimension is not unified, it may lead to skew the results if putting the index directly into computing. Therefore, before doing the data analysis, data values of each index need to be processed to make them mapped into the unified scope. Maximum minimum value method of normalization method is used to process the raw data Normalization formula is as follows:

$$x = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{3}$$

Among them, the xi is a true value indicators, xmax, xmin are the maximum and minimum values of true indices, x is the data after normalization. After the normalization, the data are shown in TABLE II as follows.

TABLE II THE NORMALIZED DATA PROCESSING RESULTS

		1	2	3	4	5	6	7	8	9	10
A	U1	0.5	0	0.662	0.385	0.694	0.676	0.4	0.667	0.375	0.5
	U2	0	0	0.625	0.385	0.694	0.338	0.688	0.493	0	0
	U3	0.7	0.28	0.7	0.846	0.972	0.946	0.625	0.867	0.875	0.9
B	U4	0.7	0.6	0.625	0.385	1	0.676	0.375	0.333	0.65	0.72
	U5	0.6	0.4	1	1	0.972	0.946	1	1	0.875	0.9
	U6	0.7	0.4	0.625	0.769	0.694	0.324	0.662	0.667	0.563	0.5
C	U7	0.48	0.6	0.125	0	0.694	0	0.375	0	0.375	0.6
	U8	0.4	0	0.5	0.938	0.694	0.946	0.25	0.667	0.65	0.8
	U9	0.7	1	0	0	1	0.676	0.5	0.333	1	1
D	U10	1	1	0.925	0.769	1	0.811	1	0.667	1	0.8
	U11	0.6	0.48	0.625	0.385	0	0.676	0.5	0.333	1	0.5
	U12	0.5	0	0.55	0.385	0.694	0.324	0	0.333	0.625	0.5
E	U13	0.8	1	0.875	0.569	0.833	1	1	0.493	1	0.5
	U14	0.7	0	0.625	0.538	0.694	0.676	0.375	0.667	0.625	0.64
	U15	0.5	0	0.3	0.385	0	0.676	0.375	0.667	0.625	0.64
	U16	0.6	1	0.625	0.538	0.694	0.676	0.375	0.333	1	0.5
U		0.25	0.15	0.5	0.75	0.35	0.4	0.2	0.7	0.1	0.3

Rows in the table represent ten samples in the engineering, columns represent 16 assessment indicators. The last line represents comprehensive security risk value given by experts according to the engineering. This paper choose the former seven samples in the table as training set and choose last three samples as a validation set evaluation index data after the normalization are selected as the input samples.

B. Parameter Optimization and Validation

According to the selected radial basis kernel function, the main parameters of support vector machine model which need to be determined are kernel bandwidth g and penalty factor C, in this paper search optimizing parameter by step length grid search method that considering two parameters at the same time. First do the rough selection, select a g and a big area scope of C, and then calculated its accuracy according to set up the parameter search step length of training in each group.

To find the better area of parameter combination according to the results of the roughly select, and then to do the further subdivision of parameter values range in this area, at last, to train each set of parameters and then calculate the accuracy(YangLi,2010.).

Due to particularity of engineering of newly building highway crossing under the existing railway, samples which can be used for reference and collected are fewer, In order to improve the effectiveness of the model of support vector regression machine, this article selects LOO cross-validation method to improve the accuracy of the selected parameters. This is to take each sample data that are samples concentrated as test set, and to take the remaining samples as training set. Afterwards, establish model according to the accuracy of the model test set generated by choosing the optimal parameters.

V. DATA ANALYSIS

A. Software

Using support vector machine (SVM) to solve regression problem is a quadratic programming. The difficulty and calculation amount of it is large, so certain calculate software is needed. Using the MATLAB software and Libsvm software package to realize the safety risk assessment research of cross-railway highway project based on support vector machine (SVM). The package supports parameter selection, sample evaluation and testing by referencing SVMcg For son Regress function of support vector machine (SVM) model in MATLAB platform. The unified format is:

$$< \text{label} >< \text{index1} > : < \text{valuel} >< \text{index2} > \cdots$$

Where < label > is the target value of the training sample data, < index1 > : < value > is normalized value (MATLAB in Chinese BBS,2010; AMARI S, WU S,1999.).

B. Parameter optimization

Radial basis kernel function has been chosen above, and now kernel bandwidth g and penalty factor C in the model need to be adjusted carefully by using training set. Specific operation is as follows: first, divide 1-7 sample data from the table 2 into seven mutually disjoint subsets, and set selection range of large area roughly to make the adjusted steps of two parameters to 1; To set up the first set of data for the validation set, and set up the rest of the six groups of data as the training set, and a set of parameters can be established. 7 set of parameters can be obtained in the same manner; Then select the optimal parameters in the region which the parameter is better combined, and the results show in Figure 2.

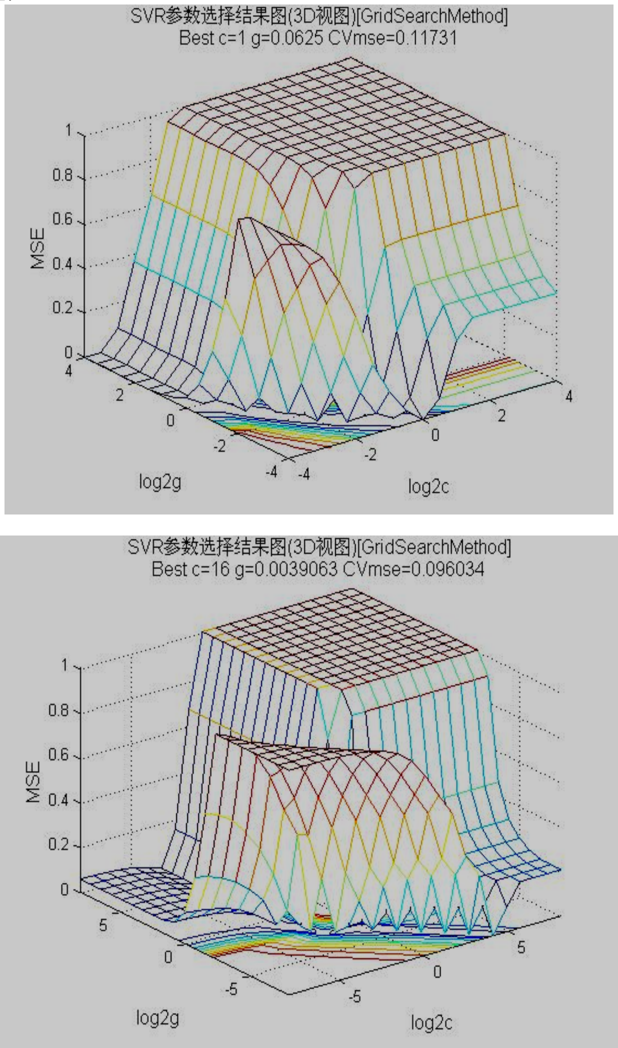

Figure.2: Parameters of Rough and Fine Results of 3D Figure

The figure 2 shows that optimal parameters of support vector machine (SVM) model by referencing radial basis kernel function are g = 0.0625, C = 1.

C. Evaluate Model Training and Inspecting

According determined the radial basis kernel function and kernel width and penalty factor, Train the 1-7 sample data in table 2 in support vector machine (SVM) model, use the obtained support vector machine (SVM) model to forecast the sample data of Group 1-7.

Then evaluate the security risk of 8-10 blocks in table 2 by trained security risk assessment model. Test the generalization ability of the model by comparing the relative error between expected output values and the network output values. The relative error formula is:

$$E_r = \frac{x_p - x_o}{x_o} \times 100\% \tag{4}$$

The x_p represents network output, xo represents expected output, test results are shown in TABLE III.

TABLE III. EVALUATE THE RESULTS OF MODEL TEST

Sample	8	9	10
Network Output	0.64	0.18	0.29
Expected Output	0.7	0.1	0.3
The Mean Square Error	-8.57%	8%	-3.33%

D. Result Analysis

The error graph of network output and the expected output of each sample can be obtained according to the model evaluation (figure 3).Under the circumstances of setting the parameters of support vector machine (SVM) model as g = 0.0625, C = 1, through the safety risk assessment of 10 sample data in the living example, the result is: the mean square error (mse) = 0.1276, the correlation coefficient = 88.7137%.

It is clear that the safety risk assessment model of this project which is established based on support vector regression is more ideal that the network output is close to the expected output; The result of calculation shows that the model consistent with experts predicted risk value. At the same time as shown, though samples are limited, the model based on support vector regression machine can still play a powerful generalized force of model.

After the model of support vector regression machine optimal trained, according to the determined parameters, an evaluation model of formula (2) can be established, through the calculation of the inputted data, the output is the safety risk value of newly building highway crossing under the existing railway.

VI. CONCLUSIONS

At present, the risk assessment study about the cross-railway highway engineering is less in our country. This paper applied support vector machine (SVM) regression method to establish evaluation model to study the safety risk assessment of cross-railway highway engineering, which improve the accuracy of the evaluation results. At the same time, an instance is used to prove that the safety risk assessment model based on support vector regression machine has strong practicality. All the engineering materials and data in this article are from a specific project. Due to the differences between various engineering, the application of safety risk assessment model for other similar engineering based on support vector machine (SVM) need to be revised and investigation.

REFERENCES

[1] LiuChong and ZhouLiAn,2014. Highway construction and regional economic development: The evidence from China's county level.[J]. Economic science,02:55-67.

[2] A.M. Hefny, H.C. Chua, J.Zhao.,2004.Parametric studies on the interaction between existing and new bored tunnels[J].Tunneling and Underground Space Technology, (19):471-478.

[3] QianQiHu and RongXiaoLi,2008. China's underground engineering safety risk management present situation, problems and proposal.27(4):649-655.

[4] ChenAiRong and RuanXin, 2005. Sutong bridge girder river cable tower construction during the construction project risk assessment [R].

[5] HouYanJuan and ZhangDingLi,2011.The risk management system of Urban tunnel construction through building.[J]. Chinese Journal of Underground Space and Engineering, 7(5):989-995.

[6] FanXinWei,2003. The research and application of support vector machine (SVM) algorithm .[D]. Zhejiang university,2003.

[7] DengNaiYang and TianYingJie,2009. Support vector machine (SVM) theory, algorithms and expand [M]. Science press, Beijing.pp:97-104.

[8] WangWei and GuoXiaoMing,2008. The selection of kernel function method [J]. Journal of Liaoning Normal University(Natural Science Edition),32(1):1-4.

[9] YangLi,2010. Based on the sample data of well gas outburst risk assessment.[D] University of Scienceand Technology of China,HeFei.

[10] MATLAB in Chinese BBS,2010. MATLAB neural network 30 case analysis [M] Beijing university of aeronautics and astronautics press.pp:112.

[11] AMARI S, WU S,1999. Improving support vector machine classifiers by modifying kernel functions[J].Neural Networks, 12:783-789.