2010 3rd International Conference on Computer and Electrical Engineering (ICCEE 2010)
IPCSIT vol. 53 (2012) © (2012) IACSIT Press, Singapore
DOI: 10.7763/IPCSIT.2012.V53.No.1.53

Prediction Control for ATE Metrology Parameters Based on Improved LS-SVM

MAO Hong-yu+, YANG Guang, YANG Hong-sheng and HU Zhuo-lin[1]

Jilin University, Changchun, China

[1]Aeronautical Equipment Measurement Master Station, Beijing, China

Abstract. An improved least squares support vector machines (LS-SVM) was proposed to improve the sparse and robust performance of LS-SVM in the small samples prediction. The sparse and robust performance could be improved through adding elements of weighted LS-SVM and robust LS-SVM. We introduced a contrast experiment for Auto Test Equipment (ATE) parameters prediction control through the three methods of neural network, LS-SVM and improved LS-SVM algorithm. Simulation results show that the improved LS-SVM algorithm has good performance in ATE parameters prediction, which succeeds in stability assessment for an aviation ATE.

Keywords: Improved LS-SVM; ATE; Parameter prediction; Non-linear;

1. Introduction

As the Auto Test Equipment is widely applied in aerospace, industrial control, and scientific experiments fields, the studies on ATE parameter prediction and measurement modification take over the domain status [1]. The reference [2] proposed the ATE parametric variation based on the analysis of entire systematic model, and it analyzed and traced the reasons which had caused the error to ATE system. This methodology is based on "whitening" principle to achieve the error analysis of ATE system, and trace the error components to the parts of ATE internal system. Therefore, it makes the measurement accuracy goes to a further step. However, the function of this movement is only limited to find out the reasons, but it has nothing to do with the prediction of parametric variation itself. Differently, reference [3] treats stable measured object as its checking standard, and records the features between measurement standard and time constant capability. This method works when the parametric variation accords with normal distribution, and it is provided with stationary random process. The disadvantage is that the hypothesis doesn't match the practical situation of parametric changes, and it lacks complexity of the changes as well.

Support Vector Machine is a new sprinted up machine learning algorithm which is based on statistical theory. It is to reduce the structural risk, and solve the prediction problems for small sample, nonlinear, and high dimension parameters. In 1999, Suykens J.A.K and his group mentioned the least squares support vector machines at the first time to figure out the parametric prediction and function estimation issues [4]. They introduced the least squares linear system to be loss function to replaced traditional support vector machines quadratic programming problem. It obviously increased research speed [5], and it is widely applied to non-stationary time series prediction and generalized control progress areas [6].

Combined with the dynamic measurement of a certain type of air ATE characteristics of nonlinear time-varying parameters, this paper proposes a kind of measurement parameters prediction ATE based on improved LS-SVM , and obtains the desired results from experiments.

+ Corresponding author.
E-mail address: 24427732@qq.com.

2. Principles and Methods for Improved LS-SVM

SVM is the last century 90's developed based on structural risk minimization algorithm. SVM is to map the input space to a high dimensional space for nonlinear problems, space in the original nonlinear problem into a feature space of the linear problem, thus reducing the computational complexity.

2.1 Assessment principles for standard LS_SVM

LS-SVM uses equality constraints instead of inequality constraints in support vector machines, and obtains parameter analytical solution by solving a set of linear equations, thus avoiding the use of quadratic programming SVM to solve function estimation. Algorithm described as follows:

For a group given training samples $\{(x_i, y_i)\}_{i=1}^{N}$, where N is the number of training samples, $x_i \in R^n$ are input vectors， $y_i \in R$ are output results. Through a nonlinear mapping function $\Phi(x_i)$, the sample from the original space R^n is mapped to high dimensional feature space Z, in high dimensional feature space, an optimal decision function is constructed:

$$y(x) = \omega^{\mathrm{T}} \cdot \varphi(x) + b \quad (1)$$

In this way, non-linear mapping function has been converted into high dimensional feature space linear estimation functions, where, ω and b is parameters to be determined. Based on structural risk minimization principle, the regression problem is transformed to meet:

$$y(x_i) = \omega^{\mathrm{T}} \cdot \varphi(x_i) + b + e_i\,,\ i=1,2,\ldots,\mathrm{N}. \quad (2)$$

And also meet Inequality (3) optimal solution:

$$\min_{\omega,e} J(\omega,e) = \frac{1}{2}\omega^T\omega + \frac{1}{2}\gamma\sum_{i=1}^{N} e_i^{\,2}, \gamma > 0 \quad (3)$$

The first control model complexity, the second control model of precision, normal number of γ is the compromise parameter between generalization ability and accuracy. e_i is the difference between predictive value and the actual value. Constructing the Lagrange function of the optimization problem:

$$L(\omega,b,e,\alpha) = J(\omega,e) - \sum_{i=1}^{N}\alpha_i(\omega^{\mathrm{T}} \cdot \varphi(x_i) + b + e_i - y(x_i)) \quad (4)$$

In the formula, $\alpha_i \in R$(i=1, 2,⋯ N)is the Lagrange factor. According to optimal conditions, we get the partial differentiation for L to variable ω, b, e_i, α_i, and obtain the matrix equation by off variable ω, e_i,

$$\begin{bmatrix} 0 & \vec{1}^T \\ \vec{1} & \Omega + \gamma^{-1}I \end{bmatrix}\begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix},$$

The vector $y = [y_1 y_2,\ldots,y_N]^T$, $\vec{1} = [1\,1,\ldots,1]^T$, $\alpha = [\alpha_1\alpha_2,\ldots,\alpha_N]^T$, Ω is a symmetric matrix $N \times N$:

$$\Omega_{ij} = \varphi(x_i)^T\varphi(x_j) = K(x_i, x_j); i,j = 1,2,\ldots,N$$

We obtain eventually the function estimation formula:

$$y(x) = \sum_{i=1}^{N}\alpha_i K(x, x_i) + b \quad (5)$$

Among, K(*,*) is kernel function.

Standard LS-SVM advantages in maintaining the standard SVM, and reduces significantly computing costs, but to the expense of loss of sparseness and robustness of the decline. In order to improve its sparse and robust, this paper proposes an improved LS-SVM algorithm.

2.2 Improved LS_SVM algorithm

The basic thought is as follows, we get the initial training data samples firstly based on LS-SVM, then adopt weighted LS-SVM training algorithm on the based of initial training for improving robustness of the training data. Finally, we adopt sparse LS-SVM training algorithm on the based of weighted LS-SVM for improving sparseness of the training data. Therefore, each LS-SVM model to achieve the selection of parameters (including the kernel function type and the parameters of), LS-SVM initial training, and robust and sparse modeling training, has been built on the final model results.

1) *Weighted LS-SVM training algorithm*

The basic thought is as follows, weighted error components form formula (3), set weighted coefficient v_i, and objective function is indicated as follows,

$$\min_{\omega,e} J(\omega,e) = \frac{1}{2}\omega^T\omega + \frac{1}{2}\gamma\sum_{i=1}^{N} v_i e_i^{\,2}, \gamma > 0 \tag{6}$$

Other constraints remain unchanged, weighted coefficient is fixed by unweighted LS-SVM error variable e_i [7], specifically expressed as:

$$v_i = \begin{cases} 1 & \text{if } |e_i/\hat{s}| \le c_1 \\ \dfrac{c_2 - |e_i/\hat{s}|}{c_2 - c_1} & \text{if } c_1 \le |e_i/\hat{s}| \le c_2 \\ 10^{-4} & \text{others} \end{cases} \tag{7}$$

$\hat{s}$ is robust estimation of standard deviation e_i, c_1、c_2 are constants. The objective function is indicated as follows,

$$\begin{cases} \dfrac{\partial L}{\partial \omega} = 0 & \Rightarrow \omega = \sum_{i=1}^{N}\alpha_i\varphi(x_i) \\ \dfrac{\partial L}{\partial b} = 0 & \Rightarrow \sum_{i=1}^{N}\alpha_i = 0 \\ \dfrac{\partial L}{\partial e_i} = 0 & \Rightarrow \alpha_i = \gamma v_i e_i \qquad i = 1,2,\ldots,N \\ \dfrac{\partial L}{\partial \alpha_i} = 0 & \Rightarrow \omega^{\mathrm{T}}\cdot\varphi(x_i) + b + v_i e_i - y_i = 0 \end{cases}$$

We obtain weighted LS-SVM equation by off variable ω, e_i,

$$y(x)\big|_{\hat{s}} = \sum_{i=1}^{N}\alpha_i K(x,x_i) + b \tag{8}$$

The kernel function,

$$K(x_i,x_j) = \exp(-\|x_i - x_j\|^2/2\sigma^2) \tag{9}$$

As the robustness estimation of the error, we obtain a better robustness for improved LS-SVM than standard LS-SVM.

2) *Sparse LS-SVM training algorithm*

In the process of calculation, we should discard some data which α_i is equal zero for improving processing speed. So we can divide into the following four steps for sparse LS-SVM.

① First, we get α_i through the initial training data samples based on standard LS-SVM.

② Arranged according to the value $|\alpha_i|$, and we should discard the smaller part of the data points(such as the relative minimum of 3% of data points removed).

③ Continue to use the LS-SVM training data points in the remainder basis;

④ Return ②,until meet the performance (such as computing speed, generalization capability).

As the sparse training algorithm, we obtain a faster computational speed for improved LS-SVM than standard LS-SVM.

2.3 The ATE metrology parameter prediction based on improved LS_SVM

ATE metrology parameters are these performances and technical parameters representing the normal operation of ATE, these parameters include the electrical parameters, time /frequency parameters, pressure, temperature and other high-precision parameters. These parameters will be drift with the pass of time, and also occur non-linear change. Especially the air ATE for motorized operation, the drift will seriously affect the flight safety.

Parameter change is a time series with typical nonlinear characteristics meeting the forecast range based on LS-SVM time series model. To the electrical of ATE parameters, for example, we establish the time series for change in electrical parameters, and suppose the history of all the observed time series of electrical parameters are:

$$\{ y(k), k = 1,2..., N \} \tag{10}$$

In order to forecast $y(k+1)$, we need establish mapping

$$F : R^n \to R, \hat{y}(k+1) = F(y(k-(m-1)),...,y(k-1), y(k)) \tag{11}$$

And $\hat{y}(k+1)$ is a forecast value for $y(k+1)$, m is the embedding dimension, $F(\bullet)$ is a non-linear function reconstructed by samples data as before. The specific expression is just like the equation (7) and (8) .We can obtain optimal predictor topology, as Fig1,

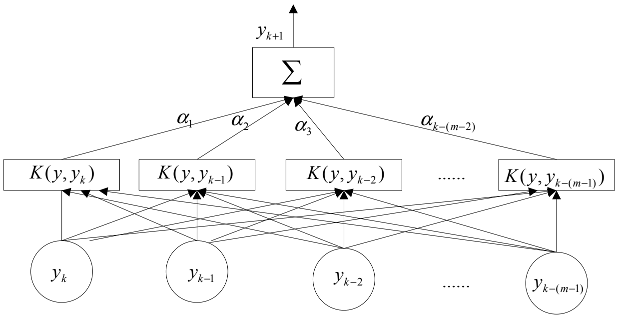

Fig.1 Topological structure of time series model on LSSVM

In the model, we select the most commonly used RBF function as kernel function, the weights is as follows,

$$v_i = \begin{cases} 1 & if \ |e_i / \hat{s}| \le 2.5 \\ \dfrac{3 - |e_i / \hat{s}|}{3 - 2.5} & if \ 2.5 \le |e_i / \hat{s}| \le 3 \\ 10^{-4} & others \end{cases} \tag{12}$$

Repeat the training several times, until the generalization performance and computational speed to meet the requirement, and then the results of this training time to achieve the parameters of the next moment predicts.

3. Experiment Study

We select the voltage parameters recorder form a certain type of air ATE-I channel form 2008.6 to 2010.2, and train and forecast changes in ATE metrology parameters by using standard LS-SVM, improved LS-SVM and traditional neural network. These samples are as follows Table 1,

Table1. Historical Verification data of ATE-I channel

Date	Voltage(mV)	Date	Voltage(mV)
2008.06	3.321	2009.04	3.330
2008.07	3.325	2009.05	3.325
2008.08	3.322	2009.06	3.326
2008.09	3.323	2009.07	3.329
2008.10	3.321	2009.08	3.331
2008.11	3.323	2009.10	3.333
2008.12	3.326	2009.11	3.326
2009.01	3.325	2009.12	3.332
2009.02	3.328	2010.01	3.338
2009.03	3.327	2010.02	3.323

After repeated experiments and considering synthetically the generalization ability and accuracy, we select γ=510, σ^2=1.5, embedding dimension m=5. Therefore, the data of Tab.1 can be divided into 16 groups. We select the first 14 groups to be the training samples and the last 2 groups to be the validation samples. Fig.2 showing the training results used respectively three algorithms (standard LS-SVM, improved LS-SVM and neural network) to fit curves. The x-axis indicates the date 2008.6~2010.2. The data of the first 18 months is fitting samples, and the last 2 months is predicting data.

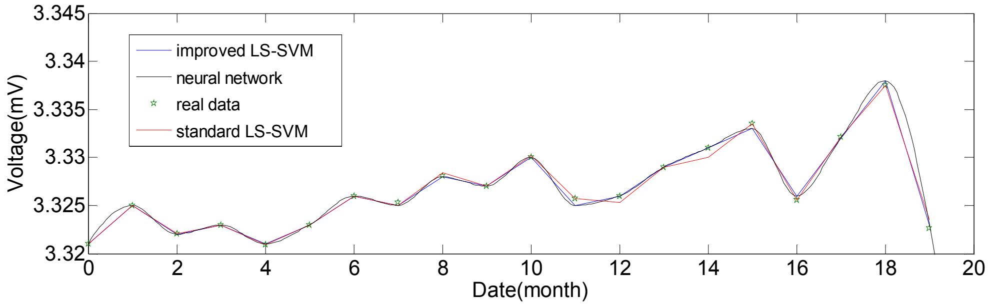

Fig.2 Voltage parameters fitting predict curve on three methods

Fig.3 shows the fitting error of three algorithms. From it we can conclude that improved LS_SVM has obvious advantage under small sample conditions. It can be seen that the improved LS-SVM is more suitable for small samples, nonlinear parameter prediction.

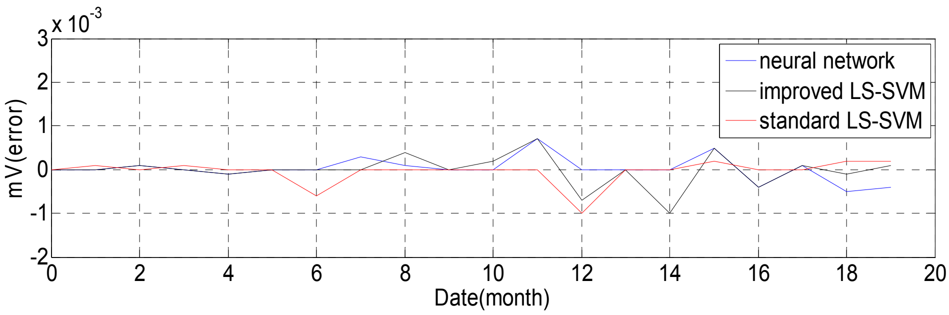

Fig.3 Voltage parameters fitting predict error on three methods

4. Conclusion

Experiments show that the improved LS-SVM has higher prediction accuracy than the standard LS-SVM and the traditional neural network. Presently, the method has been applied to the stability evaluation project for a kind of air ATE. The method has special reference to system identification in other fields of study.

5. References

[1] Lee Qiong-wei, Fee Xiao-jun, Research on System-level Automatic Metrology System for Aircraft Integrated Automatic Test Equipment[A].Electronic Measurement and Instruments[C].2007.ICEMI'07.8th,International Conference On. Xi' an, China.

[2] Mao hong-yu, Yang Hong-sheng, Wang Wen-liang, Li Qiong-wei. Research on precision model of dynamic metrology system for IATE[C]. Electronic Measurement and Instruments. 2009. ICEMI'09.8th, International Conference On. Beijing, China.

[3] Shen Jia-jie, Yang Jian-wei, He Ling. Discusses Between the Checking Stability of Measurement Standard and Verification of Measurement Standard [J], Metrology and Testing Technology, 2009，36(10): 6-8.

[4] Suykens J A K, Vandewalle J. Least squares support vector machine classifiers [J]. Neural Processing Letters (S1370-4621), 1999, 9(3): 293-300.

[5] Wang Ge-Li, Yang Pei-Cai, Mao Yu-Qing. On the application of non-stationary time series prediction based on the SVM method[J]. ACTA PHYSICA SINIC, 2008,57(2):714-718.

[6] Li Li-juan. The study of Modeling Algorithm Based on LS-SVM and Predictive Control Algorithm [D].2008.12

[7] Suykens J A K, De Brabanter J, Lukas L, Vandewalle J. Weighted least squares support vector machines: robustness and sparse approximation [J]. Neuron computing (S0925-2312), 2002, 48(1):85-105.

[8] Suykens J A K, Lukas L,Vandewalle J.Sparse approximation using least square support vector machines[C]. IEEE International Symposium on Circuits and Systems (ISCAS2000), Geneva, Switzerland, 2000.USA: IEEE, 2000, II: 757-760

[9] ZHANG Chao-yuan, CHEN Li. Improved algorithm of LS-SVM and its application to traffic flow prediction [J]. Journal of Kun-ming University of Science and Technology (Science and Technology). 2008，33(6):72-76