



NONLINEAR SYSTEM IDENTIFICATION BASE ON FW-LSSVM

^{1,2}XIANFANG WANG, ¹YUANYUAN ZHANG, ¹JIALE DONG, ³ZHIYONG DU

¹ College of Computer & Information Engineering, Henan Normal University, Xinxiang 453007, China

² Henan Province Colleges and Universities Engineering Technology Research Center
for Computing Intelligence & Data Mining, Xinxiang 453007, Henan, China

³ Henan Mechanical and Electrical Engineering College, Xinxiang 453002, Henan, China

E-mail: xfwang11@yahoo.com.cn

ABSTRACT

This paper proposed a method to identify nonlinear systems via the fuzzy weighted least squares support machine (FW-LSSVM). At first, we describe the proposed modeling approach in detail and suggest a fast learning scheme for its training. Because the training sample data of independent variable and dependent variable has a certain error, and we obtain the sample which has a certain fuzziness from measuring, influencing the accuracy of model building, this paper would put the concept of fuzzy membership into the least squares support vector machine. Using the fuzzy weighted least squares algorithm for samples, each sample in the vector is introduced into the fuzzy membership degree, this improve the anti noise ability of least squares support vector machine. The efficiency of the proposed algorithm was demonstrated by some simulation examples.

Keywords: *Mixed Kernel Function, FW-LSSVM, Identification, Nonlinear Systems.*

1. INTRODUCTION

Most systems encountered in the real world are nonlinear in nature, and since linear models cannot capture the rich dynamic behavior of limit cycles, bifurcations, etc. associated with nonlinear systems, it is imperative to have identification techniques which are specific for nonlinear systems[1-4].

SVM is employed to realize some complex nonlinear desired functions, which is presented by Vapnik and based on the statistical learning theory and learning method has obtained certain achievements. SVM has strong generalization ability in solving the small sample, nonlinear, high dimension, and local minimum of practical problems. SVM use a variety of kernel function which can be classified into global kernel function and local kernel function [5]. Smits and Jordaan use the two kernel function to make up a kind of mixed kernel function through the research on two kinds of representative global kernel function (polynomial kernel function) and local kernel function (RBF kernel function) mapping properties [6]. In addition, Suykens presents a least squares vector machine (LSSVM) method based on SVM. LS-SVM transform the traditional SVM inequality constraints into equality constraints, in order to avoid a quadratic programming problem solving

time-consuming, and LS-SVM reduce the run time, and effectively improve the speed of learning to solve at a large extent. Although the algorithm and the structure of the LS-SVM are simple, and the calculation of LS-SVM is fast, it lost the loose and robustness of SVM in a certain extent.

This paper overcome the adverse effects of the training sample noise by selecting the fuzzy weighted least squares support vector machine algorithm and putting fuzzy weighted membership on the squared error of each sample [7-9]. Selecting LS-SVM which is constructed with mixed kernel function not only has a good learning ability but also has better generalization ability. Simulation results show that the fuzzy weighted least squares support vector machines help to improve the robustness of the model and have a certain practicality.

Briefly, the paper is organized as follows: Section 2 presents brief theory of least squares support vector machine(LSSVM). Section 3 deals with the description of the fuzzy weight LSSVM. Section 4 provides an illustrative example to identify for the nonlinear systems. Finally, some conclusions are obtained in Section 5.



2. THEORY OF LSSVM

LSSVM(Least squares support vector machine) [5] flow the principle of the structural risk minimization of the SVM algorithm in training process, and the inequality constraints of the objective function in SVM is transformed into the equality constraint. Select a nonlinear mapping and input vector is mapped from the original space to a high dimensional feature space. And the optimal decision function is constructed by using the principle of the structural risk minimization in the high dimensional feature space. Dot product operation in high-dimensional feature space is replaced by the kernel function of the original space, and the first power of the empirical risk is replaced by the square in this procedure. The insensitive loss function is avoided, and the complexity is greatly reduced and that the obtained external solution which is a global optimal solution is ensured.

The article assumes that the sample training set is given. The sample training set is $\{(x_1, y_1), (x_2, y_2) \dots (x_l, y_l)\} \in R^n \times R$, and x_l is the l -th sample of the input mode and y_l is the l -th the desired output of the l -th sample, and l is the total number for the training sample. The samples are mapped to the high-dimensional feature space from the original space R^n by using the nonlinear mapping $\varphi(\cdot)$. In the feature space, the model of least square support vector machine is

$$y(x) = w^T \varphi(x) + b \tag{1}$$

where $\varphi(x)$ is the feature mapping and w^T represents the weight vector, and b is the offset vector.

Based on the structural risk minimization the recursive model of LS-SVM for solving the function (1) is

$$\min J(w, e) = \frac{1}{2} w^T w + \frac{\gamma}{2} \sum_{i=1}^l \xi_i^2 \tag{2}$$

$$s.t. \quad y_i = w^T \varphi(x_i) + b + \xi_i$$

where γ is called the adjustable constant and $\gamma > 0$, ξ_i is the error term, and $\xi_i \geq 0$.

The corresponding Lagrangian function for (2) is as follows

$$L(w, b, \xi, \alpha) = \frac{1}{2} \|w\|^2 + \frac{\gamma}{2} \sum_{i=1}^l \xi_i^2 - \sum_{i=1}^l \alpha_i \{w^T \varphi(x_i) + b + \xi_i - y_i\} \tag{3}$$

where α_i is Lagrange multiplier and $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_l]^T$. Based on KKT which is the optimal conditions, the linear equations are obtained after eliminate ξ_i and w .

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^l \alpha_i \varphi(x_i) \\ \frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^l \alpha_i = 0 \\ \frac{\partial L}{\partial \xi} = 0 \rightarrow \alpha_i = \gamma \xi_i \\ \frac{\partial L}{\partial \alpha} = 0 \rightarrow y_i = w^T \varphi(x_i) + b + \xi_i \end{cases} \tag{4}$$

Then the optimization problem is transformed into the following linear equation:

$$\begin{bmatrix} 0 & 1 & L & 1 \\ 1 & K(x_1, x_1) + \frac{1}{\gamma} & L & K(x_1, x_1) \\ M & M & L & M \\ 1 & K(x_l, x_l) & L & K(x_l, x_l) + \frac{1}{\gamma} \end{bmatrix} \begin{bmatrix} b \\ \alpha \\ M \\ \alpha_l \end{bmatrix} = \begin{bmatrix} 0 \\ y_1 \\ M \\ y_l \end{bmatrix} \tag{5}$$

where $K(x_k, x_l) = \varphi(x_k)^T \varphi(x_l)$ is the kernel function. Linear kernel, polynomial kernel and RBF kernel are the common kernel functions. In this paper, the mixture kernels which consist of polynomial kernel and RBF kernel are used. The expression is as follow

$$K_{mix} = \rho[(x \cdot x_i) + 1]^2 + (1 - \rho) \exp(-\|x - x_i\|^2 / \sigma^2) \tag{6}$$

where $\rho(0 \leq \rho \leq 1)$ is the constant to regulate the effect of the two kernel functions. It is important to select the parameters ρ, σ and the penalty-factor C . In order to get the model that has the predicted effect, it is necessary to optimize and adjust the parameters [10-11].

3. FW- LSSVM

LSSVM improves the normal SVM, but at the same time it loses the robustness by itself, and it doesn't fit the fact that the training data has the same affection in the objective function. Because the sample data have different positions, the extent of influencing by the noise and the error exists that have different and important influence in the measure. So for acquiring a better robustness, each sample data in the vector is introduced into the fuzzy membership degree[12]. The objective function (2) is replaced by the following function.

$$\min J(w, e) = \frac{1}{2} w^T w + \frac{\gamma}{2} \sum_{i=1}^l \mu_i e_i^2 \quad (7)$$

The corresponding Lagrangian function for Equation (7) is as follows:

$$L(w, b, \xi, \alpha) = \frac{1}{2} \|w\|^2 + \frac{\gamma}{2} \sum_{i=1}^l \mu_i \xi_i^2 - \sum_{i=1}^l \alpha_i \{w^T \phi(x_i) + b + \xi_i - y_i\} \quad (8)$$

The linear equations are obtained for the function (8):

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^l \alpha_i \phi(x_i) \\ \frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^l \alpha_i = 0 \\ \frac{\partial L}{\partial \xi} = 0 \rightarrow \alpha_i = \gamma_i \mu_i \xi_i \\ \frac{\partial L}{\partial \alpha} = 0 \rightarrow y_i = w^T \phi(x_i) + b + \xi_i \end{cases} \quad (9)$$

Then the optimization problem of the function (9) is transformed into the following linear equation:

$$\begin{bmatrix} 0 & 1 & L & 1 \\ 1 & K(x_1, x_1) + \frac{1}{\mu_1} & L & K(x_1, x_1) \\ \dots & \dots & \dots & \dots \\ M & M & L & M \\ 1 & K(x_l, x_l) & L & K(x_l, x_l) + \frac{1}{\mu_l} \end{bmatrix} \begin{bmatrix} b \\ \alpha_1 \\ \dots \\ \alpha_l \end{bmatrix} = \begin{bmatrix} 0 \\ y_1 \\ \dots \\ y_l \end{bmatrix} \quad (10)$$

The parameters α and b can be obtained by the Equation(10), and selecting the weighted coefficient μ_i is based on the error $\xi_i = \alpha_i / \gamma$ of the LS-SVM.

4. IDENTIFICATION VIA FW-LSSVM

Consider the nonlinear system

$$y(k) = f(y(k-1), y(k-1), L, y(k-n), u(k-1), u(k-2), L, u(k-m)) \quad (11)$$

where $u(k)$ is the system input, $y(k)$ the system output; where m and n are respectively the delays of $u(k)$ and $y(k)$. $f(\cdot)$ is a nonlinear function.

The identification of nonlinear system based on FW-LSSVM model is depicted in Figure 1.

$$y_M(k) = f_{svm}(y_M(k-1), y_M(k-1), L, y_M(k-n), u(k-1), u(k-2), L, u(k-m)) \quad (12)$$

where $y_M(k)$ is the output of the identification model, $f_{svm}(g)$ is the obtained model by training FW-LSSVM.

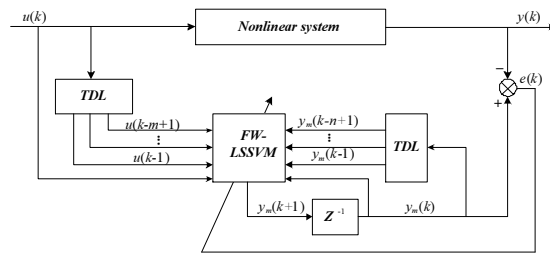


Figure 1: The Diagram Identification Of The Nonlinear System Based On SVR

The detailed process of this identification can be described as following:

Step 1 Chose the input signal $u(t)$ for the nonlinear system Equation(11).

Step 2 Chose the sampling time t_s , sampling dots N and width of window is $t = N * t_s$.

Step 3 Initialed the input and output, and obtained the values of $u(k)$ and output $y(k)$ at every sampling time k .

Step 4 Choose the delays of input m and output n ($m \geq n \geq 0$). Construct a sampling set $D = \{X, Y\}$, where

$$\begin{cases} X(k-m) = \{y(k-1), y(k-2), L, y(k-n), \\ u(k-1), u(k-2), L, u(k-m)\} \\ Y(k-m) = y(k) \\ k = (m+1), L, N \end{cases}$$

Step 5 Chose the kernel function $K(x_i, x_j)$ determine parameters of the FW-LSSVM (such as γ, μ_i)

Step 6 Computing the output Y_M when the input is, and the training error $E = Y - Y_M$.

Step 7 If the error E is less than a threshold, the training is over, then go to the next step; else return to Step 5.

Step 8 The parameter vector w would be obtained by solving quadratic regulation in Equation (9).

Step 9 The model is obtained, the identification process of nonlinear system is end.

5. THE SIMULATION EXPERIMENT

In order to prove that the opinion put forward in the paper is effectiveness, the nonlinear system is regarded as the object to emulate. The personal

computer and the software of Matlab 7.0 are needed to accomplish the experiment.

The nonlinear model is as follow [13]:

$$y(k+1) = 0.6 * y(k) / (1 + y(k) * y(k)) + 0.3 * u(k) * y(k) + 0.8 * u(k) \quad (13)$$

The data need to be trained before test the fuzzy weighted least squares support vector machine. The input sample of the training data are structured the array and every row of the array represents the input vector. The training sample is:

$$\begin{cases} X(k) = [y(k+2), y(k+1), \\ u(k+2), u(k+1), u(k)] \\ Y(k) = y(k+3) \end{cases} \quad (14)$$

In the model, the sampling time $t_s = 0.01s$, sampling dots $N = 200$, the width of window is $t = N * t_s = 2s$. The mixture kernel function

$K_{mix} = \rho[(x \cdot x_i) + 1]^2 + (1 - \rho) \exp(-\|x - x_i\|^2 / \sigma^2)$ is used.

The parameters of the kernel function has the stronger influence, so the value of ρ is 0.01 through the experience. The output of the training data in the model are compared with the actual output until the error is the minimum. The vectors of σ and γ are gained when the error is the minimum. The vector of σ is 0.2756 and the positive number γ is 200. Figure 2 is the output of the simulated nonlinear system model and the Figure 3 is the output error of modeling.

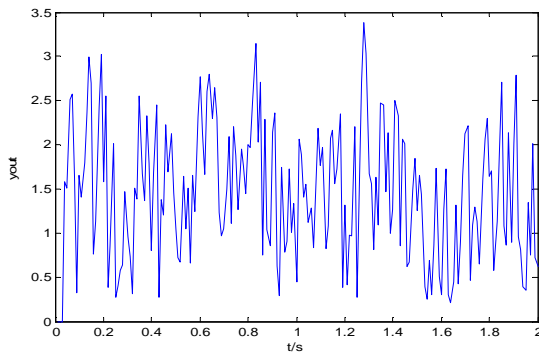


Figure 2: The Output Of The Training Model

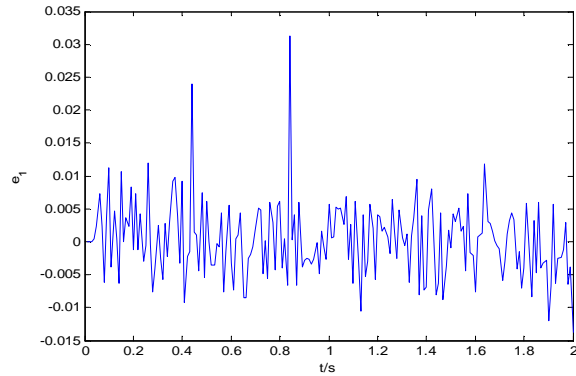


Figure 3: The Training Error Curve Of When The Input Is A Random Signal

In the fuzzy weighted least squares support vector machine of the nonlinear system, the random signal $u_r(k) = rand(1)$, unit step signal, sinusoidal signal $0.5 \sin(2\pi t + 0.5)$ and the square signal $0.5 \text{sign}(\sin(2\pi t)) + 0.5$ are used to test the generalization performance of the model. The output error curves of the fuzzy weighted LS-SVM nonlinear system identification model are the figure from Figure 4 to Figure 6.

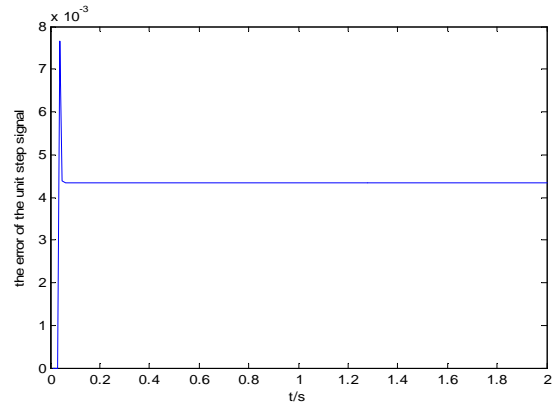


Figure 4: The Identification Error Curve When The Input Is A Unit Step Signal

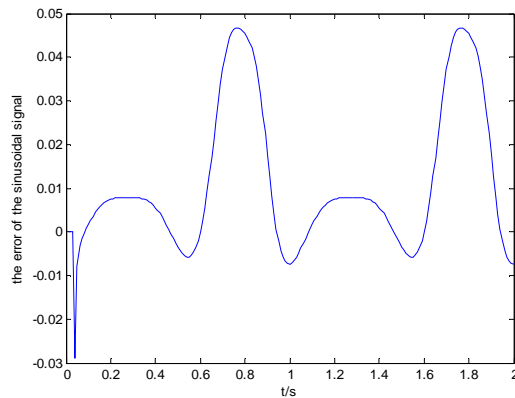


Figure 5: The Identification Error Curve When The Input Is A Sinusoidal Signal

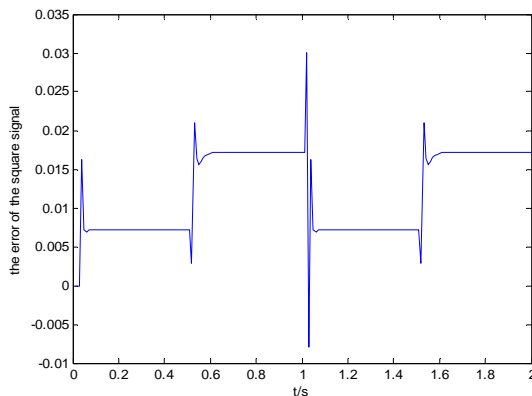


Figure 6: The Identification Error Curve When The Input Is A Square Signal

Figure 3 to Figure 6 indicate that using the fuzzy weighted LS-SVM avoids the over-fitting reduces the incompactness. The simulation experiment proves that the fuzzy weighted LS-SVM nonlinear system has a better robust.

6. CONCLUSIONS

In this paper, we present an improvement method to identify nonlinear system. The algorithm is put the fuzzy weighted to use LSSVM. Based on that the samples have different and important influence in every model of the productive process, fuzzy weighted algorithm is introduced. The influence of the abnormal training data are overcome and the over-fitting of the identification model is avoided. The simulation experiment indicates that the fuzzy weighted LS-SVM and the mixture kernel function avoids the algorithm convergent prematurely and improves the fuzzy learning precision and the generalization ability. Then the robustness of the system is improved and the fuzzy weighted LS-SVM has a better

application. The results demonstrate that the suggested method gives better error minimization for nonlinear system.

ACKNOWLEDGEMENTS

This work is supported by the National Natural Science Foundation of China (No.61173071), the Science and Technology Research Project of Henan Province(No.122102210079), the Foundation and Frontier Technology Research Programs of Henan Province(No.112400430087), the Innovation Talent Support Program of Henan Province Universities (No.2012HASTIT011), the Doctoral Started Project of Henan Normal University(No.1039)

REFERENCES:

- [1] Musa Alci, Musa H. Asyali, "Nonlinear system identification via Laguerre network based fuzzy systems", *Fuzzy Sets and Systems*, Vol. 160, No. 24, 2009, pp. 3518-3529.
- [2] Johan Paduart, Lieve Lauwers, Jan Swevers, Kris Smolders, Johan Schoukens, Rik Pintelon, "Identification of nonlinear systems using Polynomial Nonlinear State Space models", *Automatica*, Volume 46, No. 4, 2010, pp. 647-656.
- [3] Alireza Rahrooh, Scott Shepard, "Identification of nonlinear systems using NARMAX model", *Nonlinear Analysis: Theory, Methods & Applications*, Vol. 71, No. 12, 2009, pp. e1198-e1202.
- [4] Gregor Dolanc, Stanko Strmčnik, "Identification of nonlinear systems using a piecewise-linear Hammerstein model", *Systems & Control Letters*, Vol. 54, No. 2, 2005, pp. 145-158.
- [5] Suykens J A K, "Vandewalle J. Least squares support vector machine classifiers", *Neural Processing Letters*, Vol. 9, No. 3, 1999, pp. 293-300.
- [6] Mao Zhi-liang, Liu Chun-bo, Pan Feng, "Parameter Selection and Application of SVM with Mixture Kernels Based on IPSO", *Journal of Southern Yangtze University*, Vol. 8, No. 6, 2009, pp. 632-634.
- [7] Zhang Ying, Su Hong-ye, Chu Jian, "Soft sensor modeling based on fuzzy least squares support vector machines", *Control and Decision*, Vol. 20, No. 6, 2005, pp.621-624.
- [8] Zhang Zai-fang, Chu Xue-ning, Cheng Hui, "Aggregating Multi-granularity Decision



- Making Information Based on Fuzzy Weighted Least Squares Method”, *Journal of Shang Hai Jiao Tong University*. Vol. 43, No. 9, 2009, pp. 1377-1382.
- [9] Li Zheng, Song Bao-wei, Mao Shao-yong, “Fuzzy Weighted Linear Regression on Zero Failure Exponent Distribution”, *Journal of System Simulation*, Vol. 17, No. 6, 2005, pp.1373-1375.
- [10] Xiong Wei-li, Wang Xiao, Chen Min-fang, Xu Bao-guo, “Modeling for penicillin fermentation process based on weighted LS-SVM”, *CIESC Journal*, Vol. 63, No. 9, 2012, pp.2913-2918
- [11] Wu Tie-bin, Liu Yun-lian, Li Xin-jun, Yin Yong-sheng, “KPCA-fuzzy Weighted LSSVM Based Prediction Method and its Application”, *Computer Measurement and Control*, Vol. 20, No. 3, 2012, pp.617-620.
- [12] Guo Hui, Liu He-ping, Wang Ling. “Method for Selecting Parameters of Least Squares Support Vector Machines and Application”, *Journal of System Simulation*, Vol. 18, No. 7, 2006, pp. 2033-2036.
- [13] Hu Liang-mou, Cao Ke-qiang, Su Xin-bing, Li Na, Hou Yan-yan, “Research on Internal Model Control for Nonlinear System Based on Weighted LS-SVM”, *Mechanical Science and Technology*, Vol. 31, No. 1, 2012, pp. 101-106.