

Sensor Fault Diagnosis Method Based on Support Vector Regression

Shen Xianhao

School of Information Science and Engineering,
Guiling University of Technology
Guiling 54100, China

Abstract-This article design a sensor fault diagnosis system base on Support Vector Regression. SVR is trained out of line, and used online. After being trained, SVR is used to simulate dynamic characteristic of diesel engine temperature-controlling system. The simulation result shows that, SVR can simulate the system more accurately, track the output signal of sensor, and diagnose the sensor system.

Keywords-Support Vector Regression, Sensor, Fault Diagnosis system

I. INTRODUCTION

In a application system, the sensors working environment is complex, wide distribution, data volume, and other reasons which makes the sensor as the weak link in controlling process and becomes one of the most vulnerable system component. The sensor fault can cause system performance degradation and error accumulation, and even lead to paralysis of the entire system. so when the sensor fails, system should timely diagnosis failure, and isolate and signal recovery. So far, the formation of a variety of sensor fault diagnosis methods, such as hardware redundancy, analytical redundancy method, Kalman filter method, neural network methods and the observer group method based on wavelet analysis[1] and so on. In general, these methods have advantages and disadvantages, generally the result is not satisfactory.

Support vector regression (SVR) is based on statistical learning theory, a new machine learning algorithms [2], which uses the structural risk minimization (SRM) principle, and solves the small sample learning problems, with a globally unique optimal solution, generalization ability, the model automatically determined by the algorithm and a series of advantages, and in the face and text recognition, time series prediction, nonlinear system parameter identification, modeling and control other fields has been a certain application, but in the sensor fault diagnosis system are few reports. This article first introduces the principle of support vector regression and SVM regression proposed based on the principle of the sensor fault diagnosis method and steps, and finally to a diesel engine condition monitoring system, the temperature sensor output signal is verified.

II. SUPPORT VECTOR REGRESSION

Support vector machine regression using the basic idea is to support vector regression algorithm is based on statistical learning theory, a nonlinear mapping Φ by the

sample data set (x_i, y_i) ($i = 1, 2, \dots, n$) is mapped to high dimensional feature space, and space structure in this linear regression function:

$$y = w \cdot \Phi(x) + b$$

And the solution w and b can be attributed to the following convex quadratic programming (QP) problem:

$$\begin{aligned} \min_{w, b, \zeta} & \left\{ \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\zeta_i + \zeta_i^*) \right\} \\ \text{s.t.} & \begin{cases} y_i - w^* \Phi(x_i) - b \leq \varepsilon + \zeta_i \\ w^* \Phi(x_i) + b - y_i \leq \varepsilon + \zeta_i^* \\ \zeta_i, \zeta_i^* \geq 0, i = 1, 2, \dots, n \end{cases} \end{aligned} \quad (1)$$

The objective function which reflects the thinking of SRM, which first made the most smooth regression function to improve the generalization ability, a constant $C > 0$ is penalty coefficient, beyond the control of the "pipe" ε punishment for a sample of the degree, ζ_i, ζ_i^* for the slack variables.

Using Lagrange optimization method, the optimization problem can be transformed to its dual form [3]:

$$\begin{aligned} \max & \left\{ -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (a_i - a_i^*)(a_j - a_j^*) K(x_i, x_j) \right. \\ & \left. - \varepsilon \sum_{i=1}^n (a_i + a_i^*) + \sum_{i=1}^n y_i (a_i - a_i^*) \right\} \\ \text{s.t.} & \sum_{i=1}^n (a_i - a_i^*) = 0 \quad (0 \leq a_i, a_i^* \leq C) \end{aligned} \quad (2)$$

Where K is the kernel function. After solution of the optimization problem, we can get ω, b , then the regression function of the form:

$$y = w \cdot \Phi(x) + b = \sum_{x_i \in SV} (a_i - a_i^*) K(x_i, X) + b \quad (3)$$

Where SV is the support vector. Obtained using support vector machine regression function form similar to a neural network, whose output is a linear combination of a number of intermediate nodes, each node corresponds to an intermediate layer of input samples with a support vector dot product, as shown in Figure 1.

According to the most optimal necessary and sufficient condition (KKT conditions) at the optimal solution are [4]:

$$\begin{aligned} & \begin{cases} a_i (\varepsilon + \zeta_i - y_i + w \cdot \Phi(x_i) + b) = 0 \\ a_i^* (\varepsilon + \zeta_i^* - y_i + w \cdot \Phi(x_i) - b) = 0 \end{cases} \\ \text{s.t.} & \begin{cases} (C - a_i) \zeta_i = 0 \\ (C - a_i^*) \zeta_i^* = 0 \end{cases} \end{aligned} \quad (4)$$

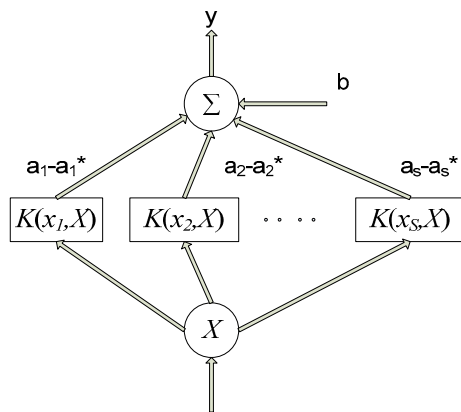


Figure 1 SVM regression network model

It can be seen in the "pipeline" ϵ sample points within the corresponding a_i and a_i^* are equal to zero, with the $a_i \neq 0$ and $a_i^* \neq 0$ corresponding to the sample, that "pipeline" or outside the boundary samples, called support vector. The entire training set to train with the results of the training sample set after the deletion of non-support vector training in the same result, that is, only those in the fitting process only works SVM (support vector which is defined sources), in general, the number of support vectors is much less than the number of training samples.

III. PRINCIPLE OF THE SENSOR FAULT DIAGNOSIS

Observer-based fault diagnosis method requires diagnosis is accurate mathematical model of the object, but in fact most of the mathematical models are nonlinear or higher, and therefore the most important sensor fault diagnosis The question is how to process the correct fitting. The support vector machine for any nonlinear system with high-end systems, can be more accurate fit [5].

Model-based fault diagnosis of residual generation and residual evaluation of two phases, to achieve the principle shown in Figure 2. Application of SVR on the dynamic process simulation to get the normal output of the sensor, compared with the actual system output residuals to create dynamic system builder. The problem is the dynamic process of fitting a multivariate function fitting problems. Sensor output as a function of the output value y , while the corresponding system input and output are some historical input values for this fitting.

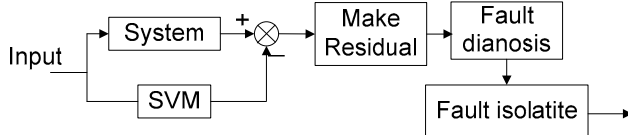


Figure 2 Schematic sensor fault diagnosis

$$\begin{cases} y(k) = f[U(k-1), Y(k-1)] \\ U(k-1) = [u(k-1), \dots, u(k-m)] \\ Y(k-1) = [y(k-1), \dots, y(k-n)] \end{cases} \quad (5)$$

$U(k-1)$ and $Y(k-1)$ were observed for the dynamic system input and output data, and m, n respectively, input and output of the delay, denoted SVR training set of input vectors and output vector prediction output

$$X(k-1) = [y(k-1), \dots, y(k-n), u(k-1), \dots, u(k-m)] \quad (6)$$

$$y(k) = f[X(k-1)] \quad (7)$$

SVR modeling process is the application of the SVR training process. First, in the normal use of system input and output data for offline training of SVR, resulting in SVR model dynamic systems. After completion of system modeling is complete, the use of the trained SVR to generate residuals. SVR-based residual generation principle shown in Figure 3, where a multi-TDL delay systems separately.

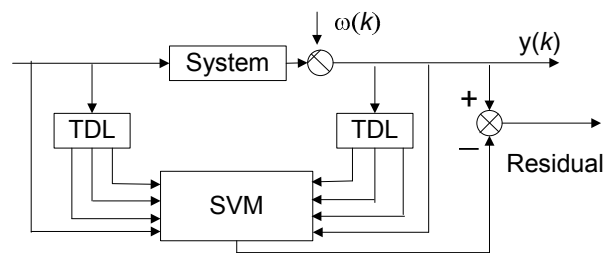


Figure 3 SVR-based residual generation system schematic

In the residual generation, you can use a variety of existing strategies for fault diagnosis and determine the separation, SVR can also be used to generate fault classifier for fault diagnosis.

IV. SIMULATION AND RESULTS ANALYSIS

engine condition monitoring sensor system, including pressure, temperature, torque, blasting, vibration, smoke and other types of sensors (need to monitor the engine is responsible for collecting the signal), in which temperature controlling systems such as diesel Figure 4. Here, using the above method is designed based on the residual generator SVR, and simulation toolbox using MATLAB-SIMULINK simulation to achieve it, in which SVR learning through the preparation of Design. M file to achieve.

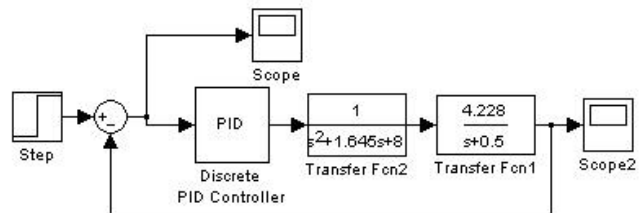


Figure 4 Diesel engine controlling system structure

Based on the above method as shown in Figure 4, the controlling system for the design of a residual generator based on SVR (Figure 5). For this controlling system input and output of the SVR training, SVR training process has 6 inputs:

$u(k-1), u(k-2), u(k-3)$ ----- the input value in $(k-1), (k-2), (k-3)$ moment

$y(k-1)$, $y(k-2)$, $y(k-3)$ ----- the sensor system output value in $(k-1)$, $(k-2)$, $(k-3)$ moment by measured

After training, application of SVR for function approximation, there is always an output, said the main steam temperature system, k the temperature sensor SVR estimates. System implementation structure shown in Figure 5, the figure of SVR training network using online applications offline work.

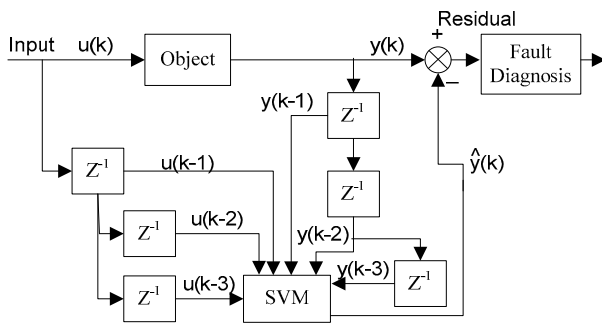


Figure 5 Fault Diagnosis Based on SVR in SIMULINK

SVR is trained off-line by normal system data, when training is ended, cut off the end of the learning process, the SVR into the fault diagnosis system, the system's actual output and the SVR model output can be obtained subtraction residuals. Taken RBF as a SVR kernel function, under normal circumstances acquisition system simulation data as training samples.

Figure 6 and Figure 7 shows, using the trained SVR fitting effect on system objects. In which the blue curve in Figure 7 curve fitting, and the points on the curve of training sample points for the SVR.

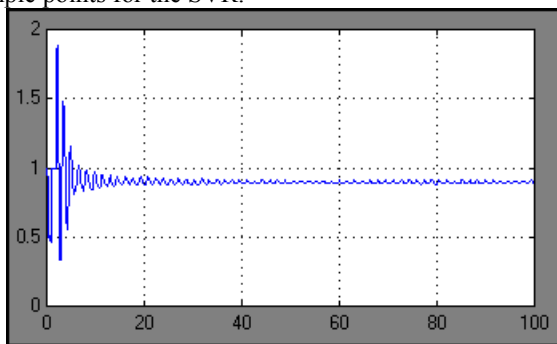


Figure 6 SVR fitting curve

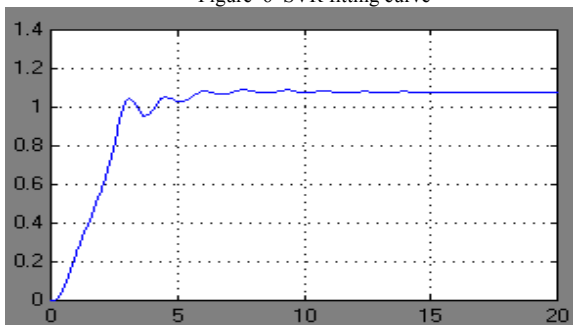


Figure 7 Output curve of Fault Diagnosis

For a given system, training of support vector machine, and further diagnosis results are as follows: Figure 7 is used for fault diagnosis of SVR output residual curve, (a) Figure is a sensor failure occurs when the output bias residual curve; (b) plans is the impact of interference in the event of the sensor output residual curve. It can be seen from the figure, when the fault occurs, the system will be significant changes in residuals, this fault diagnosis system based on SVR will be able to timely detect the fault.

V. SUMMARIZES

The proposed SVR-based fault diagnosis, simulation results show that: SVR can effectively simulate the dynamic characteristics of the sensor system can track the sensor output signal, and when the sensor occurs failure it can timely and accurate diagnosis of sensor fault. Compared with the traditional neural network, with no pre-established format of the actual system identification, SVR can avoid falling into local optimal solution, can solve nonlinear problems and higher order problem, very suitable for solving small sample problems, etc.

ACKNOWLEDGMENT

This paper is supported by Education of Guang-xi Province Scientific Research Foundation(No.: 200911LX131).

REFERENCES

- [1] ZHANG Zhen; Weng SHI Lie; Wang Yong-hong. Wavelet analysis in fault diagnosis in gas turbine applications [J]. Power Engineering, 2006,26 (2) :245-249
 - [2] HU Liang-mou; CAO Ke-qiang. Based on SVR fault diagnosis of nonlinear systems [J]. Mechanical Science and Technology, 2010 (2) :237-239
 - [3] Alex J. Smola, Bernhard Scholkop. A Tutorial of Support Vector Regression, 2003, 9.
 - [4] DI Yong-jie; SHANG Xue-Lian; HAN Pu.. SVR Sensor Fault Diagnosis in the simulation [J]. System Simulation, 2004, 16 (6) :1257-1259
- SHEN Xian-hao; Jinwei Min. Gong Lei pass. based on DNC remote diagnostic network environment and monitoring system [J]. Computer Engineering, 2007 (13): 147-152