Research on a fault diagnosis method based on improved wavelet network

Xueyan Chen, Shouchuang Sun

Department of Mechanical Engineering, Inner Mongolia University for the Nationalities, Tong Liao, 028000, China

Abstract. We improve the structure of wavelet neural network, and add a hidden layer, where different types of parameters entered to the network are classified and integrated. Their total impact and effects on the network are considered, in order to avoid the issue that network convergence is influenced due to mutual interference of various mixed parameters when inputting them to network structure with many variables. We use improved wavelet network as a nonlinear observer for the fault detection system, which can increase the speed of the network training and facilitate the implementation of online fault detection. In recent years, fault detection method based on neural network has aroused great attention. In the case that a non-linear dynamic system model is known, neural network can be used as the output observer for the system's real-time fault detection, and its essence is that neural network is used to establish the input - output model of non-linear dynamic system [1-3]. After fully trained, neural network observer can give an estimate of the output system under normal circumstances, and then the difference between the estimated output and the actual output is used as residuals to detect a fault. When the magnitude of the residual exceeds a predetermined threshold, system failure can be determined. Wavelet neural network is a very before-efficient fault detection method. In this article existing wavelet network is improved and the fault detection method with nonlinear observer based on wavelet neural network is proposed, whose feasibility is confirmed by computer simulation.

Keywords: fault diagnosis method, wavelet network, dynamic system model.

1. Improvement of wavelet network structure used as nonlinear observer

Large-scale networks have a lot of connection weights, which is not conducive to training and convergence of network. And in the network structure with many input variables, it is difficult to reflect the contribution of a certain parameter on the network output [4-5]. In addition, due to mutual interference network convergence will inevitably be influenced by various mixed parameters, which affect online real-time training and practical application. For a dynamic process, input values of a parameter at different historical moments have different degree of impact on the current output of the system, but there is a certain inherent link of its variation. Wavelet network is used to construct nonlinear observer, and the input nodes of wavelet network are control signals at different delay times and output signals of nonlinear system with the two kinds of input parameters. The following improvements of wavelet network structure are made based on the above analysis, shown in Figure 1.

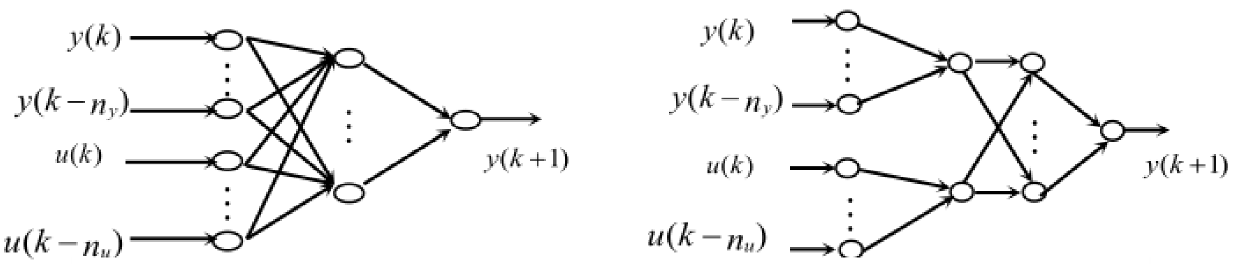

Figure 1 improvement of wavelet network structure used as nonlinear observer

As considering the establishment of the system of wavelet network model, a hidden layer is added, then the roles of the same parameters at different delay time are imposed different weights, and an overall effect of this parameter on the network in the first added hidden layer is integrated. Thus, the improved network weights are greatly simplified, so the training speed of the network can be improved.

2. Training algorithms and steps

2.1 Training algorithm

This paper adopts a simple adaptive step method: firstly we set an initial step. If the error function becomes larger after an iteration, step will be multiplied by a constant l less than 1, then iteration continues; on the contrary, step will be multiplied by a constant g greater than 1, and iteration continues. Compared with the usual adaptive change of step, there is no need for a new step to recalculate again along the original direction when the value of the error function becomes larger. This is because $\Delta\vec{r} = \lambda\nabla E$ (λ is called step size), namely vector direction of $\Delta\vec{r}$ is always same with negative gradient direction. Although the error function value may increase after an iteration, after step decreases at the corresponding iterative point, the corresponding error function is still reduced in the negative gradient direction while the error function is not recalculated along the original direction. The result is that the algorithm searches carefully repeatedly several times in the vicinity of local minima, until the network parameters are found which meet the requirements. Then a basic trick is how to select the appropriate constants g and l. Recommendation: g should be slightly larger than l which should be far less than 1.

2.2 Training step

Specific steps of the algorithm can be summarized as follows:

step1: Initialize the network parameters: assign the initial value to the wavelet stretch factor a_k, translation factors b_k, network connection weights w_{km}, w_{nk}, w_{ml_m}, and the learning rate η ($\eta > 0$) and the momentum factor λ ($0 < \lambda < 1$), and make the input sample counter m = 1;

step2: Input learning samples and corresponding desired output Y_n^p;

step3: Calculate output of hidden layer and output layer with (2) to (4);

step4: Calculate error by (1) and calculate gradient vector with (8) to (13);

step5: m=m+1, if m<P, then turn to step4; otherwise, calculate the cost function E; modify network parameters with (5) to (7), and calculate ΔE $(\Delta E = E(w_n) - E(w_{n-1}))$; if $\Delta E < 0$, then set $\eta = \eta \times 1.25$, otherwise, set $\eta = \eta \times 0.8$;

step6: When the cost function E is less than a predetermined value ε ($\varepsilon > 0$), then stop network learning. Otherwise, reset m = 1; turn to step2 and relearning.

3. The simulation results and analysis

Taking the following process model as simulation object:

$$y(t+1) = 0.75y(t) + 0.11y^2(t) - 0.15y(t-1) + 0.953u(t) - 0.862u(t-1)y(t) \quad (14)$$

System input u is control input. Nonlinear function fitting is made respectively by standard wavelet network with momentum adaptive variable step size and wavelet network with improved structure. Sequence value of $k = 1 \sim 75$ is taken as the training samples; the standard wavelet network structure is taken as 4-8-1; improved wavelet network structure is taken as 4-3-8-1; momentum factor $\lambda = 0.015$; accuracy $\varepsilon = 0.005$. The relation curves between training error and the number of steps of standard and improved wavelet network are respectively shown in Figs. 2 and 3.

From figure 2, it can be seen that the standard wavelet network diagram barely met the requirements when the step number is 145 steps, but the process is less obvious, because convergence speed is very slow after the 52 steps; and from figure 3, with structure improved, wavelet network diagram reaches the predetermined accuracy after only 32 steps. It shows that Convergence rate of wavelet network with improved structure has improved significantly than the one of standard wavelet network.

Fault detection and isolation consists of residuals generation and residuals judgment. When there is no fault in the system residuals are only caused by nonmodeling noise and disturbance, and its amplitude is close to zero. When a fault occurs, the output of residuals deviates from zero according to certain rules. Improved neural network is used to establish a nonlinear observer for the fault detection system. After fully trained, improved neural network observer can give an estimate of the output system under normal circumstances, and then the difference between the estimated output and the actual output is used as residuals to detect a fault. When the magnitude of the residual exceeds a predetermined threshold, system failure can be determined. Evaluation rules for determining residuals is:

$$G_1 : |e(t+1)| > \sigma \quad , \quad G_0 : |e(t+1)| < \sigma \quad e(t+1) = y(t+1) - \hat{y}(t+1/t) \tag{15}$$

Where: G_0 indicates that there are no fault in the system; G_1 indicates that there are faults in the system.

σ Is pre-set fault threshold, which is set as $\sigma = 0.05$ according to actual working conditions?

Assuming setting the system sudden failure at t = 89, some parameters of system configuration are mutated to 2 times the original, then the system process model (14) becomes:

$$y(t+1) = (0.75\times 2)y(t) + 0.11y^2(t) - 0.15y(t-1) + (0.953\times 2)u(t) - 0.862u(t-1)y(t) \tag{16}$$

Now start the wavelet neural network observer for output forecast and calculation of the output residuals, shown in Figure 4.

From the simulation curve, $e(t)$ exceeds the threshold at t = 89, accordingly the system failure can be determined.

We improve the basic network structure of wavelet neural network and use it as a nonlinear observer for system fault detection. Improved network input layer structure is with a more intuitive and physical sense, with the connection weights greatly streamlined, which is very beneficial to improve network speed and implement online testing.

Acknowledgment

This work is supported by the study fund of HIT at The scientific research fund project of Inner Mongolia University for the Nationalities (NMD1323).

Reference

[1] Lv Chaoxia, xiao-bei wu wei-li hu. Used for function approximation of wavelet networks [J]. Systems engineering and electronics, 2000, 22 (5).

[2] ren-bin Yang, Chen Zongmu. Computer control system fault diagnosis. Based automation. 2000, 7 (3) : 35 and 36 [3] Szu H H,Telfer B,Kadambe B. Neural Network Adaptive Wavelets for signal Representation and Classification [J]. Optical Engmeering,1992,31(9).

[4]Qinghua Zhang .Using Wavelet Network in Nonparametric Estimation. IEEE Trans, On Neural Network,1997,8(2),227-236.

[5] Jun Zhang, G G Walter, Yubo Miao, Wan Ngai Wayne Lee. Wavelet Neural Networks for Function Learning. IEEE Trans, On Signal Processing, 1995, 43(6) 1485-1497.