Journal of System Simulation

Volume 31 | Issue 11 Article 37

12-13-2019

# Fault Prediction of Satellite Attitude Control System Based on Neural Network

Xiaofan Meng
*1. School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China; ;*

Song Hua
*1. School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China; ;2. Science and Technology on Space Intelligent Control Laboratory, Beijing 100190, China;*

Follow this and additional works at: https://dc-china-simulation.researchcommons.org/journal

Part of the Artificial Intelligence and Robotics Commons, Computer Engineering Commons, Numerical Analysis and Scientific Computing Commons, Operations Research, Systems Engineering and Industrial Engineering Commons, and the Systems Science Commons

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

## Fault Prediction of Satellite Attitude Control System Based on Neural Network

### Abstract

Abstract: A new method based on *BP neural network(BPNN)*, *wavelet neural network(WNN)* and *wavelet decomposition-LSTM(wLSTM)* network is proposed for *predicting faults* in the satellite attitude control system. Normal satellite attitude data is used to train BPNN which is used as the standard model of satellite attitude control system. The real-time attitude residuals is obtained by subtracting the BPNN output attitude angle from the real-time data of satellite attitude. The time series of the residuals are used to build WNN and wLSTM models to predict the faults of satellite attitude control system. A conclusion is given according to comparing the WNN and wLSTM that both the fault prediction methods can precisely predict the fault and the wLSTM model predicts more accurately because LSTM network can selectively retain the characteristics of input data. At the same time, it also provides a novel method for the prediction of complex system.

### Keywords

fault prediction, satellite attitude control system, BPNN, WNN, wLSTM

## Recommended Citation

Meng Xiaofan, Song Hua. Fault Prediction of Satellite Attitude Control System Based on Neural Network[J]. Journal of System Simulation, 2019, 31(11): 2499-2508.

This paper is available in Journal of System Simulation: https://dc-china-simulation.researchcommons.org/journal/vol31/iss11/37

第 31 卷第 11 期
2019 年 11 月

系统仿真学报©
Journal of System Simulation

Vol. 31 No. 11
Nov., 2019

# Fault Prediction of Satellite Attitude Control System Based on Neural Network

*Meng Xiaofan*[1], *Song Hua*[1,2]

(1. School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China;
2. Science and Technology on Space Intelligent Control Laboratory, Beijing 100190, China )

**Abstract:** A new method based on *BP neural network(BPNN)*, *wavelet neural network(WNN)* and *wavelet decomposition-LSTM(wLSTM)* network is proposed for *predicting faults* in the satellite attitude control system. Normal satellite attitude data is used to train BPNN which is used as the standard model of satellite attitude control system. The real-time attitude residuals is obtained by subtracting the BPNN output attitude angle from the real-time data of satellite attitude. The time series of the residuals are used to build WNN and wLSTM models to predict the faults of satellite attitude control system. A conclusion is given according to comparing the WNN and wLSTM that both the fault prediction methods can precisely predict the fault and the wLSTM model predicts more accurately because LSTM network can selectively retain the characteristics of input data. At the same time, it also provides a novel method for the prediction of complex system.

**Keywords:** fault prediction; satellite attitude control system; BPNN; WNN; wLSTM

## 基于神经网络的卫星姿控系统故障预测

孟小凡[1]，宋华[1,2]

(1. 北京航空航天大学自动化科学与电气工程学院，北京 100191；2. 空间智能控制技术实验室，北京 100190)

**摘要：**针对*卫星姿控系统*时间序列故障预测问题，给出了 *BP 神经网络和小波神经网络、小波分解-LSTM 网络相结合的故障预测方法*。利用卫星正常运行时的数据训练 BP 神经网络，将其作为系统的标准模型，对卫星实时输出和标准模型输出之间的残差建立小波神经网络和小波分解-LSTM 故障预测模型，并进行仿真对比分析。结果表明，2 种方法均能准确的对故障进行预测，由于基于小波分解-LSTM 的方法对残差序列进行了多级小波分解，还利用了 LSTM 网络能选择性保留输入数据的特点，因此预测更准确，性能更优。

**关键词：**故障预测；卫星姿态控制系统；BP 神经网络；小波神经网络；小波分解-LSTM 网络
中图分类号：V249.32　　文献标识码：A　　文章编号：1004-731X (2019) 11-2499-10
DOI: 10.16182/j.issn1004731x.joss.19-FZ0351E

## Introduction

With the rapid development of science and technology, more and more aerospace equipments with different functions and excellent performance have emerged. For an important aerospace device, it is essential to maintain safety, reliability and maintainability of satellites. The feedback channel of satellite attitude control system composed of attitude sensors controls the attitude of the satellite by

Received: 2019-05-30　　Revised: 2019-07-19;
Fundation: Science and Technology on Space Intelligent Control Laboratory (HTKJ2019KL502010);
Biography: Meng Xiaofan(1995-), Female, Hebei, Master, Research direction is satellite attitude control system fault prediction based on deep learning neural network.



Published by Journal of System Simulation, 2019

第 31 卷第 11 期 系统仿真学报 Vol. 31 No. 11
2019 年 11 月 Journal of System Simulation Nov., 2019

changing the control torque applied to the attitude controller[1]. Due to the complex structure and the special operating environment, it is difficult to ensure that there is no failure during satellite work. It is hoped that the fault information can be known in advance before the damage to the satellite appears, so the fault prediction is proposed. As a typical closed-loop feedback control system with multiple sensors, satellite system has complex structure, poor working environment, unknown disturbances and uncertain factors, so it is difficult to obtain its precise mathematical model. In recent years, the theory of deep learning neural network has developed rapidly. Many researchers have achieved fault prediction by building a neural network model to fit the standard input and output of nonlinear complex systems. Liu[2] proposes a long-short-term memory (LSTM) method based on the historical fault data of complex systems to predict time series fault and verifies the superiority of LSTM model in prediction. However, the model will learn the noise characteristics too much when the signal-to-noise ratio is small, which will lead to deviations in the prediction results. Xiao[3] in his article uses the low-frequency signal that is close to the original signal component obtained by wavelet decomposition as the input of BP neural network. The results show that the prediction accuracy of wavelet BP neural network is higher than that of the BP neural network.

In this paper, a new method based on BP neural network[4] is proposed to fit the standard model of the satellite attitude control system. The study uses two time series fault prediction model named wavelet neural network[5] and wavelet decomposition-LSTM[6] network for predicting faults of satellite system because of the feature extraction ability of wavelet decomposition. By comparing the experimental results of the two methods, the wLSTM fault prediction model has better performance than wavelet neural network due to the strong adaptive analysis ability of LSTM network for time series data.

## 1 Satellite attitude control system

It is difficult to detect and isolate faults of the satellite attitude control system because it is a typical closed-loop control system with the complexity and nonlinearity. In this paper, a three-axis stabilized satellite is considered.

The satellite actuator consists of three orthogonal flywheels, and the attitude measuring device is a combination of star sensors and gyros. The three-axis stabilized satellite kinematics equations are as follows[2].

$$\begin{bmatrix} \dot\phi \\ \dot\theta \\ \dot\psi \end{bmatrix} = \begin{bmatrix} \omega_x\cos\theta + \omega_z\sin\theta + \omega_0\sin\psi \\ \omega_y - \tan\varphi(\omega_z\cos\theta - \omega_x\sin\theta) + \omega_0\cos\psi/\cos\varphi \\ (\omega_z\cos\theta - \omega_x\sin\theta - \omega_0\sin\varphi\cos\psi)/\cos\varphi \end{bmatrix} \tag{1}$$

where $\boldsymbol{\omega}=[\omega_x,\omega_y,\omega_z]^T$ is the projection of the three angular velocities of the satellite body coordinate system relative to the earth-centered inertial reference frame in the satellite body coordinate system. Assume that the rotational speed of the satellite orbit system with respect to the inertial system is $(0,-\omega_0,0)$. The three attitude angles of the satellite with respect to the tracking system are the roll angle $\varphi$, pitch angle $\theta$ and yaw angle $\psi$. The rotation angular velocity of the satellite relative to the orbital coordinate system is $\begin{bmatrix}\dot\phi & \dot\theta & \dot\psi\end{bmatrix}$.

The three-axis stabilized satellite dynamic equations are as follows[2]

$$\begin{cases} J_x\dot\omega_x + (J_z - J_y)\omega_y\omega_z = T_x \\ J_y\dot\omega_y + (J_x - J_z)\omega_z\omega_x = T_y \\ J_z\dot\omega_z + (J_y - J_x)\omega_x\omega_y = T_z \end{cases} \tag{2}$$



https://dc-china-simulation.researchcommons.org/journal/vol31/iss11/37
DOI: 10.16182/j.issn1004731x.joss.19-FZ0351E

where $\boldsymbol{T}$=[$T_x$, $T_y$, $T_z$]$^\mathrm{T}$, the control torques $\boldsymbol{J}$=[$J_x$, $J_y$, $J_z$]$^\mathrm{T}$, the satellite inertia moment..

Most satellite attitude control systems use PID control rate to control satellite attitude. In this paper, a PD control rate is used and the three-axis command torques[7] are

$$\begin{cases} T_x = -k_{px}(\varphi + k_{dx}\dot{\varphi}) \\ T_y = -k_{py}(\theta + k_{dy}\dot{\theta}) \\ T_z = -k_{pz}(\psi + k_{dz}\dot{\psi}) \end{cases} \tag{3}$$

where $k_{px}$, $k_{py}$, $k_{pz}$ denote the proportionality coefficients of the controller and $k_{dx}$, $k_{dy}$, $k_{dz}$ are the differential coefficients of the controller.

The control objective of the ground-three-axis stabilized satellite is to keep the three-axis attitude angle stable near zero. At this time, the reference input of the controller is [$\varphi_r$,$\theta_r$,$\psi_r$]=[0,0,0]. In the satellite attitude control simulation, the fourth-order Runge-Kutta is used to solve the kinematics and dynamics equations. The parameters used in the simulation are as follows[8]. The inertia of the three axes of the satellite are

$$\begin{cases} I_x = 1\,849.376\,5\ \mathrm{kg\cdot m^2} \\ I_y = 1\,435.234\ \mathrm{kg\cdot m^2} \\ I_z = 2\,278.882\,4\ \mathrm{kg\cdot m^2} \end{cases} \tag{4}$$

The initial attitude angle is [$\psi$, $\varphi$, $\theta$]$^\mathrm{T}$=[0.1 0.1 –0.1]$^\mathrm{T}$ and the orbital angular velocity of the orbital coordinate system relative to the geocentric inertial system is $\omega_0$=0.001 rad/s. As mentioned earlier, the PD control rate is used in this paper, so the controller parameters are

$$\begin{aligned} k_p &= [20.1637\quad 20.21\quad 20.1637] \\ k_d &= [25.51\quad 22.41\quad 28.31] \end{aligned} \tag{5}$$

The disturbance torque related parameters use

$$\begin{aligned} T_{dx} &= A_x\sin\omega_0 t, A_x = 1.4\times10^{-5}\ \mathrm{N\cdot m} \\ T_{dy} &= A_y\sin\omega_0 t, A_y = 1.5\times10^{-5}\ \mathrm{N\cdot m} \\ T_{dz} &= A_z\sin\omega_0 t, A_z = 1.6\times10^{-5}\ \mathrm{N\cdot m} \end{aligned} \tag{6}$$

Where $A_x$, $A_y$, $A_z$ are the amplitudes of the three-axis interference force, $\omega_0$=0.01 rad/s.

According to the above parameters, the attitude angle simulation diagram of the satellite attitude control system during normal operation is described as Fig. 1.

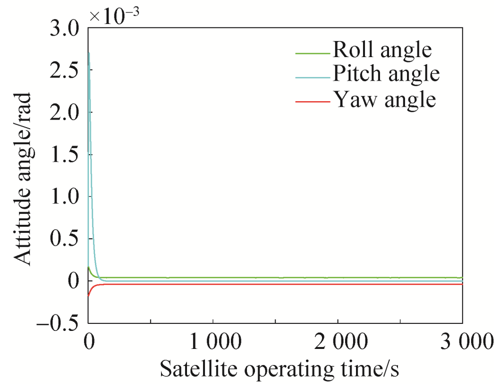


Fig. 1　Simulation results of satellite attitude angle

It can be seen from the Fig. 1 that although the attitude angles of the satellite attitude control system are biased at the initial stage, the attitude angles are gradually stabilized with the adjustment of the PD controller, and the three-axes satellite can maintain stable to the ground.

## 2　Satellite attitude control system fault prediction

The fault prediction idea of satellite attitude control system is training the neural network with the input and output data of the satellite normal operation to obtain the standard model of the system,then injecting the fault into the satellite system to obtain the fault data, and making the difference between the fault data and normal data to get the residual. The fault can be predicted by establishing a residual prediction model. Fig. 2 illustrates the fault forecasting process of three-axis stabilization satellite.

第 31 卷第 11 期 系统仿真学报 Vol. 31 No. 11
2019 年 11 月 Journal of System Simulation Nov., 2019

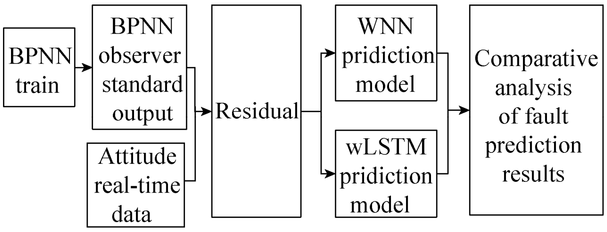


Fig. 2 Fault prediction block diagram of satellite attitude

## 2.1 Residual acquisition based on BP neural network

The satellite attitude control system model has the characteristics of nonlinearity and complexity. Because BP neural network has excellent learning ability, dynamic adaptability, and strong fault tolerance, it is often used to fit nonlinear system models. Theoretically, in the case of reasonable network parameters and structural design, the BPNN with a single hidden layer can complete any m-dimensional to n-dimensional mapping. So the BPNN is trained to get the satellite attitude control system standard model in this study.

BPNN signal propagation consists of two parts: forward propagation and back propagation. It is assumed that a BP neural network has a total of $q$ layers without the input layer and the number of neurons in the input layer, hidden layer and output layer respective are $n^{[0]}$, $n^{[1]}$,……, $n^{[q]}$, then the $i$-th layer forward propagation calculation formulas are expressed as

$$z^{[i]} = w^{[i]}a^{[i-1]} + b^{[i]} \tag{7}$$

$$a^{[i]} = \sigma(z^{[i]}) \tag{8}$$

where, $a^{[i-1]}$ is the ($i$–1)-th hidden layer output and linear calculation for $a^{[i-1]}$ to get $z^{[i]}$. The dimension of the weight matrix $w^{[i]}$ is ($n^{[i]}$, $n^{[i-1]}$) and the bias matrix $b^{[i]}$ is ($n^{[i]}$, 1). The input matrix of input layer is recorded as $a^{[0]}$, and then the activation function $\sigma(\cdot)$ is used to perform a nonlinear operation on $z^{[i]}$ to obtain the output of the $i$-th hidden layer $a^{[i]}$. Usually the activation function selects the Tanh function or the Sigmoid function. The output of the output layer is showed as

$$\hat{y} = w^{[q]}a^{[q-1]} + b^{[q]} \tag{9}$$

If the output layer does not get the desired output, the error information obtained by the loss function is backpropagated continuously to correct the weight and bias of the neural network until the loss is minimal. The loss function of this paper selects the mean square error function

$$J = \sum_{j=1}^{n^{[q]}} (\hat{y}_j - y_j)^2 \tag{10}$$

so the back propagation calculation formulas in $i$-th layer are

$$da^{[i]} = \frac{\partial J}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial a^{[q-1]}} \times \frac{\partial a^{[q-1]}}{\partial z^{[q-1]}} \times \frac{\partial z^{[q-1]}}{\partial a^{[q-2]}} \times \cdots\cdots \times \frac{\partial a^{[i+1]}}{\partial z^{[i+1]}} \times \frac{\partial z^{[i+1]}}{\partial a^{[i]}} \tag{11}$$

$$dz^{[i]} = da^{[i]} \times \sigma^{[i]'}(z^{[i]}) \tag{12}$$

$$dw^{[i]} = dz^{[i]} \times a^{[i-1]} \tag{13}$$

$$db^{[i]} = dz^{[i]} \tag{14}$$

Then the corrected weights and bias are

$$w^{[i]} = w^{[i]} - \alpha \times dw^{[i]} \tag{15}$$

$$b^{[i]} = b^{[i]} - \alpha \times db^{[i]} \tag{16}$$

For the multi-input and multi-output satellite attitude control system of this paper, the input adopts three-dimensional control torques and the output adopts three-dimensional star sensor measurement angles. The BPNN observer with one hidden layer is established for three-dimensional input and output, so the number of neurons in the input and output layers is both three. The activation function of the hidden layer is selected as the Sigmoid function and the learning rate $\alpha$ is 0.01. The weight and bias are corrected by the Stochastic Gradient Descent[9](SGD), and the number of hidden layer nodes can be expressed as

$$m = \sqrt{n + l} + \alpha \tag{17}$$



https://dc-china-simulation.researchcommons.org/journal/vol31/iss11/37
DOI: 10.16182/j.issn1004731x.joss.19-FZ0351E 4

$$m = \log_2^n \tag{18}$$

$$m = \sqrt{nl} \tag{19}$$

Where $m$, the number of hidden layer nodes, $n$, the number of input layer nodes, $l$, the number of output layer nodes between 1-10, so the range of the number of hidden layer nodes is determined to be 4-13. According to equations (20) and (21), the mean square error(MSE) and the mean absolute percentage error(MAPE) are used as indicators to evaluate the quality of the neural network model.

$$\text{MSE} = \sum_{i=1}^{m}(y - \hat{y})^2 \Big/ n \tag{20}$$

$$\text{MAPE} = \frac{1}{n}\sum\frac{|y - \hat{y}|}{y} \times 100 \tag{21}$$

In order to ensure the reliability of the results, the neural network is trained to average for 10 times, and the MSE and MAPE can be calculated between BPNN output and the actual output of the star sensors. The train result is shown in Tab. 1.

Tab. 1　MSE and MAPE under different neurons

| Number | MSE /rad$^2$ | MAPE |
|---|---|---|
| 4 | 3.879 4e-14 | 0.046 4 |
| 5 | 2.368 9e-13 | 0.055 7 |
| 6 | 2.338 8e-13 | 0.137 4 |
| 7 | 2.918 5e-14 | 0.028 8 |
| 8 | 5.966 2e-14 | 0.028 7 |
| 9 | 1.386 6e-13 | 0.047 0 |
| 10 | 2.715 1e-13 | 0.041 6 |
| 11 | 2.362 1e-13 | 0.083 4 |
| 12 | 2.281 7e-13 | 0.138 7 |
| 13 | 2.874 5e-13 | 0.134 1 |

From the order of magnitude of MSE and MAPE, it can be seen that the neural network with the number of neurons 7 and 8 has the best performance. In order to select the optimal network node, the neurons network with hidden layer nodes 7 and 8 is trained again. The results show that the network performance is optimal when the number of neurons is 7. The control torques and angles of 1 000~2 500 s are learned by using the trained BP network to obtain the standard output $\hat{y}$. The difference between the real output $y$ and the standard output $\hat{y}$ is used as the residual, expressed as $e(t)$.

$$e(t) = \sum_{i=1}^{3}(\hat{y}_i - y_i)^2 \tag{22}$$

Fig. 3 illustrates the error between the standard output and the actual output under the condition of no fault. The residual value of the observation is greater than 0 due to the square operation, and the magnitude of the residual is about 1e-12, So the BPNN observer is effective.

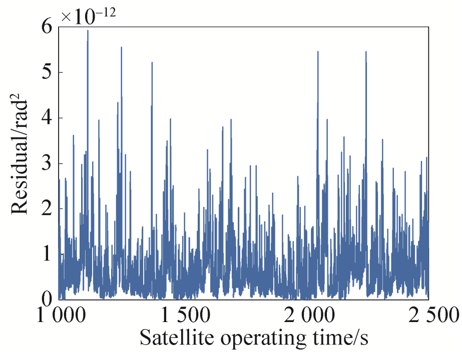


Fig. 3　BPNN observation residuals without fault

The output torques of the flywheel are used as the input of the BPNN observation. The fault of friction torque linearly increases is injected into the x-axis flywheel in 1 600 s. The fault amplitude changes according to $2\times10^{-7}\times(t-1\ 600)$ N·m, and the residual figure in this fault mode is shown in Fig. 4.

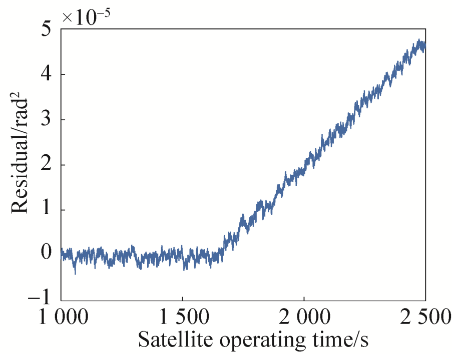


Fig. 4　BPNN observation residuals with linear fault

## 2.2 Fault prediction based on wavelet neural network

Wavelet neural network is a combination of wavelet transform and neural network. It is a neural network based on BPNN framework and using existing wavelet function to replace Sigmoid function. Wavelet transform can effectively extract the local features of the signal and the neural network has the characteristics of strong self-learning ability and good adaptability. Therefore, the wavelet neural network has good fault tolerance and high precision, and has been widely used in the field of fault prediction. The residual data obtained in the section 2.1 is subject to large noise and interference. Because the wavelet neural network is effective in approximation function and the denoising ability is strong, the wavelet neural network is selected for fault prediction to ensure the accuracy of prediction in this study.

Wavelet neural network has two structures[10]. One is a loose wavelet neural network, which first performs wavelet decomposition[11] on the input signal and then inputs the decomposed signal into other neural networks. The other is a fusion wavelet neural network, whose structure is based on BPNN and the transfer function selects the wavelet function that the corresponding weights and bias of BPNN are replaced by the scaling factor and translation factor of the wavelet function.

In this section, the fusion wavelet neural network is used. Its structure is the same as the BPNN in the previous section, but the excitation function is the Morlet wavelet function, and the signal propagates both forward and back. In the forward propagation process, the output calculation formula of the $i$-th hidden layer is

$$a^{[i]} = h\left(\frac{w^{[i]}a^{[i-1]} - b^{[i]}}{d^{[i]}}\right) \tag{23}$$

Where $a^{[i-1]}$ is the output of the ($i$–1)-th hidden layer. $d^{[i]}$ and $b^{[i]}$ respective are the scaling factor and the translation factor of the wavelet function $h(\cdot)$, $h(\cdot)$, the Morlet function, is defined as

$$y = \cos(1.75x)e^{-x^2/2} \tag{24}$$

The calculation formula for the output layer is

$$\hat{y} = w^{[q]}a^{[q-1]} \tag{25}$$

In this section, the gradient descent method is used to correct $w^{[i]}$, $a^{[i]}$ and $b^{[i]}$ of the Morlet function in the backpropagation process.

The purpose of obtaining the residual is to predict the fault. Because the residual sequence $e(t)$ obtained in the equation (22) is time series data, so the time-residual one-to-one mapping relationship is established. The input of the WNN is defined as time $t$ and the output is residual $e$. Then there is

$$e = f_{WNN}(t) \tag{26}$$

When training the WNN prediction model $WNN_x$, $\{t_{x-m+1},\ldots,t_{x-2},\ t_{x-1},\ t_x\}$ is used as the input of the model, and $\{e_{x-m+1},\ldots,e_{x-2},e_{x-1},e_x\}$ is used as the output of the model. When $WNN_x$ is trained, the next-time variable $t_{x+1}$ is brought into $WNN_x$ to obtain the predicted output $e_{x+1}$. When the real $e_{x+1}$ arrives, the data window rolls forward one step to update the input and output as $\{t_{x-m+2},\ldots,t_{x-1},t_x,t_{x+1}\}$ and $\{e_{x-m+2},\ldots,\ e_{x-1},e_x,e_{x+1}\}$, so the model $WNN_{x+1}$ is obtained after training. The $t_{x+2}$ is brought into $WNN_{x+1}$ to obtain the predicted value $e_{x+2}$. Repeatedly, and finally all residual prediction values are obtained.

## 2.3 Fault Prediction Based on Wavelet Decomposition-LSTM Network

The key problem of time series prediction is to extract as many representative features as possible from the time series and infer the future state of the



https://dc-china-simulation.researchcommons.org/journal/vol31/iss11/37
DOI: 10.16182/j.issn1004731x.joss.19-FZ0351E 6

time series based on the time series features. Wavelet decomposition is a well-known method for capturing time series features in time and frequency domains. We can use them as tools for data preprocessing before training deep model. Wavelet decomposition can decompose a time series into a set of subsequences whose frequencies are arranged from high to low, which is crucial for capturing the frequency factor of deep learning. In recent years, with the development of the concept of deep learning, various types of deep neural network models have been introduced into time series analysis. Some well-known models such as recurrent neural network[12] (RNN) and long short-term memory[13] (LSTM), use memory nodes to simulate the correlation of series of points, most of which belong to the category of time domain methods without using frequency information of time series. Therefore, the wavelet high frequency sub-series decomposed from original time series is input into a separate LSTM model, and signal reconstruction is performed on all LSTM outputs for final prediction. In this way, the frequency information of the wavelet can be seamlessly embedded into the deep learning framework. The wLSTM framework is illustrated in the Fig. 5.

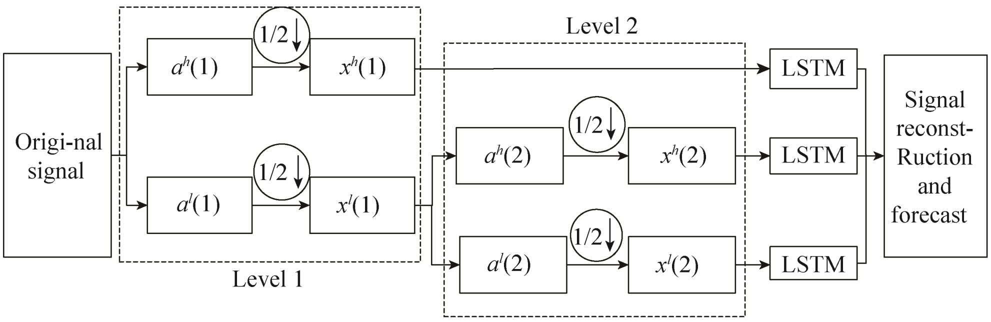


Fig. 5 Wavelet decomposition-LSTM framework

Wavelet decomposition can decompose time series data into high-frequency and low-frequency sub-series to extract time series features. The input time series is defined as $x=\{x_1, x_2,\ldots, x_T\}$, and the low and high frequency sub-series are generated in the $i$-th as $x^l(i)$ and $x^h(i)$. In the $(i+1)$-th level, wLSTM uses a low pass filter $L=\{l_1, l_2,\ldots, l_k\}$ and a high pass filter $H=\{h_1, h_2,\ldots, h_k\}$, $k << T$, to convolution low frequency sub-series obtained by the $i$-th decomposition as

$$\begin{aligned} a_n^l(i+1) &= \sum_{k=1}^{K} x_{n+k-1}^l(i) \times l_k \\ a_n^h(i+1) &= \sum_{k=1}^{K} x_{n+k-1}^l(i) \times h_k \end{aligned} \tag{27}$$

Where $x_n^l(i)$ is the $n$-th element of the low-frequency sub-series of the $i$-th decomposition, and the input sequence is set as $x^l(0)$. The low and high frequency sub-series $x^l(i)$ and $x^h(i)$in the level $i$ are generated from 1/2 down-sampled of the intermediate variable sequences $a^l(i)=\{a_1^l(i), a_1^l(i),\ldots\}$and $a^h(i)=\{a_1^h(i), a_1^h(i),\ldots\}$.

The subsequence $\chi(i)=\{x^h(1), x^h(2),..., x^h(i), x^l(i)\}$ is called the $i$-th decomposition result of $x$ and is also the input of the LSTM deep learning framework. The input data is normalized so that it is mapped into the region of [0, 1] to prevent to cause gradient disappearance or gradient explosion. Each LSTM network will use the hidden layer output calculated at the previous moment as the input of the same hidden layer at the next moment. This loop can pass information from the current step to the next step to avoid long-term dependency problems. Each node of LSTM consists of four parts: input gate, output gate, forgetting gate and Cell [14]. The Cell, core part of the LSTM node, is used to record the state of the hidden layer at the current moment. The other three gates allow the information to pass selectively by effectively adding or deleting information, and they work together to control the operation of the node. The LSTM hidden layer structure diagram is illustrated in the Fig.6.

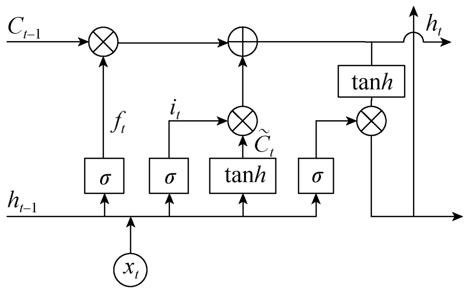


Fig. 6 LSTM hidden layer structure diagram

Where $h_{t-1}$ is the output of the hidden layer at time $t-1$, so the output of the forgotten gate is

$$f_t = \sigma(W_f \times [h_{t-1}, x_t] + b_f) \tag{28}$$

$f_t$ is a value between [0, 1] that determines what information the network can drop. The input gate determines the new information entering into the network, which is expressed as

$$i_t = \sigma(W_i \times [h_{t-1}, x_t] + b_i) \tag{29}$$

$$\tilde{C}_t = \tan h(W_C \times [h_{t-1}, x_t] + b_C) \tag{30}$$

Next, $f_t$ is multiplied by the old information $C_{t-1}$ to discard the information that needs to be discarded, and then add $i_t * \tilde{C}_t$, so a new candidate value $C_t$ is obtained.

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{31}$$

The finally output at the output layer is

$$o_t = \sigma(W_o \times [h_{t-1}, x_t] + b_o) \tag{32}$$

$$h_t = o_t \times \tan h(C_t) \tag{33}$$

## 2.4 Fault prediction based on WNN and wLSTM

The data window length of the wavelet neural network is set to $m$=120, and the sampling time is $\Delta t$=0.1 s. the WNN is trained under the number of one hidden layer. When the number of hidden layer nodes is 6, the training effect is the best. The WNN prediction result is shown in Fig. 7 in the case where the fault amplitude is a $2*10^{-7}*(t-1\ 600)$ N·m. It can be seen from the figure that the wavelet neural network prediction model can follow the original curve well and fit the fault residual.

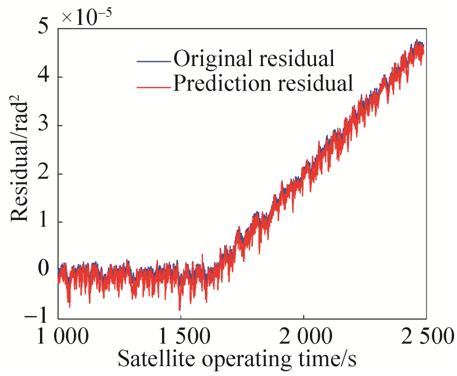


Fig. 7 WNN prediction result

The predicted values of the satellite attitude control system fault residuals are obtained by training the wLSTM network and correcting the coefficients. The filter uses Daubechies 4 wavelet. When the fault amplitude changes according to $2*10^{-7}*(t-1\ 600)$ N·m, the prediction results are shown in Fig. 8. Since the temporal correlation of nodes hidden in the time series is closely related to the frequency, so the wLSTM network prediction curve can follow the original residual well and the prediction curve is more stable, which is more conducive to fault prediction.

From the comparative analysis of the predicted curves in Fig. 7 and Fig. 8, it can be seen that the



https://dc-china-simulation.researchcommons.org/journal/vol31/iss11/37
DOI: 10.16182/j.issn1004731x.joss.19-FZ0351E

smoothness and accuracy of the wLSTM prediction curve are better than WNN. In order to better compare the performance of WNN and wLSTM methods, the optimal hyper-parameters such as learning rate, hidden layer number, and hidden layer neurons and so on of WNN and wLSTM networks trained under $2*10^{-7}*(t-1\ 600)$ N·m fault amplitude are saved. Then, under the condition that the network structure and hyper-parameter of the WNN and wLSTM are unchanged, the weights and bias of the WNN and wLSTM networks are retrained to obtain the MSE and MAPE in different fault amplitudes. By comparing MSE and MAPE, it can be seen from Tab. 2 that the wLSTM network has higher prediction accuracy. This is because LSTM can selectively retain the characteristics of the input data, so the prediction curve can better eliminate the effect of noise on the residual.

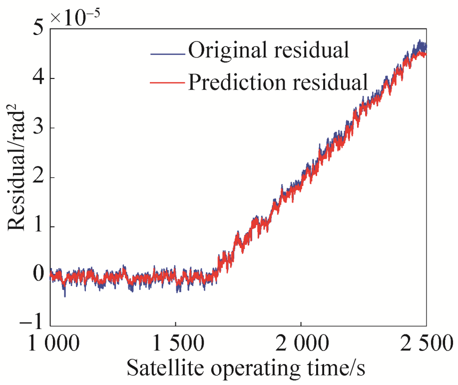


Fig. 8　wLSTM prediction result

Tab. 2　MSE and MAPE under different fault amplitudes

| Fault Amplitudes | wLSTM MSE | wLSTM MAPE | WNN MSE | WNN MAPE |
|---|---|---|---|---|
| $0.5\times10^{-7}$ | 6.824 9e-14 | 0.814 9 | 2.568 6e-12 | 5.048 5 |
| $1\times10^{-7}$ | 8.193 1e-14 | 0.693 6 | 2.399 6e-12 | 3.950 1 |
| $2\times10^{-7}$ | 3.694 5e-13 | 0.814 3 | 2.444 6e-12 | 7.264 1 |
| $3\times10^{-7}$ | 2.409 2e-12 | 1.563 8 | 2.331 5e-12 | 3.267 9 |
| $4\times10^{-7}$ | 1.740 0e-12 | 3.448 3 | 2.236 4e-12 | 5.972 4 |

## 3　Conclusion

This article introduces a fault prediction method for the satellite attitude control system. An actuator fault can affect both the system measurement output and system state. The residuals are obtained by establishing the BPNN for the control torque and attitude angle, and the data window rolling prediction is performed on the residuals by WNN and wLSTM network. The results show that the two prediction methods can predict fault well and the fault prediction method based on wavelet decomposition-LSTM network is more accurate because LSTM can selectively retain the characteristics of the input data. Through the research on the satellite attitude control system fault prediction method, people can know the fault symptom information in advance before the fault hazard appears, and then take measures to prevent it, which is of great significance for improving the safety and stability of the satellite and ensuring the smooth operation of the satellite. At the same time, the fault prediction method of this article can also be extended to other complex systems and has certain universality.

## References:

[1] Qian Yong. Studies on the Attitude Determination and Control System of High Precision Three-axis Stabilized Satellite[D]. Xi'an: Northwestern Polytechnical University, 2002.

[2] Liu Xinyue. Research on Fault Diagnosis Method of Satellite Control System Based on Deep Learning[D]. Bejing: BeiHang University, 2018.

[3] Xiao C, Jiao Z, Sun J T, et al. Fault prediction of variable pitch system of wind turbine based on wavelet BP neural network[J]. Renewable Energy Resources (S1671-5292), 2017, 35(6): 893-899.

[4] Li Q, Yu J Y, Mu B C, et al. BP neural network prediction of the mechanical properties of porous NiTi shape memory alloy prepared by thermal explosion reaction[J]. Materials Science and Engineering: A (S0921-5093), 2006, 419(1-2): 214-217.

[5] Shen X Y, Zhang F, Lü H T, et al. Prediction of Entering Percentage into Expressway Service Areas Based on

第 31 卷第 11 期
2019 年 11 月
系统仿真学报
Journal of System Simulation
Vol. 31 No. 11
Nov., 2019

Wavelet Neural Networks and Genetic Algorithms [J]. IEEE Access (S2169-3536), 2019, 7: 54562-54574

[6] Wang J Y, Wang Z, Li J F, et al. Multilevel Wavelet Decomposition Network for Interpretable Time Series Analysis[Z]. Computer Science, Statistics, 2018.

[7] Zhang R W. Attitude dynamics and control of satellite orbit[M]. Bejing: Beihang University Press, 1998.

[8] Wang M L. Research on fault diagnosis of satellite control system[D]. Bejing: BeiHang University, 2017.

[9] Shi S H, Wang T Q, Chu X W, et al. A DAG Model of Synchronous Stochastic Gradient Descent in Distributed Deep Learning[C]. 2018 IEEE 24th International Conference on Parallel and Distributed Systems (ICPADS) Parallel and Distributed Systems (ICPADS), 2018 IEEE 24th International Conference on: 425-432 Dec, 2018.

[10] Luo P. Study on wavelet neural network based on back propagation algorithm[J]. Information Technology (S1009-2552), 2012(11): 48-51, 56.

[11] Ingrid Daubechies. 1992. Ten lectures on wavelets. SIAM.

[12] Yara Rizk, Mariette Awad. On Extreme Learning Machines in Sequential and Time Series Prediction: A Non-Iterative and Approximate Training Algorithm for Recurrent Neural Networks[J]. Neurocomputing (S0925-2312), 2018.

[13] Gers F A, Schmidhuber J, Cummins F. Learning to forget: continual prediction with LSTM[C]. Artificial Neural Networks, 1999. ICANN 99. Ninth International Conference on. IET, 2002: 2451.

[14] Wöllmer M, Kaiser M, Eyben F, et al. LSTM-Modeling of continuous emotions in an audiovisual affect recognition framework[J]. Image & Vision Computing (S0262-8856), 2013, 31(2): 153-163.



https://dc-china-simulation.researchcommons.org/journal/vol31/iss11/37
DOI: 10.16182/j.issn1004731x.joss.19-FZ0351E