

Fault diagnosis method of gear based on lifting wavelet packet and combined optimization BP neural network

Shungen Xiao¹, Zexiong Zhang², Mengmeng Song³

^{1,2,3}College of Information, Mechanical and Electrical Engineering, Ningde Normal University, Ningde, People's Republic of China

^{1,2}College of Mechanical and Electrical Engineering, Fujian Agriculture and Forestry University, Fuzhou, People's Republic of China

³Corresponding author

E-mail: ¹xiaoshungen022@163.com, ²1963414543@qq.com, ³544824964@qq.com

Received 21 October 2019; accepted 28 October 2019

DOI <https://doi.org/10.21595/vp.2019.21106>



Copyright © 2019 Shungen Xiao, et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract. Aiming at the problem of weak signal signature recognition of gear faults, a gear fault diagnosis method based on lifting wavelet packet and combined optimization BP neural network is proposed. The initial non-sampling prediction and update operators are calculated by Lagrange interpolation subdivision based on the principle of lifting wavelet, and the adaptive redundancy lifting wavelet packet decomposition and reconstruction algorithm is constructed. The network parameters of the number of hidden layers and the quantity of nodes, initial weights and thresholds of BP neural network are optimized by genetic algorithm (GA). The Levenberg-Marquardt (LM) algorithm is used to improve the search space of the network. Through experimental analysis, the results show that the gear fault diagnosis method proposed in this paper not only has high diagnostic accuracy, but also increase efficiency.

Keywords: gear fault diagnosis, BP neural network, genetic algorithm, Levenberg-Marquardt, lifting wavelet packet.

1. Introduction

The gear is the most important connection and transmission component in the mechanical equipment. If the fault of the gear in the transmission process is found in time, the maintenance and repair time of the equipment can be arranged economically and reasonably to avoid accidents. How to effectively extract the fault features hidden in the gears to achieve efficient and accurate fault classification as well as diagnosis is a hot and difficult problem for current researchers. In order to overcome the problems of slow convergence and falling into local minimum of traditional BP-NN, Wu [1] uses Genetic Algorithm (GA) to optimize the initial weight of BP-NN, and uses GA global search ability to effectively avoid BP-NN is caught in a local minimum problem; Zhang [2] proposed a fault diagnosis of the fan gearbox based on genetic algorithm optimization BP neural network, which effectively diagnoses the gearbox fault diagnosis. However, because GA only optimizes the initial weight of BP-NN to speed up the determination of the search space, the basic BP algorithm is still used in the local optimization process in the search space. Obviously, it still fails to change the slow convergence of BP-NN.

2. Non-sampling lifting wavelet packet algorithm

SWELDENS [3] first proposed the wavelet transform theory of the lifting mode. The decomposition process consists of three steps: segmentation, prediction and update. Since the traditional lifting wavelet or wavelet packet transform is a transform method based on sampling operation, it is easy to cause information component loss and frequency aliasing problem. The non-sampling lifting wavelet packet algorithm proposed in [4-6] effectively solves this problem.

2.1. Non-sampling lifting wavelet packet decomposition algorithm

Firstly, the initial prediction operator and the update operator are obtained by the Lagrange interpolation subdivision principle. Let the initial prediction operator $P = \{p_m\}$, where $m = 1, 2, \dots, N$. Set the initial update operator $U = \{u_n\}$, where $n = 1, 2, \dots, \tilde{N}$. The expression of the s -layer non-sampling lifting wavelet packet prediction operator $P^{[s]}$ and the update operator $U^{[s]}$ as follow [7]:

$$\begin{aligned} P_i^{[s]} &= \begin{cases} p_m, & i = 2^s m, \\ 0, & i \neq 2^s m, \end{cases} & i = 1, 2, \dots, 2^s N, \\ U_j^{[s]} &= \begin{cases} u_n, & j = 2^s n, \\ 0, & j \neq 2^s n, \end{cases} & j = 1, 2, \dots, 2^s \tilde{N}. \end{aligned} \quad (1)$$

Compared with the traditional lifting wavelet packet decomposition process, the non-sampling lifting wavelet packet removes the segmentation link and performs the operation in a non-sampling manner. Let X_{sl} be the first frequency band signal of the original signal X decomposed in the s layer. Each sample of the upper layer signal X_{s-1} is predicted by the adjacent $2^s N$ sample signals by the non-sampling predictor $P_i^{[s]}$, and the predicted difference is the high-frequency detail signal $X_{s(l-1)}$, such as Eq. (2). The low-frequency approximation signal X_{sl} is obtained by updating the detail signal by the non-sampling prediction operator $u_j^{[s]}$, as given by Eq. (3):

$$X_{s(l-1)}(k) = X_{(s-1)(\frac{l}{2})}(k) - \sum_{i=1}^{2^s N} \left\{ p_i^{[s]} X_{(s-1)(\frac{l}{2})} \left[k - 2^{(s-1)}(N+1) + i \right] \right\}, \quad (2)$$

$$X_{sl}(k) = X_{(s-1)(\frac{l}{2})}(k) + \sum_{j=1}^{2^s \tilde{N}} \left\{ u_j^{[s]} X_{s(l-1)} \left[k - 2^{(s-1)}(\tilde{N}+1) + j \right] \right\}. \quad (3)$$

In the Eqs. (2)-(3), $l = 2, 4, 6, \dots, 2s$, $P_i^{[s]}$ is the non-sampling boost wavelet packet prediction operator coefficient of detail signal $X_{s(l-1)}$, and $u_j^{[s]}$ is the non-sampling boost wavelet packet update operator coefficient of approximating signal X_{sl} .

2.2. Non-sampling lifting wavelet packet reconstruction algorithm

The non-sampling lifting wavelet packet reconstruction algorithm is the inverse operation of the above decomposition algorithm, which consists of a recovery update link, a recovery prediction link and a merge link. The recovery update is performed by the signals X_{sl} and $X_{s(l-1)}$ to complete the recovery of the sample sequence $X_{(s-1)(l/2)}^*$, as shown in the Eq. (4). The recovery prediction link is recovered by the signal $X_{s(l-1)}$ and the completed sample sequence $X_{(s-1)(l/2)}^{**}$, as shown in Eq. (5). The merged link is obtained by averaging the signal $X_{(s-1)(l/2)}^*$ and $X_{(s-1)(l/2)}^{**}$ adding, is given us by Eq. (6), and the result is the non-sampling lifting wavelet packet reconstructed signal $X_{(s-1)(l/2)}$:

$$X_{(s-1)(\frac{l}{2})}^*(k) = X_{sl}(k) - \sum_{j=1}^{2^s \tilde{N}} \left\{ u_j^{[s]} X_{s(l-1)} \left[k - 2^{(s-1)}(\tilde{N}+1) + j \right] \right\}, \quad (4)$$

$$X_{(s-1)(\frac{l}{2})}^{**}(k) = X_{s(l-1)}(k) + \sum_{i=1}^{2^s N} \left\{ p_i^{[s]} X_{(s-1)(\frac{l}{2})}^* \left[k - 2^{(s-1)}(N+1) + i \right] \right\}, \quad (5)$$

$$X_{(s-1)(\frac{l}{2})}(k) = \frac{1}{2} \left[X_{(s-1)(\frac{l}{2})}^{**}(k) + X_{(s-1)(\frac{l}{2})}^*(k) \right]. \quad (6)$$

3. Optimize BP-NN using the combination of GA- and LM-based approach

GA is a global probability search algorithm which is derived from the population search strategy and information exchange between population individuals. It is independent of the information of BP neural network using nonlinear optimization gradient descent method [8]. It can optimize BP-NN by using this algorithm. The topology and network parameters improve the convergence speed of the network. LM is an algorithm that combines the descending gradient method and the Newton method to further improve the network optimization efficiency and avoid local minimum problems. Based on GA and LM are combined to perform BP-NN together to provide more accurate diagnosis results for gear failure.

3.1. GA optimizes the topology and network parameters of BP-NN

GA optimizes the topology and network parameters of BP-NN as follows:

Step 1: Assume the number of hidden layers and nodes between the layers in the BP-NN. The number of layers and nodes are encoded to randomly generate N encoded chromosomes.

Step 2: Decode the N encoded chromosomes into corresponding BP-NNs.

Step 3: Train each network separately with different initial weights.

Step 4: Calculate the error function of BP-NN under each code string separately, and use the error function to determine the fitness function of every individual. GA optimizes the network topology to minimize the sum of squared errors in the network output, but GA can only evolve in the direction of increasing the fitness function value. Therefore, the inverse of the squared sum of the BP-NN output error is chosen as the fitness function, as follows Eq. (7):

$$\begin{cases} E_j = \frac{1}{2} \sum_{p=1}^p \sum_{i=1}^i (\hat{y}(p) - y_i(p))^2, \\ Fit = J^{-1} \end{cases} \quad (7)$$

where p is the number of training samples, i is the number of network output nodes, $\hat{y}(p)$ is the expected output value of the network, the current output value of the network, E_j is called the evolution error, and Fit is the fitness value.

Step 5: Calculate the fitness value according to Eq. (7), and select a plurality of individuals with large fitness values to form a male parent.

Step 6: The current generation of populations are manipulated by genetic operators such as crossover and mutation to produce a new generation of populations.

Step 7: Repeat steps Step 2-Step 6 until an individual in the group can meet the network requirements.

3.2. The improved BP-NN theory of LM algorithm

The LM algorithm applies the approximate second-order derivative information, which is a combination of the descending gradient algorithm and the Gauss-Newton algorithm. It has the local characteristics of Newton's method, can generate an ideal search direction near the optimal value, and has the global characteristics of the gradient method. That is, the steps that just started the iteration drop faster. Therefore, the convergence rate is faster and more stable than the falling gradient algorithm [9]. The expression of the iteration of the LM algorithm is:

$$x(k+1) = x(k) + \Delta x, \quad (8)$$

where:

$$\Delta x = -(J^T(x)J(x) + \mu I)^{-1} J^T(x)e(x),$$

$$J(x) = \begin{bmatrix} \frac{\partial e_1(x)}{\partial x_1} & \frac{\partial e_1(x)}{\partial x_2} & \cdots & \frac{\partial e_1(x)}{\partial x_i} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial e_n(x)}{\partial x_1} & \frac{\partial e_n(x)}{\partial x_2} & \cdots & \frac{\partial e_n(x)}{\partial x_i} \end{bmatrix},$$

where $x(k)$ is the vector of the weight and threshold of the k th iteration, $e_i(x)$ is the error of the i th network node, I is the identity matrix, $J(x)$ is called the Jacobian matrix, μ is one greater than zero adaptive adjustment factor.

When $\mu \rightarrow 0$, $\mu I \approx 0$, $\Delta x = -(J^T(x)J(x))^{-1}J^T(x)e(x)$, and Eq. (8) becomes Eq. (9), namely, Gauss Newton algorithm:

$$x(k+1) = x(k) - (J^T(x)J(x))^{-1}J^T(x)e(x). \quad (9)$$

When $\mu \rightarrow \infty$, $J^T(x)J(x) \ll \mu I$, $J^T(x)J(x)$ are negligible, the LM algorithm is transformed into a gradient descent method. The LM algorithm can better achieve the combination of Gauss-Newton method and gradient descent method by adaptively adjusting μ . Practice has shown that the LM algorithm is dozens or even hundreds of times faster than the falling gradient method. At the same time, the LM algorithm is positively defined, and there is always a solution. However, it is not necessary for the Gauss-Newton method $(J^T(x)J(x) + \mu I)^{-1}$ to be full rank. The LM algorithm is better than the Gauss-Newton me

4. Experimental analysis

4.1. Fault signal acquisition

The hardware involved in the experimental device mainly includes QPZZ-II rotating mechanical vibration analysis and fault diagnosis test platform system, one signal conditioning instrument, two data collectors, and several acceleration sensors. The schematic diagram of the experimental platform in Fig. 1. The pinion of the experimental platform is a driving wheel, which relates to the motor shaft. the large gear is a driven wheel, and is connected to the magnetic powder brake through a coupling. The number of teeth of the pinion gear is 55, and the number of teeth of the large gear is 75. The two accelerometers are installed in the horizontal and vertical directions at the large gear on the outer side of the gearbox. The calibration values of the two sensors are 102 mV/g and 99 mV/g, respectively. The measuring point arrangement as shown in Fig. 2.

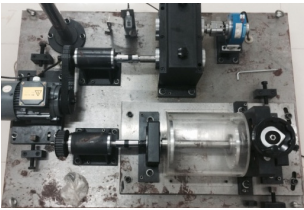


Fig. 1. The experimental platform

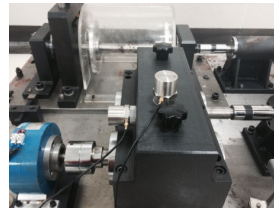


Fig. 2. Sensor test point arrangement

4.2. Diagnosis results

Based on obtaining four kinds of vibration signals: normal signal, gear broken tooth, gear crack and wear, the method proposed in this paper is used to realize the diagnosis of various gear fault types. The specific diagnosis steps are as follows:

(1) Using the redundant lifting wavelet packet method to denoise the broken gear teeth, gear surface cracks and wear signals, as depicted in Fig. 3.

(2) Calculate the energy characteristics of the signals of each frequency band decomposed by

the wavelet packet and normalize them. Fig. 4 is the energy spectrum of the broken teeth.

(3) Establish BP-NN network by using GA optimization BP-NN topology and network parameters.

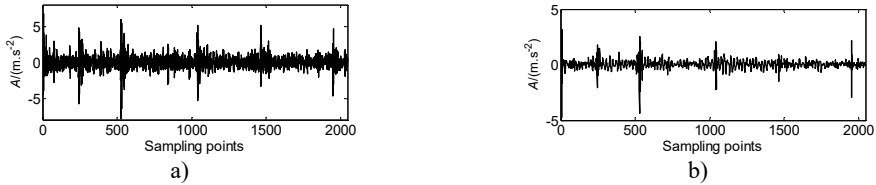


Fig. 3. Broken teeth noise reduction effect: a) the original signal of gear fault; b) the signal denoised of gear fault by redundant lifting scheme packet

The input sample of BP-NN is the fault energy characteristic vector of the gear, and the output is of four fault types: no fault (1,0,0,0), gear tooth breaking fault (0,1,0,0), gear surface Crack failure (0,0,1,0), gear wear (0,0,0,1). Therefore, there are 8 nodes in the input layer and 4 nodes in the output layer. Set the population number to 50, the maximum running algebra to 100, and the generation gap to 0.9. The error reduction curve of the optimization process is in Fig. 5. After the topology of the neural network is established, the initial weight and threshold parameters of the GA optimization network are continuously used. Fig. 6 is the error reduction curve, the error square curve and the fitness curve of the parameter optimization process.

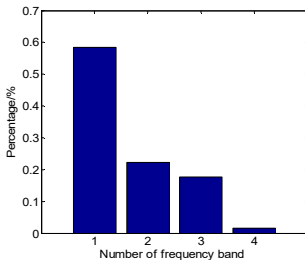


Fig. 4. Signal failure energy spectrum distribution

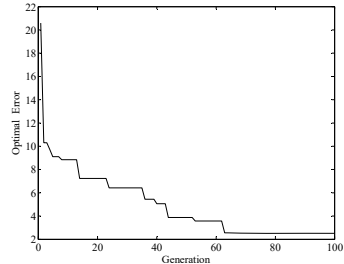
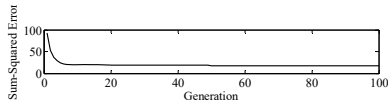
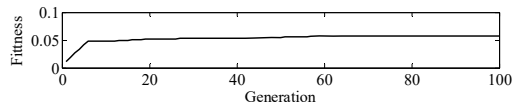


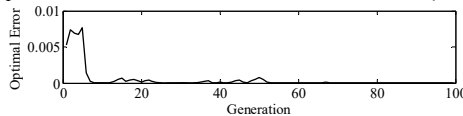
Fig. 5. Error reduction curve



a) Error sum of squares curve



b) Fitness curve



c) Optimization error descent curve

Fig. 6. Error and fitness curve

(4) Using LM algorithm to improve the efficiency of BP-NN search target, set the transfer function of hidden layer and output layer to tansig and purelin respectively, and the minimum mean square error index is set to 10^{-5} .

The results show that the GA-optimized BP-NN needs to iterate 199 steps to reach the set mean square error index, as shown in Fig. 7. 45 cases were successfully diagnosed in 50 test samples, and the cumulative error of the four types of faults was 0.0093. After GA-LM combination optimization, BP-NN only needs to iterate 124 steps to reach the set mean square error index, as shown in Fig. 8, successfully diagnosed 47. For example, the cumulative error of the four fault types is 0.0078. Therefore, the combined optimization BP-NN converges faster, namely, the diagnosis time is shorter and the diagnostic accuracy is higher.

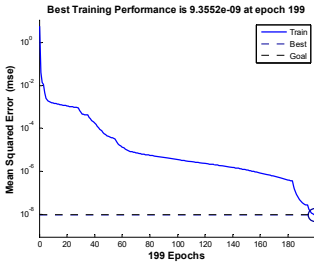


Fig. 7. GA optimization BP-NN training process

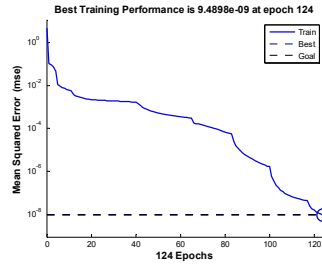


Fig. 8. GA-LM optimization BP-NN training process

5. Conclusions

1) Using the lifting wavelet packet to effectively denoise the gear broken teeth, crack and wear fault signals, and successfully extract the fault energy feature quantity as the input feature vector of BP-NN.

2) Using GA to effectively optimize the hidden layer number, initial weight and threshold of BP-NN, avoiding the traditional method of “experimental trial and error method” to obtain network topology and avoiding the blindness of the traditional random given initial weight values and threshold parameters.

3) Using LM algorithm to improve the search efficiency of BP-NN. The experimental results show that the GA and LM combination optimization BP-NN gear fault diagnosis method has higher efficiency and accuracy.

Acknowledgements

This paper was supported by the following research projects: by the Special Project of Ningde Normal University in 2018 (Grant No. 2018ZX409, Grant No. 2018Q101, Grant No. 2018ZX401) and Research project for Yong, Middle-aged Teacher in Fujian Province (Grant No. JT180601 and Grant No. JT180597). These supports are gratefully acknowledged.

References

- [1] Wu L. Research on Fault Diagnosis Algorithm Based on Genetic Neural Network. Shengyang, 2012.
- [2] Zhang X., Zheng L., Hua L. The fault diagnosis of wind turbine gearbox based on genetic algorithm to optimize BP neural network. Journal of Hunan Institute of Engineering, Vol. 28, Issue 3, 2018, p. 1-6.
- [3] Sweldens W. The lifting scheme: A construction of second-generation wavelet. SIAM Journal on Mathematics Analysis, Vol. 29, Issue 2, 1997, p. 511-546.
- [4] Chen J., Zhang L., Duan L., et al. Diagnosis of reciprocating compressor piston-cylinder liner wear fault based on lifting scheme packet. Journal of China University of Petroleum: Natural Science Edition, Vol. 35, Issue 1, 2011, p. 130-134, (in Chinese).
- [5] Duan C., Li L., He Z. Undecimated wavelet transform based of lifting scheme and its application in fault diagnosis. Journal of Mechanical Strength, Vol. 28, Issue 6, 2006, p. 796-799, (in Chinese).
- [6] An S., LV L., He Y. Fault diagnosis method of rolling bearing based on undecimated wavelet transformation of lifting scheme. Journal of Vibration and Shock, Vol. 28, Issue 1, 2009, p. 170-173, (in Chinese).
- [7] Jiang H., He Z., Duan C. Gearbox fault diagnosis using adaptive redundant lifting scheme. Mechanical Systems and Signal Processing, Vol. 20, Issue 8, 2006, p. 1992-2006.
- [8] Zhang D., Li W., Wu X., et al. Application of simulated annealing genetic algorithm-optimized back propagation (BP) neural network in fault diagnosis. International Journal of Modeling Simulation and Scientific Computing, Vol. 10, Issue 4, 2019, p. 1950024.
- [9] Song Z., Wang J. Transformer fault diagnosis based on BP neural network optimized by fuzzy clustering and LM algorithm. High Voltage Apparatus, Vol. 49, Issue 5, 2013, p. 54-59.