ISA Transactions ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

ISA Transactions

journal homepage: www.elsevier.com/locate/isatrans

Research Article

Compound feature selection and parameter optimization of ELM for fault diagnosis of rolling element bearings

Meng Luo [a], Chaoshun Li [a,*], Xiaoyuan Zhang [b], Ruhai Li [a], Xueli An [c]

[a] School of Hydropower and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, PR China
[b] College of Electrical Engineering, Henan University of Technology, Zhengzhou 450001, PR China
[c] China Institute of Water Resources and Hydropower Research, Beijing 100044, PR China

ARTICLE INFO

Article history:
Received 19 November 2015
Received in revised form
30 July 2016
Accepted 22 August 2016
This paper was recommended for publication by Dr. Steven Ding

Keywords:
Extreme learning machine
Gravitational search algorithm
Parameter optimization
Feature selection
Ensemble empirical mode decomposition
Fault diagnosis

ABSTRACT

This paper proposes a hybrid system named as HGSA-ELM for fault diagnosis of rolling element bearings, in which real-valued gravitational search algorithm (RGSA) is employed to optimize the input weights and bias of ELM, and the binary-valued of GSA (BGSA) is used to select important features from a compound feature set. Three types fault features, namely time and frequency features, energy features and singular value features, are extracted to compose the compound feature set by applying ensemble empirical mode decomposition (EEMD). For fault diagnosis of a typical rolling element bearing system with 56 working condition, comparative experiments were designed to evaluate the proposed method. And results show that HGSA-ELM achieves significant high classification accuracy compared with its original version and methods in literatures.

1. Introduction

Rolling element bearing is one of the most common components in the modern rotating machinery [1–3]. It is very important to identify the faults in the rolling element bearings accurately, easily and fast, as long as its condition deteriorates. In order to improve the classification accuracy of a fault diagnosis system, two factors should be taken into account: the amount of information contained within the extracted fault features and the ability of the classifier to correctly differentiate among faults [1].

The vibration signals carry a lot of information about the health conditions of the rotating machine [1,4], and fault features are extracted from the vibration signals using signal processing techniques. For rolling bearings fault diagnosis, because of the nonlinear factors such as loads, clearance, friction and stiffness, the vibration signal is usually non-stationary. In order to process non-stationary vibration signal, traditional time-frequency domain analysis, such as the wavelet transform (WT) [5] and Empirical mode decomposition (EMD) method [6] are widely used. As long as the variation is insignificant compared with the whole operation duration, time-frequency domain analytical techniques are reliable and efficient in computation. Although WT is a powerful method for nonlinear and non-stationary signal process, two drawbacks remain in this method. Firstly, the adaptation is extremely restricted by requirement of selection of base function and determination of decomposition scales. Secondly, the energy leakage is inevitable due to the fact that WT is essentially an adjustable windowed Fourier transform. Empirical mode decomposition (EMD) method [6], which is proposed by Huang, overcomes the above two drawbacks. Since it is based on the local characteristic time scales of signal and has good ability to decompose the complicated signal into several intrinsic mode functions (IMFs) self-adaptively. However mode mixing problem often occurs when using EMD. In order to overcome the mode mixing problem of the EMD method, ensemble empirical mode decomposition (EEMD) [7] is proposed, by adding finite while noise into the original signal. Features extracted from IMFs can reflect the characteristic of rolling element bearing more accurately and effectively [8–12]. For example, in [8], to analyzing the nonlinear and nonstationary vibration signal, a new approach based on EEMD and a self-zero space projection model (SSPM) is proposed to identify faults in a rotor-bearing system. In [9], EEMD was used to decompose bearing vibration signals, and the effective features were successfully extracted from the obtained IMFs using an autoregressive model and kernel principal component analysis for bearing fault diagnosis. Singular value [1] and energy entropy

* Corresponding author.
E-mail address: csli@hust.edu.cn (C. Li).

http://dx.doi.org/10.1016/j.isatra.2016.08.022
0019-0578/© 2016 Published by Elsevier Ltd. on behalf of ISA.

[13] calculated based on IMFs are proved to be effective and have already applied in fault diagnosis. Since different types of features have shown special effectiveness in fault diagnosis, it is reasonable to employ compound features consisting of different types of features for the purpose of avoiding information loss.

Feature combination will increase useful information for fault diagnosis, and will bring some redundant information inevitably. Redundant features may reduce the accuracy and efficiency of fault diagnosis. In order to improve performance of fault diagnosis, feature combination and feature selection are all necessary. For the problem of feature selection in fault diagnosis, a lot of methods have been proposed in references. Typically, in [14], a novel two-stage feature selection, hybrid distance evaluation technique (DET)-particle swarm optimization (PSO), is proposed to select the superior combining feature subset that discriminates well among classes. In [15], an optimization mechanism is proposed, which employs the discrete-valued PSO to optimize the input feature subset selection.

The choice of fault diagnosing model is also very important. Some popular models are used for detecting and classifying faults of the rolling element bearing, such as neural networks [16], fuzzy systems [17], and support vector machines (SVM) [18]. In this paper, we employ the extreme learning machine (ELM) [19,20] to classify fault samples of different working conditions. ELM is a novel model for a single hidden layer feed forward neural networks. In ELM, the input weights and hidden bias are chosen randomly, and the output weights are determined by using Moore–Penrose (MP) generalized inverse. ELM not only learns much faster with higher performance than the traditional gradient-based learning algorithms, but also avoids many difficulties such as stopping criteria, learning rate and local minima [21,22].

At present, some researchers have tried to improve the performance of ELM. Such as, a hybrid optimization mechanism was proposed in [15], which combined the discrete-value PSO with the continuous-value PSO to optimize the input feature subset and the number of hidden nodes to enhance the performance of ELM. In this method, the number of hidden nodes of ELM was chosen as optimization variable, and the input weights and bias of ELM were generated randomly based on the given number of hidden nodes. However, the input weights and bias directly affect the performance of ELM. So the number of hidden nodes may not be the most suitable optimization variable for optimization of ELM. In [22,23] an improved PSO were employed to optimize the input weights and bias of ELM, and the improvements were manifest. These researches have proved that optimization of ELM could improve the performance of this model significantly.

Based on above discussion, this paper is motivated by the following assumptions: (1) combination of different types of features will augment useful information for fault classification; (2) the compound features contains redundant information unavoidable, and it's necessary to eliminate redundant information by feature selection; (3) advanced optimization methods will improve accuracy of ELM. Therefore, we proposed a new fault diagnosis model for the rolling element bearings. At the first step, a compound feature set consisting of time and frequency features, the energy features and singular value features were extracted by using EEMD. And then, a hybrid GSA-ELM (HGSA-ELM) was proposed by combining a real-valued GSA (RGSA) to optimize ELM and the binary-valued of GSA (BGSA) to select the significant features.

The rest paper is presented as followings: Section 2 introduces the EEMD, and different feature extraction methods based on EEMD. Section 3 introduces the ELM algorithm. Section 4 presents the RGSA algorithm and BGSA algorithm. Section 5 describes the proposed HGSA-ELM model for fault diagnosis. The experiment arrangement and results are detailed in Section 6. Finally, Section 7 summarizes our conclusions.

2. Compound feature extraction based on EEMD

2.1. Ensemble empirical mode decomposition

The EMD method proposed by Huang [6] is an adaptive decomposition method for nonlinear and non-stationary signals. The EMD method is able to adaptively decompose non-liner and non-stationary signal into several intrinsic mode functions (IMFs) and residual trend from high frequency to low frequency. It is a versatile approach for extracting signals from data generated in noisy nonlinear and non-stationary processes. However, it might incorrectly reveal the signal's characteristic information as a result of the mode mixing, which may render the physical meaning of IMFs unclear. To overcome this drawback, the Ensemble EMD (EEMD) was proposed by Wu and Huang [24], The EEMD is a noise-assisted EMD procedure actually, using the full advantage of the Gaussian white noise's statistical characteristic of uniform distribution to improve the distribution of extreme points in original signal, fairly well solving the mode mixing problem [25]. The procedure of EEMD algorithm is as follows:

(1) Initialize the number of ensemble M and the amplitude of the added white noise, with $i=1$.

(2) Add a white noise $n_i(t)$ with the given amplitude to the original $x(t)$

$$x_i(t)=x(t)+n_i(t) \tag{1}$$

where $n_i(t)$ indicates the ith added white noise series, and $x_i(t)$ represents the noise-added signal of the ith trial ($i=1, 2, \ldots, M$).

(3) The EMD algorithm performed to decompose the $x_i(t)$ into different components

$$x_i(t)=\sum_{k=1}^{m} g_{i,k}(t)+r_i(t) \tag{2}$$

where m is the number of IMFs, $r_i(t)$ is the final residue, and $g_{i,k}(t)$ is the IMFs, which include different frequency bands ranging from high to low.

(4) Repeat (2) and (3) M times, then the final result of IMF components is obtained by computing the ensemble mean of the corresponding components.

$$c_k(t)=\frac{1}{M}\sum\nolimits_{i=1}^{M} g_{i,k}(t) \tag{3}$$

where $c_k(t)$ is the final IMF decomposed by EEMD.

2.2. Compound feature extraction

Feature extraction has a critical influence on fault diagnosis, which influence the accuracy and efficiency of the fault diagnosis model. In this section we propose three types of feature: time and frequency features extracted from the original vibration signals, energy features and singular value features extracted from the IMFs decomposed by EEMD.

2.2.1. Energy features based on EEMD

The corresponding resonance frequency components are generated in the vibration signals, while some different fault occurs in the bearing. At the same time the energy of the fault vibration signal changes with the frequency distribution. Hence, we can use the energy to specify the fault types [26].