Rotor Fault Dataset Classification Method Based on EEMD Energy Moment and Neighborhood Rough Set (Postprint)

Authors: Sun Zejin, Zhao Rongzhen

Date: 2018-12-13T00:00:00+00:00

Abstract

To address the issue of low accuracy in rotating machinery fault identification, a novel method for rotor system fault pattern identification is proposed by integrating ensemble empirical mode decomposition (EEMD), energy moment, and neighborhood rough set (NRS). Initially, EEMD is employed to adaptively decompose the collected vibration fault signals into several stationary intrinsic mode function (IMF) components, and their energy moments are calculated. These energy moments serve as conditional attributes describing the fault state to construct a fault identification decision table. Subsequently, neighborhood rough set is utilized to perform attribute reduction on the decision table, thereby eliminating redundant attributes. Finally, the reduced sensitive feature subset is fed into the designed decision tree (DT) C4.5 algorithm for pattern recognition. The effectiveness of the proposed method is validated using fault feature sets from a typical rotor test rig.

Full Text

Rotor Fault Data Set Classification Method Based on EEMD Energy Moment and Neighborhood Rough Sets

Sun Zejin, Zhao Rongzhen
(School of Mechanical & Electronical Engineering, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract: Aiming at the problem of low recognition accuracy in rotating machinery fault diagnosis, this paper proposes a rotor system fault identification method based on the combination of ensemble empirical mode decomposition (EEMD) with energy moment and neighborhood rough sets (NRS). Firstly, the method uses EEMD to decompose non-stationary vibration signals into several stable intrinsic mode functions (IMFs) and calculates the energy moment of each IMF component. Using this energy moment as the condition attribute to describe the fault state, a fault identification decision table is established. Then, neighborhood rough set theory is applied to perform attribute reduction on the decision table to eliminate redundant attributes. Finally, the reduced sensitivity feature subsets are used as input to the decision tree C4.5 algorithm for recognition. Experimental results demonstrate that this method can effectively improve the accuracy of rotor fault identification.

Keywords: ensemble empirical mode decomposition; intrinsic mode functions; energy moment; decision table; neighborhood rough set; attribute reduction; decision tree C4.5 algorithm

0 Introduction

Rotor faults in rotating machinery pose significant hazards. Since rotors typically operate in multi-vibration source environments, traditional single-diagnosis modes struggle to comprehensively and accurately capture fault characteristics. Therefore, constructing fault data sets from multi-channel signals has a positive effect on promoting the development of intelligent diagnosis technologies. However, such high-dimensional fault feature sets from multi-channel fusion inevitably contain redundant attributes, which undoubtedly increases classifier learning and training time, and may even reduce analysis accuracy. Consequently, comprehensively mining sensitive fault feature information and eliminating redundant attributes is particularly important for improving fault recognition precision.

Currently, the attribute reduction method based on neighborhood rough sets proposed in literature [4] provides a good solution to the curse of dimensionality, but the selection of condition attributes requires further research. Literature [5] proposed a liquid solenoid valve fault diagnosis method based on EMD and neighborhood rough sets, and demonstrated its effectiveness through experiments. Literature [6] used IMF energy moments after EMD decomposition to construct feature vectors and achieved precise diagnosis of rolling bearing faults. However, when EMD decomposes complex signals with multi-modal mixing, it can cause mode mixing in IMF components, leading to low decomposition precision. This results in some IMF component energy moments being unable to accurately describe the working state, leaving the problem of insufficiently high precision in rotor fault type identification to be solved. Furthermore, rotor system fault identification requires additional research. The introduction of multi-domain features allows 多元冗余信息 (multiple redundant information) to infiltrate the data, ultimately leading to low fault recognition accuracy. Therefore, dimensionality reduction methods are needed for effective secondary feature extraction to obtain low-dimensional sensitive primary feature vectors.

Empirical Mode Decomposition (EMD) [7], proposed by Huang et al., is a new adaptive time-frequency analysis method that can perform adaptive time-frequency decomposition according to the local time-varying characteristics of signals. It is highly suitable for feature extraction from nonlinear and non-stationary signals. However, current EMD decomposition suffers from endpoint effects and mode mixing phenomena. To address these issues, Wu and Huang et al. introduced noise-assisted analysis based on the EMD method and proposed the Ensemble Empirical Mode Decomposition (EEMD) method [8], which can achieve the same decomposition purpose as EMD while effectively suppressing mode mixing [9].

To achieve rapid recognition between low-dimensional vectors after neighborhood rough set attribute reduction and fault types, a simple and high-precision intelligent classifier must be selected. Such classifiers evaluate similarity between training and test samples through certain measurement rules, then assign test samples to the type that has the most neighboring sample points. In other words, fault identification is accomplished by finding the nearest known-labeled samples and assigning their label information to the samples to be classified. Decision trees [15], with their simplicity, intuitiveness, and good classification performance, have been widely applied in fault diagnosis.

Based on the above analysis, this study proposes a rotor fault data classification method that combines EEMD energy moment with neighborhood rough sets and decision tree algorithms. Leveraging EEMD's ability to effectively obtain characteristic distribution types from rotor vibration signals, the method uses energy moment indicators to select IMF components containing fault information after decomposition, constructs a feature vector matrix for rotor fault classification, and applies the advantages of neighborhood rough sets in attribute reduction and decision tree algorithms in data classification to rotor system fault diagnosis, providing a reference approach for dimensionality reduction and fault classification of rotor fault data sets.

1.1 EEMD Signal Decomposition Principle

EEMD decomposition can adaptively decompose nonlinear, non-stationary multi-modal signals into several stable single-modal IMF components and a residual term according to the signal's own characteristics [8]. In traditional EMD methods, discontinuities in IMF components cause mode mixing in adjacent waveforms, primarily due to two reasons: (a) insufficient extreme points in the signal causing decomposition to stop; and (b) non-uniform distribution intervals of extreme points when using cubic spline functions to fit them. To overcome these defects, literature [8] proposed the EEMD decomposition method.

This method utilizes the statistical characteristic of uniform frequency distribution of Gaussian white noise to compensate for the discontinuity defects mentioned above. For specific details of the EEMD decomposition process, please refer to literature [10].

Definition of IMF Energy Moment

EEMD decomposition of collected vibration signals yields several IMF components and a residual term. Research in literature [11] shows that rotor system fault information is mainly concentrated in the first few IMF components. Therefore, fault feature information is generally extracted from these primary IMF components. However, to avoid missing fault information, this experiment calculates the energy moments of 11 IMF components and one residual term.

a) For a signal sequence, the energy moment of each component after EEMD processing is calculated as shown in Equation (1).

$$E_i = \sum_{m=1}^{n} (a_t)^2 \cdot \Delta t, \quad i = 1, 2, 3, \dots, n$$

$$T = [E_1, E_2, E_3, \dots, E_n]$$

where: a_t is the intrinsic mode function; Δt is the sampling period; n is the total number of sampling points; and m is the sampling point.

b) Using the energy moment of each IMF component as elements, a feature vector T describing the fault state can be constructed.

Energy moment has unique performance in mechanical signal information quantity evaluation and information component analysis. Compared with traditional methods that use signal energy or energy entropy to describe rotor system characteristics, energy moment can reflect both the magnitude of energy and its distribution over time parameters, better revealing energy distribution characteristics and signal essence features, which is beneficial for rotor system fault feature extraction [6].

1.3 Neighborhood Rough Sets and Attribute Reduction Concepts

Pawlak's rough set theory [12] can only handle discrete precise sets and requires discretization of numerical data before processing, which may lose important information and affect final results. Hu Qinghua et al. [13] proposed the concept of neighborhood rough sets, which can directly process continuous data sets, solving the necessary discretization step in classical rough set theory and avoiding information loss caused by discretization.

Given a neighborhood decision system $NDT = \langle U, A, D \rangle$, where: - $U = \{x_1, x_2, \cdots, x_n\}$ is the experimental sample set (the universe) - $A = \{a_1, a_2, \cdots, a_m\}$ is the set of all condition attributes describing the universe - D is the classification decision attribute (i.e., fault type) - According to D's attribute value domain, U can be partitioned into N equivalence classes: $U/D = \{X_1, X_2, \cdots, X_N\}$ - $B \subseteq A$ generates a neighborhood relation N_B -

The upper and lower approximations of decision attribute D with respect to condition attribute subset B are respectively:

$$\underline{N_B}D = \bigcup_{i=1}^{N} \underline{N_B}X_i$$

$$\overline{N_B}D = \bigcup_{i=1}^{N} \overline{N_B}X_i$$

The boundary region of the decision system is:

$$BN_B(D) = \overline{N_B}D - \underline{N_B}D$$

The positive and negative domains of the neighborhood decision system are respectively:

$$POS_B(D) = \underline{N_B}D$$

$$NEG_B(D) = U - \overline{N_B}D$$

The dependency degree reflects the proportion of samples that can be correctly classified in the system, obviously $0 \le \gamma_B(D) \le 1$. The larger the positive domain, the stronger the dependency of decision attribute D on condition attribute B.

$$\gamma_B(D) = \frac{|POS_B(D)|}{|U|}$$

If $a \in B$, then the importance of condition attribute a for decision attribute D is:

$$Sig(a, B, D) = \gamma_{B\cup\{a\}}(D) - \gamma_B(D)$$

Attribute reduction is one of the core concepts in neighborhood rough set theory research. It is the concept of using indiscernibility relations to delete redundant attributes while keeping the classification ability of the decision table unchanged. This helps eliminate redundant attributes and reduce time and space complexity of data processing. Based on the attribute importance index, a greedy attribute reduction algorithm can be constructed [14].

Algorithm Input: $NDT = \langle U, A, D\rangle$
Algorithm Output: Attribute reduction subset B

Steps: a) Initialize attribute reduction subset $B = \emptyset$ b) For any remaining attribute $a_i \in A - B$, calculate the importance of this attribute relative to reduction subset B: $Sig(a_i, B, D)$ c) Select the attribute with maximum importance $Sig(a_j, B, D) = \max(Sig(a_i, B, D))$. If $Sig(a_j, B, D) > 0$, then add attribute a_j to the attribute reduction subset $B = B \cup \{a_j\}$ and return to step b) to continue calculation. Otherwise, terminate the program and output the final reduction result.

1.4 Decision Tree Algorithm

Decision tree is a widely used classification algorithm whose core is the C4.5 algorithm, proposed by Quinlan based on the ID3 algorithm [15]. The core issues of decision trees are test attribute selection and tree pruning. The C4.5 algorithm uses information gain ratio as the criterion for selecting branch attributes, overcoming the shortcoming of ID3 algorithm's bias toward attributes with many values when using information gain for attribute selection [16].

2 Design of Fault Dataset Classification Method Based on EEMD Energy Moment and Neighborhood Rough Sets

2.1 Design of Fault Diagnosis Method Based on EEMD Energy Moment and Neighborhood Rough Sets

Since Sections 1.1 and 1.2 have detailed the EEMD decomposition process and high-dimensional feature set construction process, this section will not repeat them. Attribute reduction is an important application of neighborhood rough set theory. Neighborhood rough sets compensate for the defect in classical rough set theory that requires discretization when processing continuous data. After reduction by neighborhood rough sets, redundant features in the data are deleted, reducing data processing complexity and facilitating subsequent fault classification.

Based on the EEMD-NRS algorithm principle, the rotor system fault diagnosis process is designed as shown in Figure 1 [Figure 1: see original paper].

Figure 1 Flow chart of fault diagnosis based on EEMD energy moment and neighborhood rough set theory

Specific algorithm steps: a) Perform EEMD decomposition on collected signals from five states: unbalance, misalignment, rubbing, looseness, and normal, obtaining several IMF components and a residual term for each. b) Calculate the energy moment of each IMF component using Equation (1), and construct a vector $E = (E_1, E_2, \cdots, E_m)$ using the energy moments of each IMF component as the feature vector for rotor fault states. c) Use the feature vector matrix E to construct a high-dimensional feature set and generate a decision table. Apply neighborhood rough sets for attribute reduction on the decision table, calculating the importance of each condition attribute using Equation (11), and select the attribute set with relatively high importance to add to the reduction set. d) Use the reduced low-dimensional sensitive feature subset as input to train and build a decision tree model for fault type identification.

3 Application Results Analysis

The fault data set used in this work originates from the rotor test bench shown in literature [17]. Experimental data were collected at a sampling frequency of 5000 Hz with a drive motor speed of 2800 r/min. The following four typical faults were simulated: {support looseness, rotor-stator rubbing, mass unbalance, and shaft misalignment}, along with the normal state.

3.1 Status of Extracted Fault Features

The time-domain waveforms of raw vibration signals measured under four fault conditions and normal state are shown in Figure 2 [Figure 2: see original paper] (due to space limitations, only the raw vibration signal of the first channel under misalignment state is listed). For feature extraction analysis of the Figure 2 signal, EEMD processing was performed on vibration signals from each channel (12 channels total) for each state. To select real feature components containing fault information, the energy moment evaluation index was introduced. This paper uses misalignment fault data with a signal length of 2048 points as an example to illustrate the evaluation method. In this experiment, the auxiliary white noise standard deviation for EEMD was set to 0.2 times the original standard deviation with $M = 100$. Taking the misalignment fault signal as an example, its decomposition results are shown in Figure 3 [Figure 3: see original paper]. It can be seen that 11 IMF components and one residual term were decomposed, and Equation (1) was used to calculate the energy moments of these components.

As shown in Figure 4 [Figure 4: see original paper], the energy of the 11 decomposed IMF components and one residual term varies significantly. Therefore, to avoid missing fault information, the energy moments of 12 components were selected to establish the diagnosis model in the experiment. Fifty sample groups were collected for each fault, plus the normal state, totaling 250 groups to generate the decision table. The obtained energy moment features were set as condition attributes in the decision table, while decision attribute values 1, 2, 3, 4, and 5 correspond to the five states of misalignment, unbalance, rubbing, looseness, and normal, respectively. Then attribute reduction was performed on the decision table as shown in Table 1 . In Table 1, a_1 to a_{144} are used as condition attributes of the neighborhood rough set, and the last column showing each sample's fault type is used as the decision attribute of the neighborhood rough set.

Typically, the neighborhood size can be obtained according to Equation (13) [18]:

$$\delta(a_i) = \lambda \cdot Std(Z_i)$$

where $Std(Z_i)$ represents the standard deviation of attribute a_i, and λ is a set parameter used to adjust neighborhood size. Different λ values correspond to different neighborhood radii, typically $\lambda \in [2, 4]$ [4]. As shown in Table 2 , to achieve optimal diagnosis results, λ was varied with a step size of 0.1 to obtain different neighborhood radii and reduction sets. The reduced results were input into the decision tree classifier, and classification results were compared. It was found that under different lower limits of attribute importance, the reduced attributes varied when $\lambda = 3$, and the best classification effect was achieved when the importance lower limit was 0.01 and $\lambda = 3$.

After calculating the dependency degree of each condition attribute using Equation (10), the optimal feature reduction set output by the attribute reduction algorithm is $A = \{a_4, a_{41}, a_{113}\}$.

Decision Tree Construction: For the 250 collected sample groups, training was performed using random allocation. MATLAB's `randperm` function was used to randomly generate 250 random numbers from 1 to 250 (i.e., a random permutation of 1-250). The first 200 groups were assigned as training samples and the remaining 50 as test samples. This experiment was repeated 10 times. To compare the fault identification effect of this method, both EMD decomposition and EEMD decomposition were performed. The classification effect comparison of the two methods is shown in Figure 5 [Figure 5: see original paper] and Table 3 . Figures 5(a)-(d) show the decision tree model classification effect diagrams constructed with sample data before and after neighborhood rough set reduction, respectively. Table 3 shows the diagnosis results before and after reduction for both methods.

Table 1 The decision table before and after neighborhood rough set reduction

Condition Attributes A	…	Decision Attribute D
399.9463064	…	2.40E-05
398.7100662	…	3.52E-05
…	…	…

Table 2 Comparison of classification accuracy under different importances and λ

[Table content showing classification accuracy variations]

From Figures 5(a)-(d), it can be seen that although both before and after reduction using EEMD decomposition have four leaf nodes and the same tree size, the pre-reduction version used 4 attribute values to distinguish the five states, while the post-reduction version distinguished all states using only 3 attribute values. For EMD decomposition, there were four leaf nodes before reduction but only three after reduction, with the pre-reduction version requiring 4 attribute values to distinguish the five states, while the post-reduction version distinguished all states using only 2 attribute values. Table 2 shows that the recognition accuracy after EEMD decomposition is higher than that after EMD decomposition. The reason is that EMD decomposition causes mode mixing in IMF components, which leads to the deletion of useful mixed attribute values during attribute reduction, resulting in information loss. Although the decision tree model constructed with EMD requires fewer leaf nodes and condition attribute values than EEMD, the final fault recognition accuracy is relatively lower.

Figure 5 Schematic diagram of decision trees before and after neighborhood rough set reduction

Table 3 Recognition accuracy before and after neighborhood rough set reduction

Method	Misalignment	Unbalance	Rubbing	Looseness	Normal	Average Recognition Rate (%)
EEMD-NRS-DT						
EEMD-DT						
EMD-DT						
EMD-NRS-DT						

To further verify the superiority of the EEMD-NRS-DT method, this paper compares the recognition results of BP neural network and decision tree methods using low-dimensional vectors after attribute reduction and original high-dimensional feature sets under different training sample quantities, as shown in Table 4 . The results show that: (a) Whether using decision tree or neural network methods, recognition efficiency improves after feature reduction, and the comprehensive diagnosis rate does not decrease significantly, because neighborhood rough set attribute reduction eliminates redundant features and improves system efficiency; (b) The recognition accuracy of decision trees is higher than that of BP neural networks, because conventional BP neural networks can be seen as a "black box" with unknown internal structure, making the learning process unobservable and output results difficult to interpret. BP networks also require numerous parameters such as network topology, weights, and threshold initial values. Using gradient descent methods, they are prone to falling into local minima, have long training times, and slow convergence. Moreover, there is no unified and complete theoretical guidance for BP network structure selection, which generally can only be determined empirically. In contrast, decision trees are easy to understand and interpret, have low requirements for data preparation, while many other classification algorithms have certain constraints on data attributes and require generalization before application. Neural networks also need to convert discrete attributes to numerical attributes before processing, whereas decision trees require no such conversion. Additionally, decision trees are a "white box" model that can conveniently derive corresponding logical rules. Since decision trees can be constructed for data sets with many attributes, they scale well to large databases. Their size is independent of database size and can produce feasible and good results on large data sources in relatively short time.

By comparing the classification results of EMD-NRS-DT and EEMD-NRS-DT, it is evident that using the EEMD energy moment and neighborhood rough set diagnosis model for rotor system fault diagnosis is reasonable.

Table 4 Recognition accuracy (%) corresponding to different sample numbers

Training/Test Sample Numbers	100/150	150/100	200/50
EMD-DT			
EMD-BP			
EEMD-DT			
EEMD-BP			
EMD-NRS-DT			
EMD-NRS-BP			
EEMD-NRS-BP			
EEMD-NRS-DT			

4 Conclusion

a) To effectively identify rotor system faults and improve fault recognition rate, this paper proposes a method combining EEMD energy moment with neighborhood rough sets and decision trees. Through experiments simulating five states of the rotor system (normal, support looseness, rotor-stator rubbing, mass unbalance, and shaft misalignment), the rationality of this method for fault diagnosis of rotor system vibration signals is verified.

b) To address the large amount of redundant information in IMF components after decomposition, a greedy attribute reduction algorithm based on neighborhood rough sets is applied to reduce the condition attribute set. Finally, a decision tree is built using the reduced attribute subset. After training, the diagnostic accuracy of the obtained model reaches 100%, achieving the expected results.

c) Experimental results from rotor test bench fault data sets demonstrate that the fault classification mode based on EEMD energy moment combined with neighborhood rough set attribute reduction provides better fault identification performance compared with traditional EMD-neighborhood rough set diagnosis methods, offering a solution approach for intelligent diagnosis of complex rotating machinery faults.

References

[1] Zheng Zhi, Jiang Wanlu, Hu Haosong, et al. Research on fault diagnosis method of rolling bearing based on integrated empirical mode decomposition and kurtosis criterion [J]. Proceedings of the CSEE, 2012, 32(11): 106-111.

[2] Liang Shengjie, Zhang Zhihua, Cui Lilin, et al. Dimensionality reduction method based on principal component analysis and kernel independent component analysis [J]. Systems Engineering and Electronics, 2011, 33(9): 2144-2148.

[3] Zhao Xiaoli, Zhao Rongzhen. Research on dimensionality reduction method of rotor fault data set based on global and local discriminant information fusion [J]. Acta Automatica Sinica, 2017, 43(4): 560-567.

[4] An Ruoming, Suo Mingliang. Application of attributes reduction and weights calculation through neighborhood rough set [J]. Computer Engineering and Applications, 2016, 52(7): 160-165.

[5] Tan Yangbo, Cheng Jinjun, Liu Shuai. Fault diagnosis of liquid solenoid valve based on EMD and neighborhood rough set [J]. Computer Engineering and Applications, 2017, 53(12): 255-260.

[6] Zhang Min, Cui Hailong, Chen Yuhui, et al. Fault diagnosis method of rolling bearing based on IMF energy moment and HSMM model [J]. Combined Machine Tool & Automatic Processing Technology, 2015(10): 101-103.

[7] Huang N E, Zheng Shen, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis [J]. Proceedings Mathematical Physical & Engineering Sciences, 1998, 454(1971): 903-995.

[8] Wu Zhaohua, Huang N E. Ensemble empirical mode decomposition: a noise-assisted data analysis method [J]. Advances in Adaptive Data Analysis, 2011, 1(1): 1-41.

[9] Hu Aijun, Ma Wanli, Tang Guiji. Fault feature extraction method of rolling bearing based on integrated empirical mode decomposition and kurtosis criterion [J]. Proceedings of the CSEE, 2012, 32(11): 106-111.

[10] Ma Zaichao, Zhao Rongzhen, Yang Wenbiao. KPCA-BFDA method for rotor fault feature data classification [J]. Vibration, Test & Diagnosis, 2013, 33(2): 192-198.

[11] Li Linfeng, Zhao Jun, Guo Tiantai, et al. Gear fault diagnosis method based on eemd, fuzzy entropy and SVM [J]. Mechanical Transmission, 2014(2): 147-151.

[12] Pawlak Z. Rough Sets: Theoretical Aspects of Reasoning about Data [M]. Kluwer Academic Publishers, 1991.

[13] Hu Qinghua, Yu Daren, Xie Zongxia. Numerical attribute reduction based

on neighborhood granularization and rough approximation [J]. Journal of Software, 2008, 19(3): 640-649.

[14] Hu Qinghua, Zhao Hui, Yu Daren. Fast reduction algorithm for symbol and numerical properties based on neighborhood rough sets [J]. Pattern Recognition & Artificial Intelligence, 2008, 21(6): 730-738.

[15] Quinlan J R. C4.5: programs for machine learning [M]. San Francisco: Morgan Kaufmann Publishers Inc., 1992.

[16] Li Liuxing, Zhou Huangli. Bearing fault identification based on sample entropy and decision tree adjustment algorithm [J]. Journal of Anhui Agricultural University, 2017, 44(5): 936-940.

[17] Huo Tianlong, Zhao Rongzhen, Hu Baoquan. SVM rotor fault diagnosis based on entropy band method and PSO optimization [J]. Journal of Vibration, Testing and Diagnosis, 2011, 31(3): 279-284.

[18] Zhang Dongwen, Wang Peng, Qiu Jiqing. Attribute reduction algorithm based on neighborhood rough set and ant colony optimization [J]. Journal of Hebei University of Science and Technology, 2011, 32(5): 403-408.

Note: Figure translations are in progress. See original paper for figures.