Journal of Sound and Vibration 425 (2018) 137–148

Contents lists available at ScienceDirect

Journal of Sound and Vibration

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

# Variable predictive model class discrimination using novel predictive models and adaptive feature selection for bearing fault identification

Tao Tang, Lin Bo*, Xiaofeng Liu, Bing Sun, Daiping Wei

*The State Key Laboratory of Mechanical Transmission, Chongqing University, Chongqing 400044, PR China*

A R T I C L E I N F O

*Article history:*
Received 3 November 2017
Received in revised form 1 March 2018
Accepted 30 March 2018

Handling Editor: K. Shin

*Keywords:*
Variable predictive model
Rolling bearing diagnosis
Neural-network
Feature selection
Affinity propagation
RReliefF

A B S T R A C T

A complete fault diagnosis for the rolling bearing is proposed in this paper. Variable predictive model class discrimination (VPMCD) is a conventional pattern recognition method; however, in practice, when the fault diagnosis method is applied to small samples or in multi-correlative feature space, the stability of the VPM constructed based on the least squares (LS) method is not sufficient. Based on affinity propagation (AP) clustering, RReliefF, and sequential forward search, the ARSFS is proposed to select the significant subset of original feature set and to reduce the dimension and multiple correlations of the feature space. Further, this paper uses two kinds of Gaussian Neural Network, namely the Radial Basis Function Neural Network (RBF) and the Generalized Regression Neural Network (GRNN), instead of the LS method to construct predictive models of VPMCD, called AOR-VPMCD. Compared with the conventional VPMCD and its improvements, based on sufficient experiments, the entire process presented in this paper can effectively identify the fault of the rolling bearing.



## 1. Introduction

Vibration signals gathered from bearings usually involve a great deal of information about the state of the machine's operation [1,2]. The variable predictive model-based class discrimination as a distinct method of pattern recognition was proposed by Raghuraj, R et al. [3], and this method has been wildly used in fault diagnosis of the rolling bearing. According to the significant differences of the intrinsic correlation between characteristics in different bearing states, different samples can be distinguished. Compared with other algorithms, VPMCD avoids the iterative process of the artificial neural network (ANN) [4] and the optimization process in the SVM [5], and there are no preset parameters. However, practically in the face of the varied effects of non-linear factors such as load, friction, clearance, and stiffness, it is difficult to establish precise variable predictive models (VPMs) with merely linear or quadratic models.

Aiming at the shortcomings of the conventional models in abnormal data processing, a robust regression-variable predictive model are proposed by Yu Yang et al. [6], which makes up the problem that least square method is sensitive to abnormal data. Similarly, Songrong Luo in Ref. [7] combined ANN with the mean impact value (MIV) to reduce feature dimensions and enhance the stability of the VPMCD output. Nevertheless, they both ignore VPMCD estimated by the Least

* Corresponding author.
*E-mail address:* bolin0001@aliyun.com (L. Bo).

https://doi.org/10.1016/j.jsv.2018.03.032

Squares (LS) method [8] is sensitive not only to the abnormal data but also to the existing multi-correlation between variables. Recently by Cui et al. [9], Partial Least Squares method (PLS) instead of the LS method was used to estimate the model parameters of VPMCD to address the negative influence of multiple correlations. Besides, to strengthen the robustness of recognition, GA-VPMCD was put forward by combining the genetic algorithm (GA) with four conventional VPMs by Luo SR et al. [10]. It does break the limitations of one single VPM, but the GA method can also contribute to considerable time costs.

To overcome these deficiencies, an adaptive optimal radial model-based VPMCD is proposed, called AOR-VPMCD. The new VPMs are constructed by two artificial neural networks, namely the Radial Basis Function Neural Network (RBF) [11] and the Generalized Regression Neural Network (GRNN) [12]. Unlike other methods of improvement, we managed to redefine the new VPMs rather than strengthening the original ones and to verify its superiority by comparing other improvements.

These features extracted from the original or preprocessing signals may have multiple correlations or large dimensions, which is not conducive to pattern recognition in VPMCD. Most of the feature selection (FS) methods can be roughly divided into two broad categories, namely, filter and wrapper. Regarding filter methods, features are scored and graded according to statistical criteria. For example, in Ref. [13], Zexian Wei et al. have designed a feature selector based on affinity propagation (AP) clustering algorithm to extract the cluster centers as optimal features at the cost of considerable information loss, as just these central points may be insufficient to carry all the state information. In Ref [14], considering the varied contributions of features, Abrahamsson, V et al. combined F-statistics feature selection with ReliefF algorithm, but the artificial thresholds have great randomness. On the wrapper methods, Alrajab, M et al. [15] proposed a three-phase combinatorial feature selection algorithm, including particle swarm optimization (PSO) and Support Vector Machine (SVM) algorithm. Although this method has been identified to do well in feature selection, it also has higher computational costs.

This paper uses sequential forward selection (SFS) method and two filtering feature selection methods, namely AP [16] and RReliefF [17], to create our unsupervised feature selector called ARSFS. First, the raw features are evaluated by the RReliefF algorithm and are clustered into several subsets using AP clustering, which is to distinguish the useful correlation from multi-correlated features. Second, based on AP clustering and SFS method, a feature evaluation loop is designed to extract key features from each subset. Features extracted from all feature subsets can inherit the correlation between subsets. Without any manual parameters, ARSFS can use feature selection loops to differentiate between critical and unrelated features. Besides, the number of key features can be determined automatically. For further comparison, several feature methods have been applied to verify the ability of ARSFS, such as the feature center method and ReliefF-SFS method.

The remainder of this paper is organized as follows. In Section 2, the fundamental theories are briefly introduced, including feature extraction, AP, and RReliefF. In Section 3, the proposed feature selector is illustrated. Section 4 summarizes the VPMCD and AOR-VPMCD method respectively. Section 5 presents and analyses the engineering application of the proposed techniques. Finally, some conclusions are brought in Section 6.

## 2. Brief review of basic theories

### 2.1. Feature extraction

Recurrence Quantification Analysis (RQA) [18,19] is a method of nonlinear time series analysis. The primary focus lies on recurrence plots (RP) and their quantification. Wherein RPs are briefly defined as

$$\mathbf{R}_{i,j}^{m,\varepsilon_i} = H\left(\varepsilon_i - \left\|\vec{\mathbf{x}}_i - \vec{\mathbf{x}}_j\right\|\right), i,j \in [1, N-(m-1)\tau] \tag{1}$$

where $H()$ is the Heaviside function, and ε is a predefined threshold (1.3 times the standard deviation of the series), and $\vec{\mathbf{x}}_\mathbf{i}$, $\vec{\mathbf{x}}_\mathbf{j}$ are phase space trajectories in m-dimension phase space [20,21], these trajectories can be reconstructed from single time series by using a time delay $\tau$ [22,23].

In this study, 11 RQA parameters (listed in Table 1) are used to characterize those non-linear and non-stationary bearing states, and detailed parameters are listed in Table 3.

**Table 1**
The extract eleven-feature parameters.

| Feature parameters | Equation | Feature parameters | equation |
| --- | --- | --- | --- |
| 1.Recurrence Rate | $RR = \frac{1}{N^2}\sum_{i,j=1}^{N}\mathbf{R}_{i,j}^{m,\varepsilon}$ | 7.Trapping Time | $TT = \sum_{v=v_{\min}}^{N} v\mathbf{P}^{\varepsilon}(v)/\sum_{v=v_{\min}}^{N}\mathbf{P}^{\varepsilon}(v)$ |
| 2.Determinism | $DET = \sum_{l=l_{\min}}^{N} l\mathbf{P}^{\varepsilon}(l)/\sum_{i,j=1}^{N}\mathbf{R}_{i,j}^{m,\varepsilon}$ | 8.Maximal vertical line length | $V_{\max} = \max(\{v_i, i = 1,2,\cdots,N_v\})$ |
| 3.Mean diagonal line length | $L = \sum_{l=l_{\min}}^{N} l\mathbf{p}(l)/\sum_{l=l_{\min}}^{N}\mathbf{p}(l)$ | 9.Recurrence Time of 1st type | $T_j^1 = \left\|\{i,j : \vec{\mathbf{x}}_i, \vec{\mathbf{x}}_j \in \Re_i\}\right\|$ |
| 4.Maximal diagonal line length | $L_{\max} = \max(\{l_i, i = 1,2,\cdots,N_l\})$ | 10.Recurrence Time of 2nd type | $T_j^2 = \left\|\{i,j : \vec{\mathbf{x}}_i, \vec{\mathbf{x}}_j \in \Re_i; \vec{\mathbf{x}}_{i-1} \notin \Re_i\}\right\|$ |
| 5.Entropy | $ENTR = -\sum_{l=l_{\min}}^{N} v\mathbf{p}(l)\ln\mathbf{p}(l)$ | 11.Recurrence Time Entropy | $RTE = -\mathbf{p}(TT)\ln\mathbf{p}(TT)$ |
| 6.Laminarity | $LAM = \sum_{v=v_{\min}}^{N} v\mathbf{P}^{\varepsilon}(v)/\sum_{v=1}^{N} v\mathbf{P}^{\varepsilon}(v)$ | | |

Notes: 1. $\mathbf{P}^{\varepsilon}(l) = \{l_i, i = 1,2,\cdots,N_l\}$ is the frequency distribution of the lengths $l$ of diagonal structures; 2. $\mathbf{P}^{\varepsilon}(v) = \{v_i, i = 1,2,\cdots,N_v\}$ is the frequency distribution of the lengths $l$ of vertical structures; 3. $N_l$ and $N_v$ are the absolute number of diagonal and vertical lines, respectively; 4. $\mathbf{p}(l) = \mathbf{P}^{\varepsilon}(l)/\sum_{l=l_{\min}}^{N}\mathbf{P}^{\varepsilon}(l)$; 5. $\Re_i$ denotes the recurrence points which belong to the trajectory $\vec{\mathbf{x}}_i$.