

DOI: 10. 20040/j. cnki. 1000-7709. 2023. 20222298

RFOA 优化 EEMD 阈值和 SampEn 的水电机组 振动信号重构与特征提取

董利江^{1,2}, 朱霄珣³, 刘 伟², 杨春旭², 林 翔², 高晓霞³, 吕朝阳³, 胡乔良², 苏海鹏²

(1. 新疆新能集团有限责任公司乌鲁木齐电力建设调试所, 新疆 乌鲁木齐 830011; 2. 国网新疆电力有限公司电力科学研究院能源技术研究所, 新疆 乌鲁木齐 830011; 3. 华北电力大学动力工程系, 河北 保定 071003)

摘要: 针对 EEMD 在水电机组振动信号降噪处理中的不足, 提出一种基于改进果蝇算法(RFOA)优化 EEMD 噪声 IMF 分量阈值的降噪算法。通过 EEMD 算法将噪声信号分解, 得到 IMF 分量, 进而通过相关系数法确定噪声信号与有效信号, 利用 RFOA 确定噪声信号 IMF 分量阈值; 将求得的 IMF 分量的样本熵当作特征向量输入 GRNN 算法中, 进行振动模式识别。研究表明, 与小波阈值法、EEMD-GA 方法相比, 所提算法降噪时信噪比最高, 降噪效果最佳。

关键词: 振动信号提取; 集合经验模态分解; 样本熵; 特征提取; 广义回归神经网络模型

中图分类号: [TV734. 1] **文献标志码:** A **文章编号:** 1000-7709(2023)11-0178-05

1 引言

水电机组的振动信号能反映机组的安全运行状态^[1]。全面、有效、准确提取故障信号并识别, 对水电机组故障诊断起重要作用^[2]。非线性和非平稳性是振动信号最明显的特征, 毕杨等^[3]对振动信号进行小波阈值处理, 虽增强了信噪比但缺乏自适应时频分析能力, 容易造成模态混叠现象^[4]; 徐艳春等^[5]将改进的小波阈值降噪法与 EMD 方法相结合应用于振动信号降噪, 降噪效果优于小波阈值法, 由于 EMD 方法存在端点效应和模态混叠, EEMD^[6]方法进行了改进, 有效减弱了 EMD 的不足; 潘光贞等^[7]采用果蝇算法对 IMF 分量进行去噪易陷入局部最优。对于 EEMD 算法的缺点, 本文提出改进果蝇算法(RFOA)优化 EEMD 噪声阈值方法, 通过加入样本熵对振动信号降噪与特征提取, 利用 GRNN^[8]对机组进行振动模式识别; 通过改进 EEMD 阈值(RFOA-EEMD)降噪法进行降噪处理, 将得到的 IMF 分量与有效信号 IMF 分量重构得到降噪后的振动信号, 用 EEMD 算法分解, 最后在 GRNN

算法中输入得到的 IMF, 进行振动模式识别; 实例分析表明与小波阈值法、EEMD-遗传算法相比, 本文算法降噪信噪比最高; 提出的 RFOA-EEMD-SampEn 方法可准确提取不同振动状态特征信号。该方法为后续机组故障诊断奠定了基础。

2 算法基本

2.1 EEMD 算法

EEMD 算法^[9-10]如下。①由于随机噪声多次加入, 使得原始的水电机组振动信号 $X(t)$ 采集待处理的信号 $X_i(t)$:

$$X_i(t) = X(t) + N_i(t) \quad (1)$$

式中, $N_i(t)$ 为第 i 次加入的白噪声信号。②通过 EMD 算法将待处理信号 $X_i(t)$ 分解得到 IMF 分量:

$$X_i(t) = \sum_{j=1}^k C_{i,j}(t) + R_i(t) \quad (2)$$

式中, $C_{i,j}(t)$ 为在分解第 i 信号之后得到的第 j 个分量; $R_i(t)$ 为第 i 个残余量; k 为分量个数。③设 $i = i + 1$, 循环式(1)、(2), 当 $i = M$ 时循环结束; M 为总数平均数。④所得总体平均值为分解结果:

收稿日期: 2022-11-02, **修回日期:** 2023-02-24

基金项目: 重庆市科学基金重点项目(035679); 2002 年高等学校博士学科点专项科研项目(20020183061)

作者简介: 董利江(1986-), 男, 高级工程师, 研究方向为能源与动力、信号处理, E-mail: 309165699@qq.com

通讯作者: 朱霄珣(1985-), 男, 博士、讲师、博导, 研究方向为风功率预测, 电网暂态故障影响及汽轮机故障检测, 风电机电传动系统故障诊断, 信号处理及深度学习智能诊断等, E-mail: zhuxiaoxun@ncepu.edu.cn

$$C_j = \frac{1}{M} \sum_{i=1}^M C_{i,j} \quad (3)$$

式中, C_j 为第 j 个 IMF 分量。

在进行信号分解时,首先需确定白噪声的噪声水平 ϵ 和整体平均数 M 两个参数。关系式为:

$$\delta = \epsilon / \sqrt{M} \quad (4)$$

式中, δ 为初始待测信号标准差。

当 $0.01 \leq \epsilon \leq 0.4$ 、 $100 \leq M \leq 300$ 时, EEMD 算法的分解效果较好^[9]。取 $M=100$, $\epsilon=0.1$ 。

2.2 样本熵原理

样本熵是一种新的度量时间序列复杂度的方法,样本熵和信号复杂度呈正比例关系^[11]。样本熵可以快速捕捉故障信号产生时的复杂度变化,准确分析数据。样本熵计算过程如下。

步骤 1 将 $\{x(n)\}$ 组成一组 m 维向量序列 $X_m(1), X_m(2), \dots, X_m(N-m+1)$, 其中 $X_m(i) = [x(i), x(i+1), \dots, x(i+m-1)]$, 其中 $i=1, 2, \dots, N-m+1$ 。

步骤 2 $d[X_m(i), X_m(j)]$ 为两向量之间的距离,表示 $X_m(i)$ 与 $X_m(j)$ 中最大差值的绝对值。公式为:

$$d[X_m(i), X_m(j)] =$$

$$\max\{|x(i+k-1) - x(j+k-1)|\} \\ k=1, 2, \dots, m; i \geq 1; j \leq N-m+1; i \neq j \quad (5)$$

步骤 3 用 $B_i^m(r)$ 表示统计距离 $d \leq r$ (相似容限) 的数目除以总距离 $N-m+1$, 即:

$$B_i^m(r) = B_i / (N-m) \quad (6)$$

步骤 4 对 $N-m+1$ 个 $B_i^m(r)$ 取平均值得到 $B^m(r)$, 表示为:

$$B^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} B_i^m(r) \quad (7)$$

步骤 5 向量为 $m+1$ 维时,重复步骤 1~4, 得到 $B_i^{m+1}(r)$ 。

步骤 6 在 r 给定条件下,样本熵计算公式为:

$$\text{SampEn}(m, r) = \lim_{N \rightarrow \infty} \left\{ -\ln \left[\frac{B^{m+1}(r)}{B^m(r)} \right] \right\} \quad (8)$$

当 N 为有限值时,计算公式可表示为:

$$\text{SampEn}(m, r, N) = -\ln \left[\frac{B^{m+1}(r)}{B^m(r)} \right] \quad (9)$$

由式(8)、(9)可知,样本熵取决于维数和相似容限。本文取 $m=2$, $r=0.2\text{Std}$ ^[12] (Std 为原时间序列标准差)。

2.3 果蝇算法的改进

果蝇算法中搜索步长 L 固定,若 L 太大,由于收敛精度低,难以得到最优解;反之,收敛时间

增加,会陷入局部最优解^[13];对此,对 FOA 算法引入动态步长搜索策略。可在迭代前期扩大搜索半径,解决相应局部最优解问题;在迭代后期半径减小,提高算法收敛速度;此动态步长改进方法可有效改善果蝇算法的局部及全局寻优能力。 L 计算式为:

$$L = L_{\max} e^{\log\left(\frac{L_{\min}}{L_{\max}}\right)^p c \frac{M}{M_{\max}}} \text{rand} \quad (10)$$

式中, p 为步长调节因子; c 为权重因子; M, M_{\max} 分别为当前迭代次数及最大迭代次数; rand 为 $[-1, 1]$ 间的随机数。

当 $M_{\max}=300$, $L_{\max}=5$, $L_{\min}=0.1$, $p=1.2$, $c=1$ 时, L 的变化曲线见图 1。随着迭代增加 L 减小,动态步长在一定程度上改善了步长为固定值的缺陷。

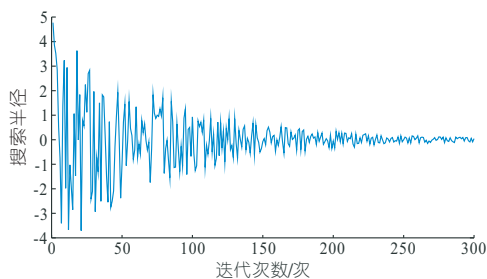


图 1 L 的变化曲线

Fig. 1 L curve of radius search

3 基于 RFOA-EEMD-SampEn 的水电机组振动信号降噪与特征提取流程

3.1 噪声 IMF 的确定

IMF 分量通过 EEMD 算法对原始信号进行分解得到,本文引入皮尔森互相关系数 R 来判断有效分量与噪声分量。其表达式为:

$$R(I_{\text{IMF}_i}, x) = \frac{\text{cov}(I_{\text{IMF}_i}, x)}{\sqrt{\text{COV}(I_{\text{IMF}_i}, I_{\text{IMF}_i})} \sqrt{\text{COV}(x, x)}} \quad (11)$$

在理想状态时,真实信号 R 和 高频噪声分量均为 0。取 $R < 0.1$ 的 IMF 分量作为噪声主导分量^[14],可以区分主导分量 IMF 和有效主导分量 IMF,再利用 Savitzky-Golay 滤波^[15]进一步平滑这两个 IMF 分量。

3.2 噪声 IMF 阈值的选取

利用阈值截断法分割噪声主导分量,确定噪声分量的阈值,并通过 RFOA 算法对阈值进行迭代优化。公式^[16]为:

$$I_{\text{IMF}_i}(t) =$$

$$\begin{cases} \text{sgn}(I_{\text{IMF}_i}(t)) | I_{\text{IMF}_i}(t) - T_i | & | I_{\text{IMF}_i}(t) | > T_i \\ 0 & | I_{\text{IMF}_i}(t) | \leq T_i \end{cases} \quad (12)$$

其中 $T_i = C \sqrt{E_i 2 \ln N}$ (13)
 $E_i = E_1 \rho^{-i} / \beta$

式中, T_i 为噪声信号分量 IMF 的阈值; C 为阈值系数; E_i 为计算能量; N 为信号长度; β, ρ 均为迭代参数。

通过 RFOA 寻优寻求 ρ, β, C 的最优解, 得到每个 IMF 的最佳阈值 T_i 。

3.3 振动信号重构

振动信号重构后的信号 $d(t)$ 为:

$$d(t) = \sum_{k=1}^n I_{\text{IMF}_k}(t) + \text{Res}(t) \quad (14)$$

式中, $\text{Res}(t)$ 为 EEMD 分解的剩余分量。

3.4 信号降噪与特征提取基本流程

信号降噪和特征提取步骤如下。

步骤 1 首先使用 EEMD 集合经验模态分解法对原始振动信号做分解处理, 得到各 IMF 分量及一个剩余模态分量。

步骤 2 计算各 IMF 分量与互相关系数 R , 通过 S-G 滤波平滑信号。

步骤 3 利用 RFOA 算法寻求三个参数 ρ, β, C 的最优解, 并通过阈值滤除带噪声的 IMF 分量。

步骤 4 重构水电机组振动信号 $d(t)$ 。

步骤 5 用 EEMD 算法分解信号 $d(t)$, 求得 IMF 分量。

步骤 6 将 IMF 分量的样本熵作为特征向量。

步骤 7 将所求得特征向量集输入 GRNN 算法中进行识别。

4 仿真试验分析与实例验证

4.1 仿真试验

正常震动机组一般包含 1/2、1、2 倍频等特征信号。对正常信号进行仿真分析数据模型为:

$$\begin{aligned} x(t) = & 0.5 \sin\left(0.5 \times 2\pi \frac{125}{60} t\right) + \\ & 5 \sin\left(2\pi \times \frac{125}{60} t\right) + 10 \sin\left(2 \times 2\pi \frac{125}{60} t\right) + \\ & \text{rand}(1, 2048) \end{aligned} \quad (15)$$

若机组转速为 125 r/min, 采样频率与点数分别为 500 Hz 和 2 048 个, 故障信号为正常信号加上 10 dB。

EEMD 分解算法对端点效应和模态混叠做出了极大调整。

由于 IMF1、IMF3、IMF7、IMF8 互相关系数 $R < 0.1$, 故噪声主导信号分别为 1、3、7、8, 其余信号作为有效信号。Savitzky-Golay 滤波处理所有信号后, 使用 RFOA 算法寻优 ρ, β, C 三个参数的最优解 ($\rho = 2.0103, \beta = 0.7183, C = 0.7001$), 得到 IMF 分量最佳阈值。最后采用阈值截断对噪声分量进行处理^[17], 见图 2。

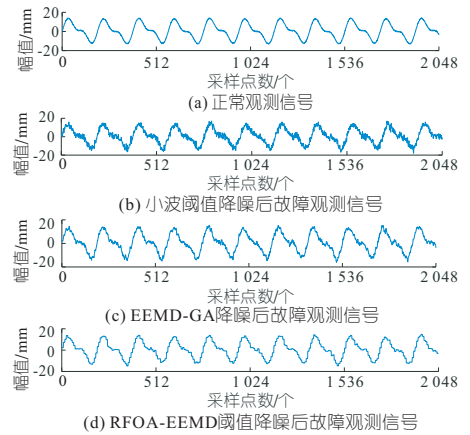


图 2 不同算法降噪效果图

Fig. 2 Noise reduction effect diagram of different algorithms

由图 2 可知, 处理后的小波阈值的振动数据更加密集, 噪声更加明显。EEMD-GA 算法较小波阈值法更优, 但存在一定噪声。总体而言, 小波阈值法和 EEMD-遗传算法的降噪效果不如 RFOA-EEMD 算法。

利用信噪比 S_{SNR} 和均方根误差 (R_{RMSE}) 评估降噪结果。不同算法降噪效果对比见表 1。由表 1 可知, 本文方法信噪比最高, 降噪效果和稳定性均更佳。

表 1 不同算法降噪效果对比

Tab. 1 Comparison of noise reduction effects of different algorithms

降噪方法	R_{RMSE}	S_{SNR}
小波阈值	0.885 2	12.975 1
EEMD-GA	0.677 6	14.727 5
本文方法	0.491 3	16.514 2

4.2 实例验证与结果分析

采集型号为 GZA684-WP-690, 额定功率为 30 MW 的水力发电厂的 #8 水力发电机组在正常运行、报警、故障等三种振动状态下的信号, 每个信号采集 40 个, 每组信号采集长度为 2 048。通过 EEMD 算法分解, 每个模态分量见图 3。

不同方法的降噪效果见表 2。由图 4 与表 2 可知, 本文算法降噪效果最优。

计算可得 IMF1~IMF8 的样品熵值、IMF 元件样品熵特征向量。选择 40 组 EEMD 熵值, 预测输入 GRNN 模型的样本熵值。状态的前 20 组作为训练集后 10 组留作测试集。通过多次试验, 最终取 $\sigma = 0.01$ 。在 GRNN 中输入培训样本后

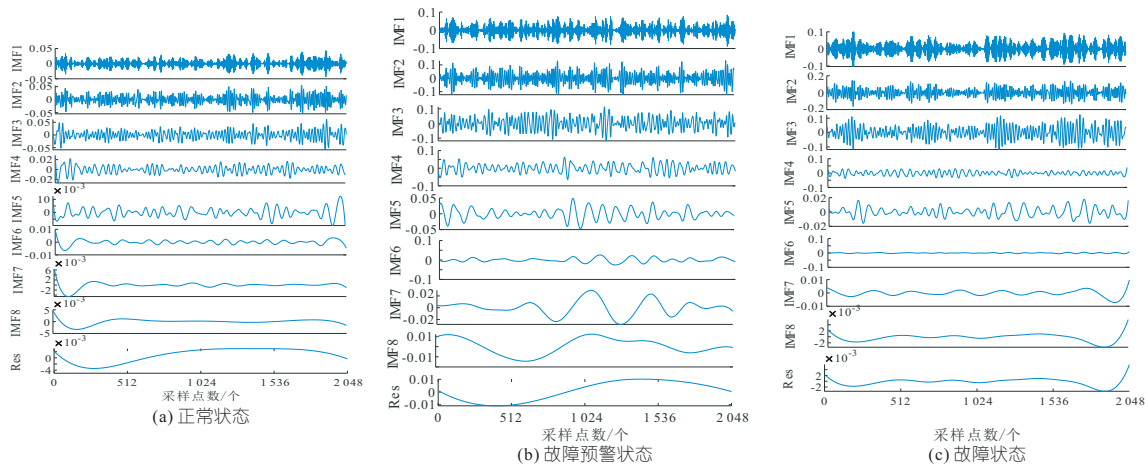


图 3 三种状态 IMF 分量

Fig. 3 Three kinds of IMF components

表 2 不同算法降噪后结果

Tab. 2 Results after noise reduction by different algorithms

降噪方法	正常		故障预警		故障	
	R_{RMSE}	S_{SNR}	R_{RMSE}	S_{SNR}	R_{RMSE}	S_{SNR}
小波阈值	0.037 1	11.011 2	0.048 1	14.580 2	0.035 2	9.661 2
EEMD-GA	0.023 3	17.318 1	0.032 3	19.721 1	0.023 8	17.364 3
本文方法	0.010 7	23.060 3	0.015 8	26.910 2	0.014 4	27.712 4

对结果进行分类,见图 5。图 5 中,状态 1、2、3 分

别表示机组正常工作状态、故障预警工作状态、故障工作状态。如图 5 所示,测试集样本被分类为 GRNN 模型后,预测测试集分类结果与实际测试集分类结果非常相似。GRNN 模型可准确识别所有组正常振动状态。对故障状态和故障修复后的状态识别率在 90% 以上(表 3),可有效识别水电站各种运行状态。

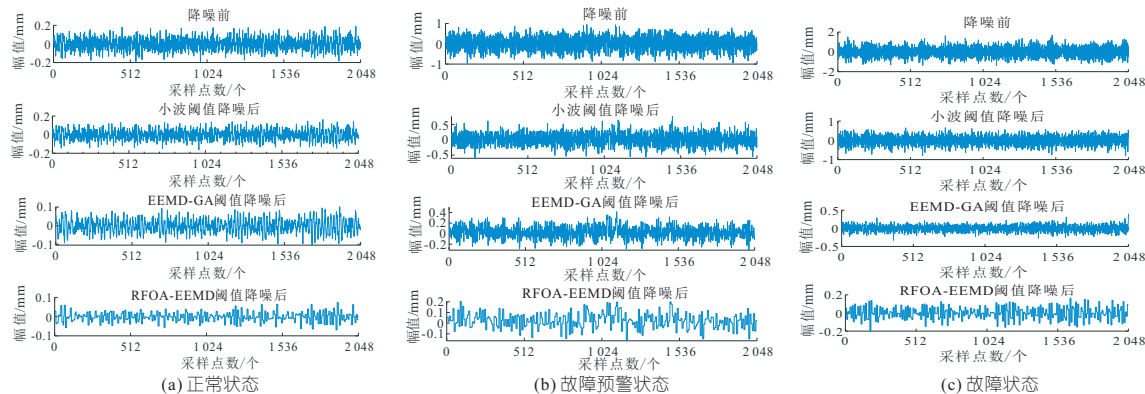


图 4 不同运行状态下机组振动信号降噪前后对比

Fig. 4 Comparison of unit vibration signals before and after noise reduction under different operation modes

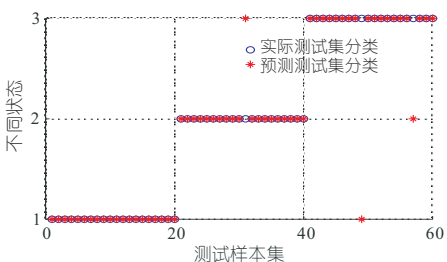


图 5 输入 RFOA-EEMD-SampEn 振动信号分类

Fig. 5 Enter RFOA-EEMD-SampEn vibration signal classification

表 3 不同状态识别效果

Tab. 3 Different state recognition effects

状态	准确率/%	平均识别率/%
正常	100	95
故障	95	
故障维修后	90	

5 结论

- 将果蝇优化算法与 EEMD 相结合,提出信号降噪算法,提高了算法效率,计算互相关系数,改进果蝇算法,IMF 分量阈值问题得到解决。
- 对比小波阈值法、EEMD-遗传算法和所提算法,RFOA-EEMD-SampEn 算法降噪信噪比最高。
- RFOA-EEMD-SampEn 的特征提取方法对振动信号最大识别精度最高,进一步证明该算法降噪效果更好,为后续机组故障诊断工作奠定了基础。

参考文献:

[1] 卢娜,张广涛,刘付鑫,等.基于 LTSA 与谱聚类的

- 水电机组振动故障诊断方法[J]. 武汉大学学报(工学版), 2021, 54(11):1064-1069.
- [2] 赵佳佳, 贾嵘, 武桦, 等. 基于 FastICA-EEMD 的振动信号特征提取[J]. 水力发电学报, 2017, 36(3): 63-70.
- [3] 毕扬, 郑波, 张亚武, 等. 基于 MIC 与 BiGRU 的水电机组振动趋势预测[J]. 水利学报, 2021, 52(5): 612-621, 632.
- [4] 周小龙, 徐鑫莉, 王尧, 等. 基于变分模态分解和最大重叠离散小波包变换的齿轮信号去噪方法[J]. 振动与冲击, 2021, 40(12): 265-274, 289.
- [5] 徐艳春, 方绍晨. EMD 自相关阈值去噪法在水电机组振动信号中的研究[J]. 中国农村水利水电, 2017(7): 189-192, 195.
- [6] WU Z, HUANG N E. Ensemble empirical mode decomposition: A noise-assisted data analysis method[J]. Advances in adaptive data analysis, 2009, 1(1): 1-41.
- [7] 潘广贞, 王凤, 孙艳青. 改进 EEMD 算法在心电信号去噪中的应用[J]. 数据采集与处理, 2018, 33(4): 646-653.
- [8] IMANAHADI AKHLAGHI, SAEED KAHR-OBAEE, KIAN KARAMI NEZHAD. An accurate non-destructive method for determining mechanical properties of plain carbon steel parts using MHL and GRNN[J]. Nondestructive testing and evaluation, 2021, 36(3): 278-296.
- [9] YU DONGYANG, SUN FENGCHANG, DENG XIAOCHUAN, et al. Wind speed time series forecasting method in wind farm based on EEMD[J]. IOP conference series: Earth and environmental science, 2021, 791: 012140.
- [10] ZHAO LIHUA, HONG GUO, WANG ZELONG, et al. Research on fault vibration signal features of GIS disconnecter based on EEMD and kurtosis criterion[J]. IEEE transactions on electrical and electronic engineering, 2021, 16(5): 677-686.
- [11] 施莹, 林建辉, 庄哲, 等. 基于振动信号时频分解-样本熵的受电弓裂纹故障诊断[J]. 振动与冲击, 2019, 38(8): 180-187.
- [12] 但长林, 李三雁, 张彬. 基于样本熵和 SVM 的滚动轴承故障诊断方法研究[J]. 中国测试, 2020, 46(11): 37-42.
- [13] AMBIKA AGGARWAL, DIMRI PRITI, AGARWAL AMIT, et al. Self adaptive fruit fly algorithm for multiple workflow scheduling in cloud computing environment[J]. Kybernetes: The international journal of systems & cybernetics, 2020, 50(6): 1704-1730.
- [14] WANG G, CHEN X Y, QIAO F L, et al. On intrinsic mode function[J]. Advances in adaptive data analysis, 2010, 2(3): 277-293.
- [15] 胡润志. 水电机组非平稳信息特征提取及其振动预测[D]. 邯郸: 河北工程大学, 2021.
- [16] XU PENGFEI, JIA YINJIE, JIANG MINGXIN. Blind audio source separation based on a new system model and the Savitzky-Golay filter[J]. Journal of electrical engineering, 2021, 72(3): 208-212.
- [17] 位秀雷, 林瑞霖, 刘树勇, 等. 改进 EEMD 方法及混沌降噪应用研究[J]. 振动与冲击, 2017, 36(17): 35-41.

Vibration Feature Extraction for Hydropower Units Based on RFOA Optimized Ensemble Empirical Mode Decomposition Threshold and Sample Entropy

DONG Li-jiang^{1,2}, ZHU Xiao-xun³, LIU Wei², YANG Chun-xu², LIN Xiang², GAO Xiao-xia³,
LV Zhao-yang³, HU Qiao-liang², SU Hai-peng²

(1. Urumqi Electric Power Construction and Commissioning Institute of Xinjiang Xinneng Group, Urumqi 830011, China; 2. Institute of Energy Technology, Electric Power Research Institute of State Grid Xinjiang Electric Power, Urumqi 830011, China; 3. Power Engineering Department, North China Electric Power University, Baoding 071003, China)

Abstract: Aiming at the shortcomings of hydropower unit vibration signal denoising using ensemble empirical mode decomposition (EEMD), a denoising algorithm based on an improved fruit fly optimization algorithm (RFOA) for optimizing the EEMD noise IMF component threshold was proposed. Firstly, the noise signal was decomposed using the EEMD algorithm to obtain the IMF components, and then the correlation coefficient method was used to determine the noise signal and the effective signal. Then, the RFOA was used to determine the noise signal IMF component threshold. Finally, the sample entropy of the obtained IMF components was used as a feature vector input of the GRNN algorithm for vibration mode recognition. Compared with the wavelet threshold method and the EEMD-GA method, the results show that the proposed algorithm has the highest signal-to-noise ratio and the best denoising effect.

Key words: extraction of vibration signal; ensemble empirical mode decomposition; sample entropy; feature extraction; generalized regression neural network model