

DOI: 10.19650/j.cnki.cjsi.J2006136

风力发电机组发电机前轴承故障预警及辨识*

尹 诗^{1,2}, 侯国莲¹, 胡晓东¹, 周继威², 弓林娟¹

(1. 华北电力大学控制与计算机工程学院 北京 102206; 2. 中能电力科技开发有限公司 北京 100034)

摘 要:为实现风电机组发电机前轴承故障预警及辨识,将监控和数据采集系统(SCADA)时间序列数据和状态监测系统振动数据相结合,提出了一种时频域建模方法。首先,利用 SCADA 数据建立基于门控循环单元神经网络的发电机前轴承温度模型,并计算其温度残差特征;其次,提取发电机前轴承振动信号时域特征和频域特征;最后,将温度残差特征和振动信号时频域特征相融合,建立基于极限梯度提升的前轴承故障辨识模型,从而辨识发电机前轴承正常、内圈损伤、外圈损伤、轴不平衡、滚动体损伤 5 类情况。实验研究表明,该方法比单独利用振动信号特征开展前轴承故障预警辨识的准确率高,其正常、内圈损伤、外圈损伤的平均辨识准确率从 87%、58.5%、65%,分别提升到 88.5%、67.5% 和 74%。

关键词:风电机组;发电机前轴承;故障辨识;特征提取;时频域建模

中图分类号: TH133.3 TK83 **文献标识码:** A **国家标准学科分类代码:** 480.6040

Fault warning and identification of front bearing of wind turbine generator

Yin Shi^{1,2}, Hou Guolian¹, Hu Xiaodong¹, Zhou Jiwei², Gong Linjuan¹

(1. School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China;

2. Zhongneng Power-Tech Development Co., Ltd., Beijing 100034, China)

Abstract: In order to realize the fault warning and identification of front bearing of wind turbine generator, in this paper a time-frequency domain modeling method is proposed, which integrates the time series data of supervisory control and data acquisition (SCADA) system with the vibration data of condition monitoring system (CMS). Firstly, the temperature model of generator front bearing based on gated recurrent unit (GRU) neural network is established using the SCADA data, and the temperature residual features are calculated. Secondly, the time domain features and frequency domain features of the vibration signal of generator front bearing are extracted. Finally, the temperature residual features and the time-frequency domain features of the vibration signal are fused, and the extreme gradient boosting (XGBoost) based fault identification model of the front bearing is established, which can identify five working conditions of the generator front bearing, including normal, inner ring damage, outer ring damage, shaft imbalance and rolling body damage. Extensive experiment results demonstrate that the proposed method can achieve higher identification accuracy compared with the front bearing fault warning identification method using the vibration signal characteristics alone. The average identification accuracy for normal, inner ring damage and outer ring damage conditions increase from 87%, 58.5% and 65% to 88.5%, 67.5% and 74%, respectively.

Keywords: wind turbine; generator front bearing; fault identification; feature extraction; time-frequency domain modeling

0 引 言

风能作为一种可再生的清洁能源,已被各国广泛开发利用。根据国家能源局发布的数据,2019 年度风电并网新增装机为 2 579 万千瓦,截至到 2019 年 12 月,全国

累计并网风电容量达到 21 005 万千瓦。在全球大型并网型风电机组中,双馈风力发电机组占整个风力发电机组的比重超过 70%。由于风电机组发电机轴承长期运行在变工况和电磁环境中,故障通常由超负荷、润滑不当、电腐蚀、密封失效等造成。尽管发电机类故障在风电机组故障中频次占比不高,但由于设备复杂、停机时间长、

维修成本高,造成经济损失巨大^[1]。因此,在风电机组发电机故障发生前进行有效的预警和识别是风电运营商和设备制造商关注和亟待解决的问题。

当前,风电机组设备故障诊断研究主要集中在监控和数据采集系统(supervisory control and data acquisition, SCADA)数据分析方法、振动频谱分析方法、润滑油检测方法、声发射信号、应变传感信号等检测方法^[2]。齐咏生等^[3]提出了一种改进的自适应变分模态分解(adaptive variational mode decomposition, AVMD)方法用于滚动轴承故障诊断,与经验模态分解(empirical mode decomposition, EMD)、集合经验模态分解(ensemble empirical mode decomposition, EEMD)、小波变换相比,具有分解速度快、准确度高的特点。刘帅等^[4]将高斯混合模型、动态时间规整及熵权值算法三者紧密结合,提出了一种基于群体多维特征相似性的故障预警策略,能提前预警风电机组齿轮箱潜在故障。董兴辉等^[5]以风电机组轴承为研究对象应用最小二乘曲面拟合算法,建立轴承健康状态劣化模型,并利用EEMD将轴承劣化趋势分解并预测。赵洪山等^[6]提出一种逐层编码网络的深度学习方法,选用阈值检测误差趋势变化作为判定风电机组主轴承故障的决策准则。王梓齐等^[7]提出了一种基于模糊软聚类 and 集成非线性状态估计(nonlinear state estimate technology, NSET)的风电机组齿轮箱状态监测方法,通过模型残差均值和基于残差构造的健康指数进行故障预警。顾煜炯等^[8]将齿轮箱振动非平稳时域信号转化成平稳角域信号,提出了无量纲指标趋势分析法以实现风电机组齿轮箱故障早期诊断。Fuentes等^[9]提出使用声发射监测和高斯混合概率模型方法来检测发电机轴承亚表面损伤问题。

以上这些研究方法各有特点,在不同程度上对风电机组关键核心部件状态监测和故障预警起到了推动促进作用,但在以下方面仍需进一步拓展研究。

1) 在利用SCADA数据进行分析建模的过程中,通常只是考虑模型输入与输出的关系,往往忽略了SCADA数据中时间参数对状态监测模型和故障预警影响,导致模型建立的精度不高。

2) 在设备异常决策或者异常样本选择过程中,大部分算法往往通过对阈值的人为设定或者选择异常发生前一段时间的SCADA数据作为异常样本数据,决策模型增加了建立的主观性,降低了模型的可信度和泛化能力。

3) 在状态监测系统(condition monitoring system, CMS)振动信号分析中,往往是通过振动信号的时频域特征进行建模分析,但由于故障产生的随机性和前期征兆的不确定性,一种故障往往由多种原因造成,单一SCADA数据或者振动信号数据往往存在故障诊断误报率高,且很难覆盖风电机组某一部件内的各类故障。

综上所述,本文在总结前人的基础上,根据发电机前轴承实际运行过程中主要故障(内圈损伤、外圈损伤、轴不平衡、滚动体损伤),提出了将发电机前轴承温度残差特征和发电机前轴承振动数据时频域特征融合,建立故障预警辨识模型,从而辨识发电机前轴承正常、内圈损伤、外圈损伤、轴不平衡、滚动体损伤5种状态。

1 建模方法和思路

风电机组发电机前轴承故障预警及辨识的整体建模方法如图1所示。

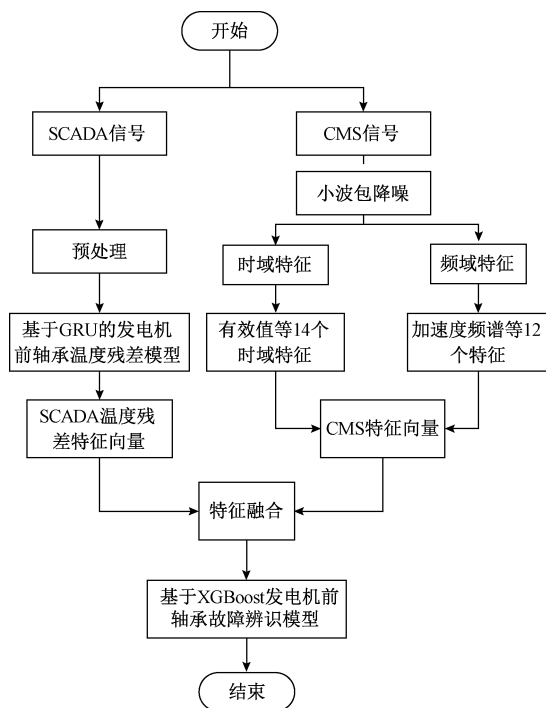


图1 整体建模方法流程

Fig.1 Flow chart of the overall modeling method

首先,利用温度特征建立基于门控循环单元(gated recurrent unit, GRU)神经网络的发电机前轴承温度模型,并计算SCADA数据与温度模型输出之间的残差。

温度异常是判断机组关键设备异常的重要征兆,发电机轴承在运行过程中由于磨损、润滑不良、屏蔽不良等原因会导致过热,过高的温度会影响设备正常运行^[10]。发电机前轴承温度模型建立前需要对参与模型训练和测试的SCADA数据进行预处理,并确定温度模型的输入和输出。

其次,将振动信号进行小波包降噪,提取发电机前轴承振动信号的时域特征和频域特征。按照振动信号时间序列找到对应时间区间的发电机前轴承温度残差10 min均值,与之结合最终得到后续参与辨识模型的输入特征

向量。振动信号特征提取分为时域和频域,时域特征提取加速度和速度高频振动信号的有效值、峰峰值、峰值指标、波形指标、脉冲指标、裕度指标、峭度指标共 14 个时域特征向量;频域特征提取加速度频谱、速度频谱、加速度包络在内的频域特征共 12 个频域特征向量。输出特征向量按照诊断平台标记的故障类型进行标记,分为 5 类状态,分别是正常、发电机前轴内圈损伤、发电机前轴外圈损伤、发电机轴不平衡、发电机滚动体损伤。

最后,建立基于极限梯度提升(extreme gradient boosting, XGBoost)的发电机前轴承故障辨识模型,从而辨识发电机前轴承 5 类状态。在不同算法下进行对比分析,并测试增加温度残差特征作为发电机前轴承辨识模型输入时对测试样本数据准确率的影响。

GRU 神经网络能够保存长期时间序列中的信息,因此更加适用于故障劣化渐变过程的时序数据建模问题。将发电机前轴承温度残差特征与振动信号时域特征和频域特征相结合能够最大限度挖掘温度特征和振动特征,可以避免单一数据对决策结果的局限和偏差,提升了建模精度。利用诊断平台将发电机前轴承故障进行确认标记,降低了利用 SCADA 数据建模过程中人为设定阈值参与模型决策的情况,同时也降低了选择异常发生之前一段时间的 SCADA 数据作为异常样本数据的不确定性,提升了决策模型的可信度和泛化能力。

2 方法原理及实现

2.1 基于 GRU 神经网络的发电机前轴承温度模型

1) 数据预处理

由于风电机组所处运行环境受机组故障、传感器损坏、SCADA 系统宕机、数据传输网络等因素造成 SCADA 数据中夹杂着很多异常噪声数据。因此,首先对 SCADA 数据进行预处理,以提升后续建立模型的准确性。

首先,将数据集进行单点阈值过滤,将超出阈值范围数据点进行删除,并将风电机组故障停机、弃风限电停机工况下无法体现温度及振动特征的数据剔除;其次,按拉依达准则(3σ 准则)进行去除离奇值,消除大幅度的脉冲干扰,并在此基础上进行中值滤波直至数据满足要求;最后将参与训练和测试的数据进行归一化处理,计算公式如下:

$$x_n = \frac{x - x_{\text{mean}}}{x_{\text{max}} - x_{\text{min}}} \quad (1)$$

式中: x 为每一个 SCADA 数据; x_{mean} 为数据的均值; x_{max} 为数据的最大值; x_{min} 为数据的最小值; x_n 为归一化以后的数据。

风电机组实际运行过程中振动信号夹杂大量干扰和噪声,需要对振动信号进行预处理,采用基于小波包对振

动信号进行降噪,其步骤如下:

- (1) 选择小波基并确定分解层次,对振动信号进行小波包分解;
- (2) 通过熵标准确定最优小波包基;
- (3) 对于每一个小波包分解系数,选择恰当的阈值对系数进行量化;
- (4) 将振动信号进行小波包重构。

小波包能够对低频部分和低频部分同时进行分解,能有效区分振动信号中突变部分和噪声,优于小波降噪后处理的结果。

2) 发电机前轴承温度模型

神经网络是由很多神经元连接而成的高度复杂的非线性信息处理系统,具有强大的学习能力。在传统神经网络中,模型只关注当前时刻信息,不关注上一时刻信息对下一时刻的影响。长短期记忆(long short-term memory, LSTM)神经网络通过门控机制使神经网络不仅能记忆以往的信息,还能选择将不重要的信息进行遗忘,使其具备时序数据处理能力,网络结构包括输入门、输出门、遗忘门和记忆单元^[11-12]。

GRU 神经网络是 LSTM 神经网络的加强版,其原理与 LSTM 相似,利用门控机制控制输入、记忆等信息并在当前时刻预测输出。GRU 结构中具有重置门(reset gate)和更新门(update gate),重置门决定了如何将新的输入信息与前记忆信息相结合,更新门定义了前记忆信息保存到当前时刻的信息量。两个门控向量决定了哪些信息最终作为门控循环单元的输出,使其能够保存长时间序列中的信息,且不会随着时间或者预测不相关而移除^[13]。GRU 神经网络结构如图 2 所示。

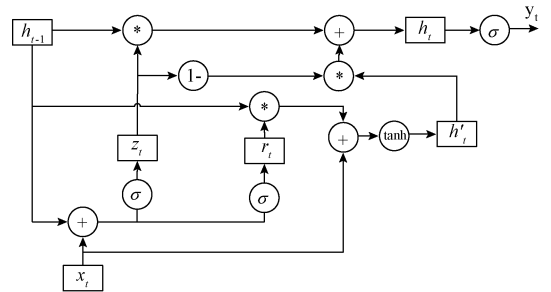


图 2 GRU 神经网络结构

Fig.2 The structure diagram of GRU neural network

对于每个 GRU 单元,给定输入序列 $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_t)$, 通过式(2)和(3)获取两个门控信号,即:

$$\mathbf{z}_t = \sigma(\mathbf{W}^{(z)} \mathbf{x}_t + \mathbf{U}^{(z)} \mathbf{h}_{t-1}) \quad (2)$$

$$\mathbf{r}_t = \sigma(\mathbf{W}^{(r)} \mathbf{x}_t + \mathbf{U}^{(r)} \mathbf{h}_{t-1}) \quad (3)$$

式中: $\mathbf{z}_t$ 为更新门; $\mathbf{r}_t$ 为重置门; $\mathbf{x}_t$ 为第 t 个时间的输入向量; $\mathbf{U}^{(z)}$ 、 $\mathbf{U}^{(r)}$ 、 $\mathbf{W}^{(z)}$ 、 $\mathbf{W}^{(r)}$ 为权重矩阵; $\mathbf{h}_{t-1}$ 为 $t-1$ 时刻记

忆信息; σ 为 Sigmoid 激活函数。

使用重置门得到新的记忆信息 h'_t 为:

$$h'_t = \tanh(Wx_t + r_t * Uh_{t-1}) \tag{4}$$

式中:“ $*$ ”表示矩阵中对应的元素相乘; W 、 U 为权重矩阵; $\tanh$ 为双曲正切函数。

最后,计算最终 GRU 输出记忆信息 h_t :

$$h_t = z_t * h_{t-1} + (1 - z_t) * h'_t \tag{5}$$

$$y_t = \sigma(W_o h_t) \tag{6}$$

式中: W_o 为权重矩阵; y_t 为输出层的输出。

GRU 神经网络作为 LSTM 神经网络的变体,简化了网络结构,解决了 LSTM 神经网络训练时间长、参数较多、内部计算复杂的缺点。目前,GRU 神经网络在语义情感分类、自然语言处理、交通流量预测、电力负荷预测等领域已有应用案例^[14-18],但在故障预警及辨识相关领域研究较少,因此本文将 GRU 神经网络引入,以解决故障劣化渐变过程的时序数据建模问题。

2.2 基于振动信号的时频域分析

风电机组发电机前轴承加速度高频振动信号由 CMS 诊断平台采集,数据和诊断结果均通过专业人员和风电场审核确认。加速度高频振动信号通过傅里叶变换得到加速度频谱,进行希尔伯特变换得到加速度包络频谱,对加速度信号进行积分得到速度振动信号,对其进行傅里叶变换得到速度频谱。

1) 基于振动信号的时域特征提取

时域特征提取是振动信号分析中最为直接便捷的特征提取方法,本文采用的发电机前轴承振动信号的 7 个时域特征指标及其表达式如表 1 所示。

表 1 时域特征指标

Table 1 Time domain characteristic indexes	
特征名称	表达式
有效值	$x_{rms} = \sqrt{\frac{1}{N} \sum_{n=1}^N x^2(n)}$
峰峰值	$x_{p-p} = x_{max} - x_{min}$
峰值指标	$C_f = x_{max}/x_{rms}$
波形指标	$S_f = x_{rms} / \left(\frac{1}{N} \sum_{n=1}^N x(n) \right)$
脉冲指标	$I_f = x_{max} / \left \frac{1}{N} \sum_{n=1}^N x(n) \right $
裕度指标	$CL_f = x_{max} / \left(\left(\frac{1}{N} \sum_{n=1}^N \sqrt{ x(n) } \right) \right) 2$
峭度指标	$K_V = \frac{1}{N} \sum_{n=1}^N x^4(n) / x_{rms}^4$

注: x 为特征序列向量; N 为特征序列向量长度; x_{max} 、 x_{min} 、 x_{rms} 分别为特征序列向量最大值、最小值、有效值。

在时域特征指标中,有效值反映振动信号的强度和

能量,适用于轴承磨损产生的不规则连续性缺陷引起的振动;峰峰值指标反映轴承局部故障点产生冲击振动的大小;峰值指标反映振动信号波形尖峰程度,能够反映出风电机组发电机轴承的擦痕、刻伤等故障;波形指标反映发电机轴承点蚀等故障;脉冲指标和裕度指标在故障早期呈现明显增加趋势,对冲击类故障较为敏感;峭度指标反映了信号偏离正态分布的程度,峭度值越大,说明轴承受到的冲击程度加剧^[19]。本文提取的时域特征包含加速度振动信号、速度振动信号在内的 7 个时域指标共计 14 个时域特征。

2) 基于振动信号的频域特征提取

轴承振动频谱反映信号在不同频率下幅值分布情况,发电机轴承正常机组和异常机组的频谱结构有很大的差别,通过频谱分析可以判断轴承是否发生故障及故障部位。根据傅里叶级数理论,假设离散化后的振动信号时间序列,采样频率为 f_s ,采样点数为 N ,频谱为 $s(k)$, $k = 1, 2, \dots, K$, K 为谱线数。本文采用 4 个频域指标,其表达式如表 2 所示。

表 2 频域特征指标

Table 2 Frequency domain characteristic indexes	
特征名称	表达式
F_1	$F_1 = \sum_{k=1}^K s(k) / K$
F_2	$F_2 = \sum_{k=1}^K (s(k) - F_1)^2 / (K - 1)$
F_3	$F_3 = \sum_{k=1}^K (s(k) - F_1)^3 / (K (\sqrt{F_2})^3)$
F_4	$F_4 = \sum_{k=1}^K (s(k) - F_1)^4 / (K (F_2)^2)$

注:频域特征指标 F_1 反映振动能量的大小; F_2 反映信号频谱与信号频谱均值的偏离程度; F_3 反映了信号频谱相对均值不对称的程度; F_4 反映了信号在均值处峰值的大小。

除了频谱能很好的反映风电机组运行时出现的故障信息外,包络频谱也能很好的反映设备冲击类故障信息。包络频谱常常用来检测轴承缺陷,当轴承表面因疲劳或者集中应力产生剥落或者损伤时,就会产生周期性的冲击振动信号。这种冲击振动可大致分为由于轴承部件在运行过程中反复受到冲击而产生的低频脉冲和冲击振动产生的固有振动,因此对包络频谱分析就能精准的诊断出典型轴承故障^[20]。本文提取的频域特征包含加速度频谱、速度频谱、加速度包络频谱在内的 4 个频域指标 (F_1 、 F_2 、 F_3 、 F_4) 共计 12 个频域特征。

2.3 基于 XGBoost 的发电机前轴承故障辨识

集成算法是通过构建并结合多个学习器完成学习任务。XGBoost 算法是在自适应增强 (adaptive boosting, adaBoost) 算法和梯度提升迭代决策树 (gradient boosting

decision tree, GBDT)算法基础上优化形成,通过不断拟合优化目标函数,从而达到准确预测分类结果。XGBoost 算法在原函数的基础上增加了正则项,减少了过拟合的可能性,同时加快了收敛速度^[21-22]。XGBoost 是一个树集成模型,将 K 颗树的结果进行求和作为最终的预测值。算法目标函数定义如下^[23]:

$$L(\varphi) = \sum_{i=1}^n l(y_i, \tilde{y}) + \sum_{k=1}^K \Omega(f_k)$$

(7)

$$\Omega(f_k) = \gamma T + \frac{1}{2} \|\omega\|^2$$

(8)

式(7)中,等式右边第一部分是模型的训练误差,第二部分是正则化项。 f_k 为第 k 树的模型;正则化项是 K 颗树正则化项相加。式(8)中, T 为叶子节点数; ω 为叶子权重值; γ 为叶子惩罚系数。

每次建树优化如下目标:

$$\tilde{L}^{(t)} = \sum_{i=1}^n \left[g_i f_i(x_i) + \frac{1}{2} h_i f_i^2(x_i) \right] + \Omega(f_t)$$

(9)

其中:
$$g_i = \partial_{\tilde{y}_i} l(y_i, \tilde{y}_i^{(t-1)}), h_i = \partial_{\tilde{y}_i}^2 l(y_i, \tilde{y}_i^{(t-1)}),$$

$$\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2$$

最终得到叶子权重值为:

$$\omega_j^* = - \frac{G_j}{H_j + \lambda}$$

(10)

式中: G_j 叶子结点 j 所包含样本的一阶偏导数累加之和; H_j 叶子结点 j 所包含样本的二阶偏导数累加之和; λ 为叶子权重惩罚系数。

最终目标值为:

$$\tilde{L}^* = - \frac{1}{2} \sum_{j=1}^T \frac{G_j^2}{H_j + \lambda} + \gamma T$$

(11)

基于 XGBoost 的发电机前轴承状态辨识模型的输入特征向量如表 3 所示。

表 3 辨识模型的输入特征向量
Table 3 Input eigenvectors of the identification model

名称	特征
SCADA 数据特征(1 个)	发电机前轴承温度残差
振动信号时域特征(14 个)	加速度有效值、加速度峰峰值、加速度波形指标、加速度峰值指标、加速度脉冲指标、加速度裕度指标、加速度峭度指标、速度有效值、速度峰峰值、速度波形指标、速度峰值指标、速度脉冲指标、速度裕度指标、速度峭度指标
振动信号频域特征(12 个)	加速度频谱特征 F_1 、加速度频谱特征 F_2 、加速度频谱特征 F_3 、加速度频谱特征 F_4 、速度频谱特征 F_1 、速度频谱特征 F_2 、速度频谱特征 F_3 、速度频谱特征 F_4 、加速度包络频谱特征 F_1 、加速度包络频谱特征 F_2 、加速度包络频谱特征 F_3 、加速度包络频谱特征 F_4

XGBoost 采用基于决策树的梯度提升算法,其优点是特征在构建提升树的过程中可以自动获取特征的重要性。一般来说如果一个特征在所有树中作为划分属性的次数越多,那么该特征就越重要。因此,XGBoost 在训练过程中可以自动计算出输入特征向量中各个特征的评分,从而能够有效地进行特征筛选。

辨识模型输出分为 5 类状态,分别是正常、发电机前轴承内圈损伤、发电机前轴承外圈损伤、发电机轴不平衡、发电机滚动体损伤,对应的数据标签分别是 0、1、2、3、4,并按照式(1)进行归一化处理。

发电机前轴承故障辨识的判断标准是:辨识模型输出为分类器 5 类状态的辨识准确率,选择分类器输出占比最大的类别作为辨识的最终决策依据。

3 案例分析及验证

本实验所采用的 SCADA 数据和 CMS 振动数据为 2019 年 1 月到 2019 年 6 月,选取的风电机组为河北某风

电场 1.5 MW 双馈异步发电机组,共计 46 台;其中包含 20 台发电机前轴承正常机组,26 台发电机前轴承异常机组,在发电机前轴承异常机组中包含发电机前轴承内圈损伤故障机组 10 台,发电机前轴承外圈损伤故障风机 10 台,发电机轴不平衡机组 3 台,发电机轴滚动体损伤机组 3 台。

GRU 神经网络中选取发电机前轴承正常的 10 台机组进行模型的训练和和测试。SCADA 系统采集的数据包括,风速、发电机转速、叶轮转速、风向角、偏航角度、齿轮箱油温等近百个参数。本文通过研究与发电机运行状态相关性,选取了风速、发电机转速、环境温度、机舱温度、A 相电流、无功功率、有功功率、桨角、发电机后轴承温度 9 个指标的 10 min 数据最小值、最大值、平均值共 27 个参数作为温度模型输入,以发电机前轴承温度 10 min 平均值作为温度模型输出。按照 2.1 节论述的步骤进行数据预处理后,基于 GRU 神经网络的发电机前轴承温度模型的训练和测试的 SCADA 数据共计 25 946 条。发电机前轴承振动加速度信号为

12 800 Hz,采样点选择 12 800,采用 db1 小波对振动信号进行 3 层小波包的分解与重构,降噪前和降噪后的振动信号如图 3 所示。

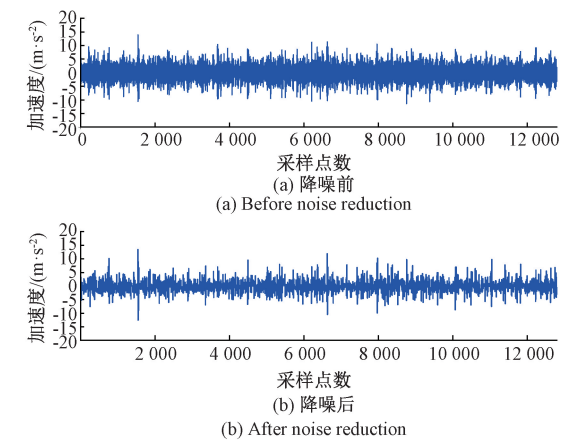


图3 降噪前和降噪后的振动信号
Fig.3 Vibration signal diagrams before and after noise reduction

图3 中上半部分为原始振动信号,下半部分为利用小波包降噪后的信号。从图3 中可以看出,经过小波包降噪后的振动信号,保留了信号中突变部分,并有效地消除了噪声成分。

GRU 神经网络输入层有 27 个神经元(分别是九个指标的最大值、最小值、平均值),隐层设置 3 层记忆细胞,其中第 1 层 100 个神经元,第 2 层有 35 个神经元,第 3 层有 15 个神经元,输出层共 1 个神经元,激活函数选用双曲正切函数(tanh),误差选用均方误差,批处理数量为 200,超参数设定为样本训练次数为 30,Dropout 的值设为 0.2,用于防止网络训练过拟合,训练数据样本和测试数据样本比例为 7:3。基于 GRU 神经网络的发电机前轴承温度模型仿真结果如图 4 所示,其中虚线为温度模型输出,三角形标记的实线分别为 SCADA 训练数据和测试数据,圆形标记的点划线为温度模型输出和 SCADA 数据之间的残差。

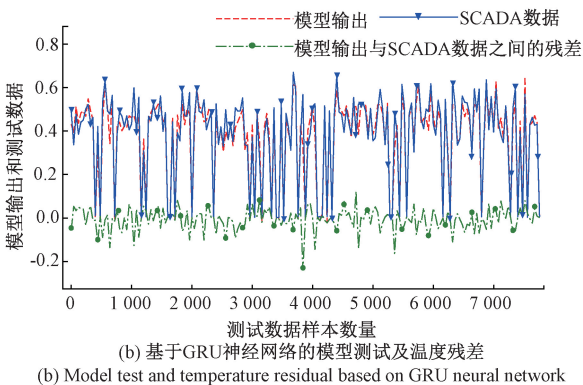
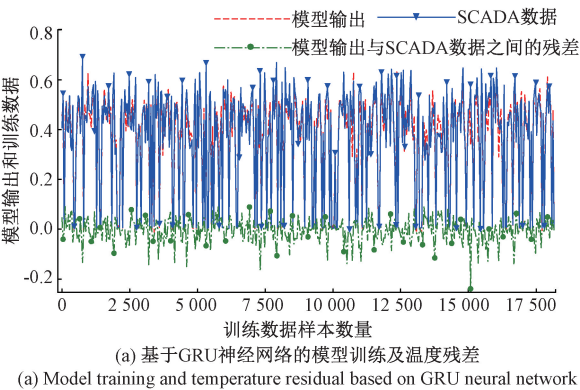


图4 温度模型的输出、训练数据、测试数据及残差
Fig.4 The output, training data, test data and residual of the temperature model

对比分析基于 GRU 神经网络、循环神经网络(recurrent neural network, RNN)、BP(back propagation)神经网络、LSTM 神经网络的发电机前轴承温度模型训练评价指标,结果如表 4 所示。

表4 不同算法在温度模型中的训练评价指标			
Table 4 Training evaluation indexes of temperature model for different algorithms			
算法	时间/s	R^2	均方误差
GRU	46	0.94	0.039
RNN	49	0.92	0.045
BP	36	0.55	0.053
LSTM	107	0.95	0.037

测试结果表明,GRU 神经网络在发电机前轴承温度模型训练中均方误差接近于 LSTM 神经网络,但是在训练时间上大大优于 LSTM 神经网络,因此 GRU 神经网络在模型训练精度和训练时间上较为均衡。

发电机前轴承故障辨识模型选取与温度模型中参与训练和测试不同的剩余 36 台机组。其中辨识模型训练和测试选取发电机前轴承正常机组 8 台,发电机前轴承内圈损伤故障机组 8 台,发电机前轴承外圈损伤故障风机 8 台,发电机轴不平衡机组 3 台,发电机轴滚动体损伤机组 3 台,共计 30 台机组,将振动信号时频域特征和温度特征融合后数据 3 012 条,按 70% 训练集、30% 测试集进行划分。36 台机组中剩余 6 台机组用来最终验证辨识模型的准确率和泛化能力。

按照 2.2 节说明将振动信号提取时频域特征并按照时间序列提取相对应的发电机前轴承温度残差组成新的特征,作为辨识模型训练及测试的输入。

XGBoost 算法的参数选取为最大树的数目($n_{estimator}$)设置为 400、树的最大深度(max_depth)设置为 50、每一步迭代步长($learning_rate$)设置为 0.1、静默模式

(silent) 设置为 True、学习目标参数 objective 设置为 binary:logistic、样本权重(scale_pos_weight) 设置为 1。

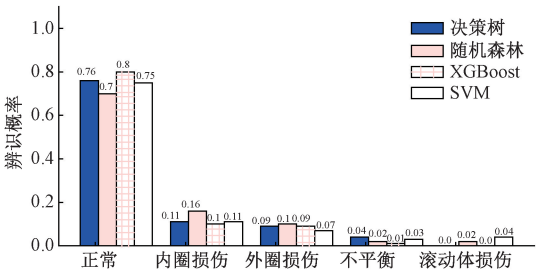
在本文实验中选择决策树、随机森林、支持向量机(support vector machine, SVM)和 XGBoost 算法进行对比分析。观察在不同的算法上实验结果准确率,同时对比分析温度残差特征作为决策模型输入时对测试集辨识准确率的影响,对比结果如表 5 所示。

表 5 测试样本数据辨识准确率

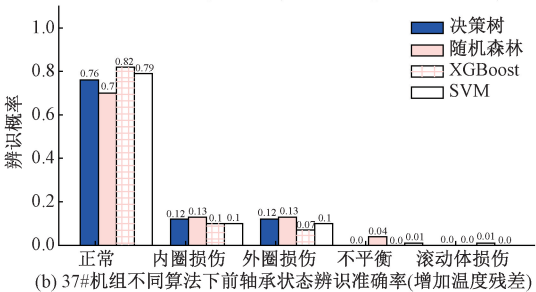
Table 5 Identification accuracy of test sample data		
算法	不加温度残差	增加温度残差
决策树	0.845	0.856
随机森林	0.878	0.881
XGBoost	0.899	0.919
SVM	0.696	0.718

测试结果表明,XGBoost 算法在测试集上的准确率最高,SVM 算法在测试集中的准确率最低,在增加发电机前轴承温度残差特征后辨识模型的准确率均能进一步提升。

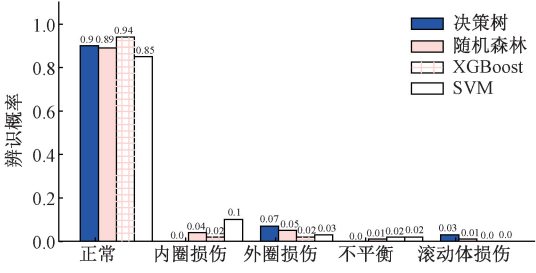
下一步验证辨识模型在剩余 6 台测试机组中的泛化能力。由于轴不平衡和轴滚动体损伤案例较少,6 台测试机组中选择 2 台发电机前轴承正常(37#机组、52#机组)、2 台发电机前轴承内圈损伤(69#机组、85#机组)、2 台发电机前轴外圈损伤(31#机组、93#机组),对比分析决策树、随机森林、XGBoost、SVM 不同算法在发电机前轴承的状态辨识准确率,并进一步对比将温度残差特征和振动信号时频域特征相融合后对辨识精度的影响,测试机组在不同算法下的模型辨识准确率结果如图 5 所示。



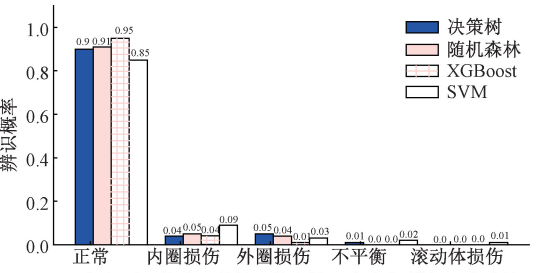
(a) 37#机组不同算法下前轴承状态辨识准确率(无温度残差)
(a) State identification accuracy of front bearing of the wind turbine 37# for different algorithms (without temperature residual)



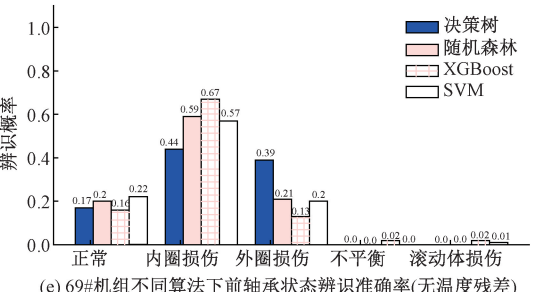
(b) 37#机组不同算法下前轴承状态辨识准确率(增加温度残差)
(b) State identification accuracy of front bearing of the wind turbine 37# for different algorithms (with temperature residual)



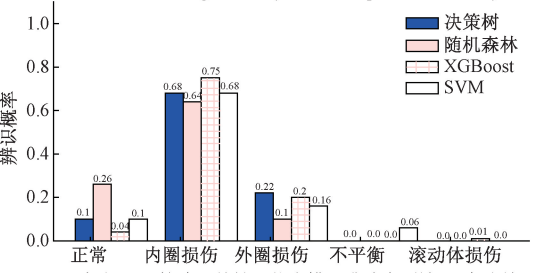
(c) 52#机组不同算法下前轴承状态辨识准确率(无温度残差)
(c) State identification accuracy of front bearing of the wind turbine 52# for different algorithms (without temperature residual)



(d) 52#机组不同算法下前轴承状态辨识准确率(增加温度残差)
(d) State identification accuracy of front bearing of the wind turbine 52# for different algorithms (with temperature residual)



(e) 69#机组不同算法下前轴承状态辨识准确率(无温度残差)
(e) State identification accuracy of front bearing of the wind turbine 69# for different algorithms (without temperature residual)



(f) 69#机组不同算法下前轴承状态辨识准确率(增加温度残差)
(f) State identification accuracy of front bearing of the wind turbine 69# for different algorithms (with temperature residual)

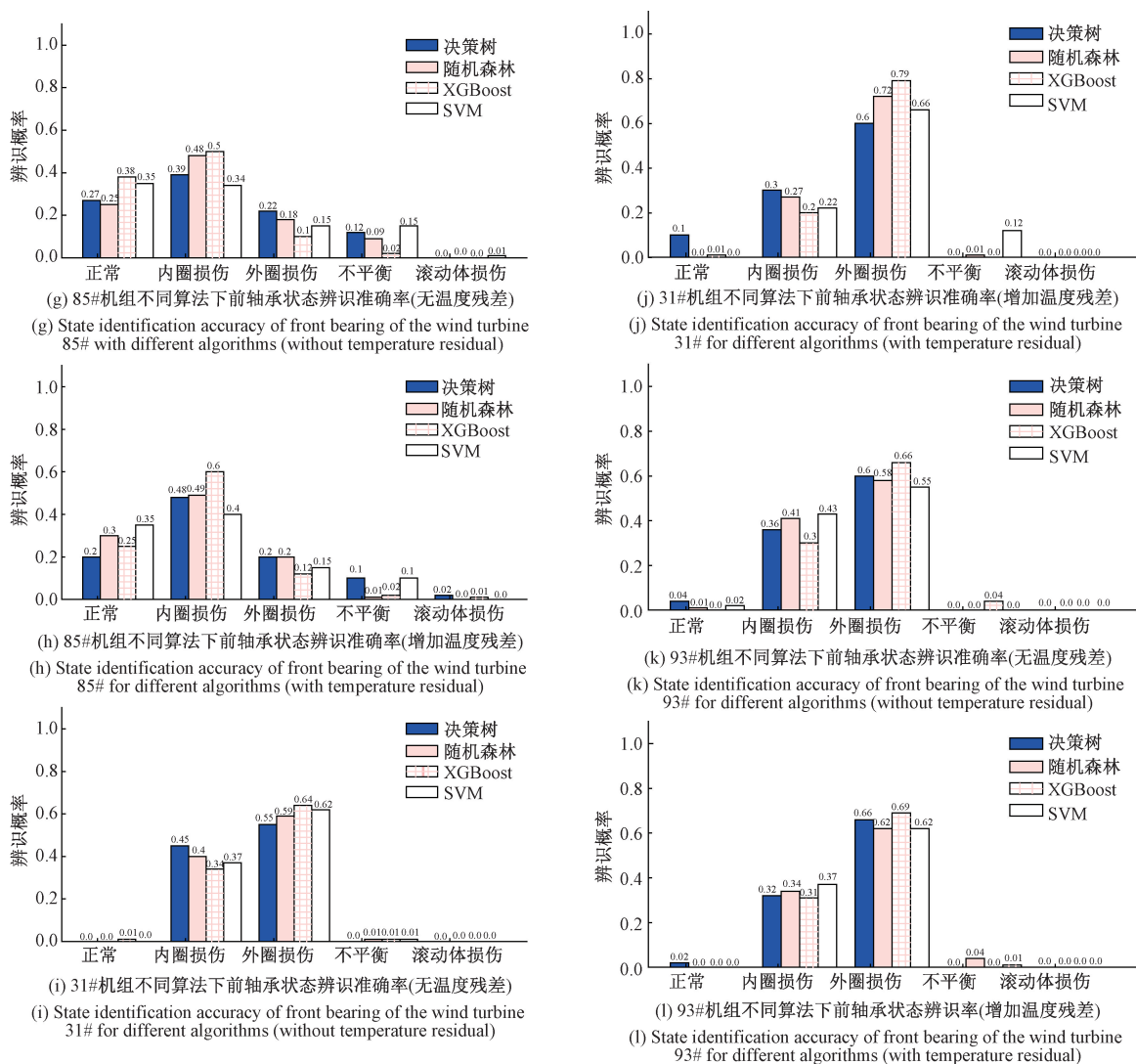


图5 测试机组在不同算法下的模型辨识准确率

Fig.5 Model identification accuracy of test wind turbines for different algorithms

测试结果表明,6台测试机组利用XGBoost算法根据判断标准均能判断正确。XGBoost算法在风电机组发电机前轴承故障辨识中的效果优于决策树、随机森林和SVM算法。在不增加前轴承温度特征时,基于XGBoost的发电机前轴承故障辨识正常机组的平均辨识准确率达为87%;内圈损伤的平均辨识准确率为58.5%;外圈损伤的平均辨识准确率为65%,在增加发电机前轴承温度特征后,测试机组的辨识准确率分别提升到88.5%、67.5%和74%。

在风电场实际应用过程中,首先选取机组连续正常运行的离线历史SCADA数据,经过数据预处理建立基于GRU神经网络的发电机前轴承温度模型,并计算其温度

残差;其次,利用CMS诊断平台选取发电机前轴承正常和异常更换的离线历史振动信号数据,经过小波包降噪,提取其时频域特征;最后,将发电机前轴承振动时频域特征与温度残差特征在同一时间尺度下相融合,建立基于XGBoost的发电机前轴承状态辨识模型。至此,离线模型建立完成。

根据SCADA数据和振动数据的采集情况,离线模型后续可用于该风电场同型号机组的离线分析或在线分析中,若同型号机组在其他风电场故障样本较少,也可将该离线模型推广至同型号机组的不同风电场。随着故障样本逐渐增多,对已训练好的离线模型进行定期增量训练,使模型学习更多的故障模式,提升其辨识精度。

4 结 论

本文以风电机组发电机前轴承为研究对象,提出了一种 SCADA 数据和振动数据相结合的时频域组合建模方法。基于 GRU 神经网络建立风电机组发电机前轴承温度模型,并计算温度残差特征,将该特征和振动信号的时频域特征相结合,建立基于 XGBoost 的发电机前轴承故障辨识模型,从而辨识发电机前轴承故障状态和类别。实验结果表明:

1) 以发电机前轴承温度为特征的 GRU 神经网络能够更好地跟踪识别故障劣化渐变过程的长时间序列数据,提升了温度建模精度和效率。

2) 基于 XGBoost 的发电机前轴承故障辨识,相较于决策树、随机森林、SVM 算法能更好的辨识发电机前轴承正常、内圈损伤、外圈损伤、轴不平衡、滚动体损伤 5 类状态。

3) 将发电机前轴承温度模型残差特征和振动数据时频域特征相结合,提升了发电机前轴承故障辨识建模精度和泛化性。

参考文献

- [1] 卢源. 双馈式风电机组发电机轴承失效分析及预防措施[J]. 设备管理与维修, 2020(1): 71-72.
LU Y. Bearing failure analysis and preventive measures of doubly fed wind turbine generator[J]. Plant maintenance engineering, 2020(1): 71-72.
- [2] 曾军, 陈艳峰, 杨苹, 等. 大型风力发电机组故障诊断综述[J]. 电网技术, 2018, 42(3): 849-860.
ZENG J, CHEN Y F, YANG P, et al. Review of fault diagnosis methods of large-scale wind turbines[J]. Power system technology, 2018, 42(3): 849-860.
- [3] 齐咏生, 白宇, 高胜利, 等. 基于 AVMD 和谱相关分析的风电机组轴承故障诊断[J]. 太阳能学报, 2019, 40(7): 2053-2063.
QI Y SH, BAI Y, GAO SH L, et al. Fault diagnosis of wind turbine bearing based on AVMD and spectral correlation analysis [J]. Acta Energiæ Solaris Sinica, 2019, 40(7): 2053-2063.
- [4] 刘帅, 刘长良, 甄成刚, 等. 基于群体多维相似性的风机齿轮箱预警策略[J]. 仪器仪表学报, 2018, 39(1): 180-189.
LIU SH, LIU CH L, ZHEN CH G, et al. Fault warning strategy of wind turbines gearbox based on group multi-dimensional similarity [J]. Chinese journal of scientific instrument, 2018, 39(1): 180-189.
- [5] 董兴辉, 马晓双, 程友星, 等. 风电机组轴承健康劣

化趋势建模与仿真[J]. 系统仿真学报, 2019, 31(1): 151-157, 165.

DONG X H, MA X SH, CHENG Y X, et al. Modeling and simulation of health degradation trend for wind turbine bearing[J]. Journal of system simulation, 2019, 31(1): 151-157, 165.

- [6] 赵洪山, 刘辉海. 基于深度学习网络的风电机组主轴轴承故障检测[J]. 太阳能学报, 2018, 39(3): 588-595.
ZHAO H SH, LIU H H. Fault detection of wind turbine main bear based on deep learning network [J]. Acta Energiæ Solaris Sinica, 2018, 39(3): 588-595.

- [7] 王梓齐, 刘长良, 刘帅. 基于集成 NSET 和模糊软聚类的风电机组齿轮箱状态监测[J]. 仪器仪表学报, 2019, 40(7): 138-146.

WANG Z Q, LIU CH L, LIU SH. Condition monitoring of wind turbine gearbox based on ensemble nonlinear state estimation technique and soft fuzzy clustering [J]. Chinese Journal of Scientific Instrument, 2019, 40(7): 138-146.

- [8] 顾煜炯, 贾子文, 尹传涛, 等. 无量纲指标趋势分析法在风电机组齿轮箱故障预警诊断中的应用[J]. 振动与冲击, 2017, 36(19): 213-220.

GU Y J, JIA Z W, YIN CH T, et al. Non-dimensional parameters trend analysis method in the early warning and diagnosis of wind turbine gearbox failures[J]. Journal of Vibration and Shock, 2017, 36(19): 213-220.

- [9] FUENTES R, DWYER R S, MARSHALL M B, et al. Detection of sub-surface damage in wind turbine bearings using acoustic emissions and probabilistic modelling[J]. Renewable Energy, 2020, 147(1): 776-797.

- [10] 杨文珺, 何婷, 王贵军. 风力发电机组的故障及维护策略[J]. 机械研究与应用, 2019, 32(5): 190-193, 198.

YANG W J, HE T, WANG G J. Faults and maintenance strategy of wind turbine [J]. Mechanical research & application, 2019, 32(05): 190-193, 198.

- [11] GREFF K, SRIVASTAVA R K, KOUTNIK J, et al. LSTM: A search space odyssey [J]. IEEE Transactions on Neural Networks & Learning Systems, 2015, 28(10): 2222-2232.

- [12] 于洋, 何明, 刘博, 等. 基于 TL-LSTM 的轴承故障声发射信号识别研究[J]. 仪器仪表学报, 2019, 40(5): 51-59.

YU Y, HE M, LIU B, et al. Research on acoustic emission signal recognition of bearing fault based on TL-LSTM [J]. Chinese Journal of Scientific Instrument,

- 2019, 40(5): 51-59.
- [13] CHUNG J, GULCEHRE C, CHO K H, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[J]. Eprint Arxiv, 2014(12): 1-9.
- [14] JIN F, ZHAO B C. Short-term traffic flow prediction based on road network topology [J]. Journal of Beijing Institute of Technology, 2019, 28(3): 383-388.
- [15] KADARI R, ZHANG Y, ZHANG W N. Gated recurrent unit model for a sequence tagging problem [J]. High Technology Letters, 2019, 25(1): 81-87.
- [16] 王震宇, 张雷. 基于深度卷积和门控循环神经网络的传感器运动识别[J]. 电子测量与仪器学报, 2020, 34(1): 1-9.
- WANG ZH Y, ZHANG L. Deep convolutional and gated recurrent neural networks for sensor-based activity recognition[J]. Journal of Electronic Measurement and Instrumentation, 2020, 34(1): 1-9.
- [17] 余培, 行鸿彦, 刘刚. 中文评论情感分析方法研究[J]. 电子测量与仪器学报, 2018, 32(12): 197-203.
- YU P, XING H B, LIU G. Chinese commentary analysis method [J]. Journal of Electronic Measurement and Instrumentation, 2018, 32(12): 197-203.
- [18] 赵兵, 王增平, 纪维佳, 等. 基于注意力机制的 CNN-GRU 短期电力负荷预测方法[J]. 电网技术, 2019, 43(12): 4370-4376.
- ZHAO B, WANG Z P, JI W J, et al. A short-term power load forecasting method based on attention mechanism of CNN-GRU [J]. Power system technology, 2019, 43(12): 4370-4376.
- [19] 彭进, 王维庆, 王海云, 等. 基于 EEMD 峭度-相关系数准则的多特征量风电机组轴承故障诊断[J]. 可再生能源, 2016, 34(10): 1481-1490.
- PENG J, WANG W Q, WANG H Y, et al. Fault diagnosis method of wind turbine's bearing based on EEMD kurtosis-correlation coefficients criterion and multiple features [J]. Renewable energy resources, 2016, 34(10): 1481-1490.
- [20] 邓巍, 韩斌, 王建亮, 等. 风力发电机轴承故障包络谱扫频分析诊断[J]. 热力发电, 2019, 48(7): 126-130.
- DENG W, HAN B, WANG J L, et al. Envelope spectrum sweeping analysis method for bearing fault diagnosis of wind turbine[J]. Thermal power generation, 2019, 48(7): 126-130.
- [21] CHEN T, GUESTRIN C. XGBoost: A Scalable Tree Boosting System [C]. 22nd ACM SIGKDD international conference on knowledge discovery and data mining, ACM, 2016: 785-794.
- [22] DONG H, XU X, WANG L, et al. Gaofen-3 PolSAR Image Classification via XGBoost and Polarimetric Spatial Information [J]. Sensors, 2018, 18(2): 1-20.
- [23] ZHANG R, LI B B, JIAO B. Application of XGBoost algorithm in bearing fault diagnosis[C]. 2nd International Seminar on Computational Intelligence, Engineering and Technology (SCIET2018), 2018: 531-535.

作者简介



尹诗(通信作者), 2012 年于华北电力大学获得硕士学位, 现为华北电力大学博士研究生, 主要研究方向为风电机组故障预警。

E-mail: yinshi502@163.com

Yin Shi (Corresponding author) received his M. Sc. degree from North China Electric Power University in 2012. Now, he is a Ph. D. candidate in North China Electric Power University. His main research interest includes wind turbine fault warning.



侯国莲, 1988 年、1991 年、1997 年于华北电力大学分别获得学士学位、硕士学位和博士学位。现为华北电力大学教授, 主要研究方向为风电机组故障诊断、复杂系统优化控制、新能源技术及控制策略研究。

E-mail: hgl@ncepu.edu.cn

Hou Guolian received her B. Sc., M. Sc. and Ph. D. degrees all from North China Electric Power University in 1988, 1991 and 1997, respectively. Now, she is a professor in North China Electric Power University. Her main research interests include wind turbine fault diagnosis, optimization and control of complex system, new energy resource technology and control strategy research.