

基于大数据分析的智慧电厂设备状态监测研究

王建设, 李奎, 印伟, 郑斌, 夏晨阳

(国家电投集团协鑫滨海发电有限公司, 江苏 盐城 224000)

摘要: 为提高发电厂设备运行可靠性及安全性, 提出了基于大数据分析的智慧电厂设备状态监测方法。利用无线传感器采集发电厂设备振动信号, 振动信号经小波包分解得到子频信号能量值, 将子频信号的能量与总能量的比例作为态特征向量输入到最小二乘支持向量机中, 完成发电厂设备异常状态分类, 并且利用改进蚁群算法优化分类过程, 获取最佳设备状态监测结果。试验结果表明, 该方法可有效提取到不同设备状态下的信号频谱图, 且设备状态监测结果与实际设备状态结果高度一致; 该方法应用后可极大程度提高智慧电厂设备运行安全性及可靠性, 客户满意度高。

关键词: 大数据分析; 智慧电厂; 设备状态监测; 改进蚁群算法; 支持向量机; 能量值

Research on equipment condition monitoring of smart power plant based on big data analysis

WANG Jianshe, LI Kui, YIN Wei, ZHENG Bin, XIA Chenyang

(GCL Coastal Power Generation Co., Ltd., Yancheng 224000, China)

Abstract: In order to improve the operation reliability and safety of smart power plant equipment, a smart power plant equipment condition monitoring method based on big data analysis is proposed. The wireless sensor is used to collect the vibration signal of smart power plant equipment. The vibration signal is decomposed by wavelet packet to obtain the energy value of sub frequency signal. The ratio of sub frequency signal energy to total energy is input into the least squares support vector machine as a state feature vector to complete the abnormal state classification of smart power plant equipment, and the improved ant colony algorithm is used to optimize the classification process. Obtain the best equipment condition monitoring results. The experimental results show that this method can effectively extract the signal spectrum under different equipment states, and the equipment state monitoring results are highly consistent with the actual equipment state results; The application of this method can greatly improve the operation safety and reliability of smart power plant equipment, with high customer satisfaction.

Key words: big data analysis; smart power plant; equipment condition monitoring; improved ant colony algorithm; support vector machine; the energy value

0 引言

随着物联网、大数据以及智能控制的快速发

展, 使得智慧电厂在电力领域日益重要。为提高用户的用电安全及质量, 需不断提升智慧电厂相

关技术的水平,使其成为更安全可靠、高效环保的智能发电技术^[1-2]。

当前,由于智慧电厂在实际运行中不定时地出现某些故障,这些故障引发某些设备故障,若不能及时发现解决,必然会造成严重的影响及经济损失^[3]。智慧电厂设备状态监测可实时分析发电厂不同设备的工作机能,协助维修人员立即察觉设备故障的位置,减少设备损耗,为发电厂安全运行提供保障。

近年来关于智慧电厂设备状态监测技术的研究成为国内外研究的重点。文献[4]研究的基于双通道图像的智能监测方法,提高了发电厂设备状态监测效率,该方法虽然通过两个图像进行实际监测,但操作过程过于复杂。文献[5]研究的全维度设备状态监测方法^[5],可多维度、全方位地自动监测发电厂设备状态,但收敛效率不高,导致设备状态监测实时性差。

大数据具有海量、实时以及高维的特征,海量数据信息经过大数据分析挖掘,可以产生对用户有价值的信息,利用大数据原理和数据挖掘可实时监测设备状态^[6]。支持向量机(Support Vector Machine, SVM)是一种分类及回归工具,基于统计学习原理开发出来,可利用结构风险最小化理论使泛化能力得到提升,更好地处理非线性、易陷入局部极小点等问题。最小二乘支持向量机(Least Squares Support Vector Machine, LSSVM)是在SVM的基础上进行改进,提升了求解速度及收敛准确度,大大减少了建模时间。

针对当前智慧电厂设备状态监测效率低的问题,本文研究基于大数据分析的智慧电厂设备状态监测,可有效提高收敛效率,实现智慧电厂设备状态的实时监测^[7]。

1 大数据分析技术

通过最小二乘支持向量机对发电厂设备异常状态进行分类,实现设备异常状态监测,并且利用改进蚁群算法优化分类过程,获取最佳的设备状态监测结果^[8]。

1.1 最小二乘支持向量机

LSSVM的根本理论为: $\{x_i, y_i\}_{i=1}^n$ 表示 n 个训练设备状态样本,其中,训练输入设备状态样本有 n 个维度,该设备状态样本用 $x_i \in R^n$ 描述;设备状态样本目标值用 $y_i \in R$ 描述。为创建最佳决策函数,需采用非线性映射函数 $\varphi()$ 将 $\{x_i, y_i\}_{i=1}^n$ 映射到高维特征空间,该最佳决策函数公式为

$$y(x) = \omega^T \varphi(x) + b \quad (1)$$

式中, $\omega \in R^m$ 为向量;通过 $\varphi(): R^n \rightarrow R^m$ 映射 $x_i \in R^n$ 至高维特征空间; b 为偏置量。

最小二乘支持向量机的优化问题公式为

$$\begin{cases} \min_{w,b,e} J(w,e) = \frac{1}{2} \|\omega\|^2 + \frac{1}{2} \gamma \sum_{i=1}^n e_i^2 \\ \text{submit } y_k = \omega^T \varphi(x_k) + b + e_k, i=1,2,\dots,n \end{cases} \quad (2)$$

式中, $e_k \in R$ 为误差变量; γ 为正则化参数。

对于最小二乘支持向量机的优化问题,其相应的拉格朗日函数公式为

$$\begin{aligned} L(\omega, b, e; \alpha) = & \frac{1}{2} \omega^T \omega + \frac{1}{2} \gamma \sum_{i=1}^n e_i^2 - \\ & \sum_{i=1}^n \alpha_i \{ \omega^T \varphi(x_i) + b + e_i - y_i \} \end{aligned} \quad (3)$$

式中, $\alpha_i \in R$ 为拉格朗日乘子, $i=1,2,\dots,n$ 。

计算 L 对 ω 、 b 、 e_i 、 α_i 的偏导数值为0,需优化式(3),优化后为

$$\begin{cases} \frac{\partial L}{\partial \omega} = 0 \rightarrow \omega = \sum_{i=1}^n \alpha_i \varphi(x_i) \\ \frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^n \alpha_i = 0 \\ \frac{\partial L}{\partial e_i} = 0 \rightarrow \alpha_i = \gamma e_i \\ \frac{\partial L}{\partial \alpha_i} = 0 \rightarrow \omega^T \varphi(x_i) + b + e_i - y_i = 0 \end{cases} \quad (4)$$

清除 ω 与 e_i 变量后获得矩阵方程公式为

$$\begin{pmatrix} 0 & I_n^T \\ I_n & Q + \gamma^{-1} I_n \end{pmatrix} \begin{pmatrix} b \\ \alpha \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix} \quad (5)$$

式中, I_n 为单位矩阵; $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_n]^T$ 为拉格朗日乘子; $Q = [Q_{kj}]$, $Q_{kj} = \varphi(x_k)^T \varphi(x_j)$, $k, j=1,2,\dots,n$; $y = [y_1, y_2, \dots, y_n]^T$ 。

基于Mercer要求,核函数公式为

$$k(x_k, x_j) = \varphi(x_k)^T \varphi(x_j) \quad (6)$$

为获得最小二乘支持向量机的预测模型,需算出 a 与 b 的值,该值可联合式(5)与式(6)求得,该模型为

$$y(x) = \sum_{i=1}^n a_i k(x, x_i) + b \quad (7)$$

1.2 改进蚁群算法

利用改进蚁群算法优化求解最小二乘支持向量机,提高算法的收敛速率及监测准确度^[9]。蚁群算法原理为:先出发的蚂蚁随机选择不同的出发道路,并将信息素散发在出发道路上,循环往复,导致各条道路上的信息素浓度存在差异,后出发的蚂蚁沿着高浓度信息素的道路行走^[10]。以上蚁群算法的收敛准确度与效率受到信息素浓度累积速度的影响,累积速度过快或过慢分别导致蚁群算法陷入局部最优或死循环,导致蚂蚁算法收敛准确度低,收敛效率下降,因此要对蚁群算法里的算子进行优化改进^[11]。改进蚁群算法过程如下:

1) 将蚁群规模、迭代次数等基本参数进行原始化处理,并在最小二乘支持向量机的优化模型中输入这些参数。

2) 设置原始信息素浓度^[12]。

3) 为选择最终的出发路径,每个蚂蚁将采集所有路径上的信息素浓度,依据概率进行判断。计算概率公式为

$$p_i^j(t) = \begin{cases} \frac{\rho_i^s(t) \eta_i^v(t)}{\sum_{i \in A_j} \rho_i^s(t) \eta_i^v(t)} & i \in A_j \\ 0 & i \notin A_j \end{cases} \quad (8)$$

式中, A_j 为蚂蚁选择的路径,且该蚂蚁为第 j 只蚂蚁; ρ_i 为路径上的信息素浓度,且该路径仅为第 i 个路径; η_i 为路径的期望程度; ζ 为启发算子; v 为期望算子。

4) 计算并保存不经过路径相应的目标函数值。

5) 迭代信息素的优化可根据上一步算出目标函数值,迭代公式为

$$\rho_i(k+1) = (1-\delta_0) \rho_i(k) + \delta_0 \Delta \rho_i^k \quad (9)$$

式中, $\rho_i(k+1)$ 与 $\rho_i(k)$ 分别为迭代前后的路径; $\delta_0 \in [0, 1]$ 为信息素浓度的挥发速率; $\Delta \rho_i$ 为信息素在第 i 条路径上的增量,公式为

$$\Delta \rho_i^k = \begin{cases} I/K(x) & K(x) \in R \\ 0 & K(x) \notin R \end{cases} \quad (10)$$

式中, I 为信息常用函数; $K(x)$ 为目标函数。

6) 信息素的约束条件公式为

$$\begin{cases} N_i = N_{\min} & N_i \leq N_{\min} \\ N_i = N_{\min} & N_{\min} \leq N_i \leq N_{\max} \\ N_i = N_{\max} & N_i \geq N_{\max} \end{cases} \quad (11)$$

式中, N_i 为第 i 条路径上的信息素浓度; $N_{\max}$ 与 $N_{\min}$ 分别为最大与最小信息素浓度。

7) 启发及期望算子的迭代公式分别为

$$\zeta(k) = \theta_1 \zeta(k-1) \quad (12)$$

$$v(k) = \theta_2 v(k-1) \quad (13)$$

式中,满足 $\theta_1 > 1$ 与 $\theta_2 > 1$ 。

8) 若以上算法达到了最大迭代次数,则进入下一步;反之,则返回步骤3)。

9) 输出最小二乘支持向量机的最优参数值。

2 基于大数据分析的设备状态监测

2.1 设备状态数据采集

大多数设备参数受设备状态的影响,其中设备状态变化的关键参数是振动信号。为获得有效的设备状态数据,可利用无线传感器采集设备的振动信号^[13],经预处理去除无效信号后,进行设备状态特征提取。

2.2 设备状态特征提取

设备振动信号的提取可利用小波包分解完成, $f_{ij}(t)$ 与 $x(t)$ 分别描述第 i 层分解设备状态信号与智慧电厂设备振动信号, $x(t)$ 为

$$x(t) = \sum_{j=0}^{2^L-1} f_{ij}(t_j) \quad (14)$$

式中, $f_{ij}(t)$ 为第 i 层节点(i, j)上设备状态信号的小波包分解的重构信号。

设备振动信号的能量随着设备状态的改变而

改变, 监测设备状态可依据能量谱完成。能量谱运算公式为

$$E_{ij}(t_j) = \int_T |f_{ij}(t_j)|^2 dt = \sum_{j=1}^m |x_{ij}|^2 \quad (15)$$

式中, $f_{ij}(t_j)$ 的幅值是 x_{ij} 。

设备状态振动信号的总能量公式为

$$E = \sum_{j=1}^{2^i-1} E_{ij}(t_j) \quad (16)$$

设备振动信号的小波包分解在第 i 层上的子频带能量在总能量中的占比公式为

$$P_{ij} = \frac{E_{ij}(t_j)}{E} \times 100\% \quad (17)$$

式中, P_{ij} 为设备状态的监测特征, 且支持向量机分类的输入与输出向量分别为 P_{ij} 及设备状态。

2.3 设备状态监测流程

基于大数据分析的智慧电厂设备状态监测方法的详细步骤如下。

- 1) 利用无线传感器采集的设备振动信号, 经预处理去除冗余。
- 2) 为提取设备监测的特征向量, 需算出子频信号的能量与总能量的比例, 其中, 每个子频信号的能量值可通过小波包分析处理提取之后的有效振动信号得到。
- 3) 将步骤2) 得到的设备监测特征向量输入到最小二乘支持向量机中^[14], 并输出该设备状态, 构建设备监测样本集。
- 4) 随意挑选一些样本集内的样本, 并训练这些样本, 样本集里的剩余样本视为验证集, 且训练样本数需远大于验证样本。
- 5) 构建智慧电力设备状态监测模型, 需确定惩罚参数与核宽度的数值, 该数值可利用改进蚁群算法优化设备监测训练样本参数与支持向量机器学习获得。

6) 基于步骤5) 构建的智慧电力设备状态监测模型, 利用验证样本算出该模型的监测准确率与错检率, 判断设备的监测效果^[15]。

3 试验结果与分析

为检验本文方法对发电厂设备状态的监测效

果, 以某智慧电厂为例进行发电厂设备状态监测仿真试验。

对比不同迭代次数下传统蚁群算法与改进蚁群算法的最小路径长度, 分析本文方法采用的改进蚁群算法的收敛效率与准确度, 结果如图1所示。由图1可得, 随着迭代次数的不断递增, 两种蚁群算法的最小路径长度均有所减短, 直至趋于稳定。但传统蚁群算法的最小路径长度拐点尖锐且数量密集, 在迭代180次时才达到稳定; 对比该算法, 本文方法采用的改进蚁群算法的迭代次数与路径长度均大幅度变短, 拐点平滑且数量少, 其最小路径长度仅在迭代20次时就已达到稳定。试验表明, 该方法采用的改进蚁群算法具有更高的收敛效率与准确度。

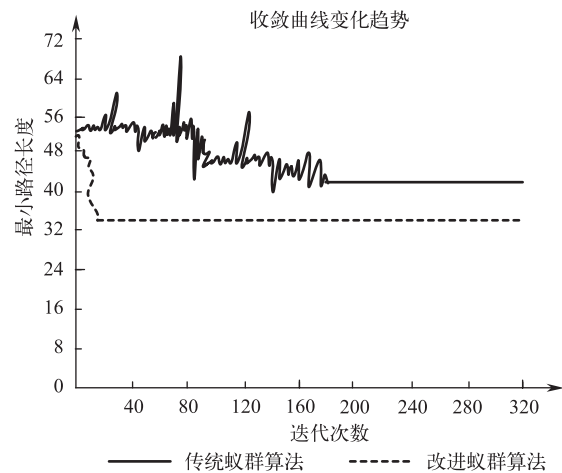


图1 对比迭代次数及路径长度

Fig. 1 Compare the number of iterations and path length

以该发电厂的发电机为例, 采用本文方法提取该设备的振动信号频谱图, 分析该发电机的运行状态, 提取到的发电机正常信号与异常信号频谱图如图2所示。由图2a可以看出, 发电机在正常运行时的频谱图整体比较稳定, 仅在1 000 ~ 1 500 Hz区域内出现较为显著的峰值信息, 最大幅值低于0.12。由图2b可看出, 频谱图在500 Hz以后发生大幅波动, 其幅值高达0.32, 为发电机异常监测提供了依据。试验表明, 本文方法可有效提取到设备不同状态下的振动信号频谱图, 为有效监测设备异常状态提供有效的数据支撑。

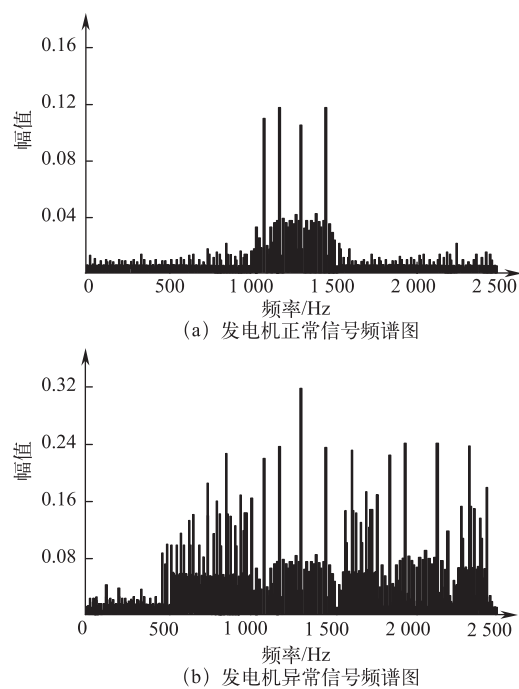


图2 发电机信号频谱图
Fig. 2 Generator signal spectrum

采用相对误差（RPE）作为评价发电机状态监测指标，相对误差公式为

$$RPE = \left| \frac{A(i) - F(i)}{A(i)} \right| \times 100\% \tag{18}$$

式中， $A(i)$ 为发电机振动实际值； $F(i)$ 为采用本文方法监测发电机振动的监测值。

采用本文方法的发电机振动监测结果相对误差如图3所示。由图3所知，采用本文方法监测发电机振动结果与其实际结果相差不大，监测值与实际值之间的相对误差全部低于3%。试验表明，采用本文方法可精准监测发电机振动情况，具有良好监测效果。

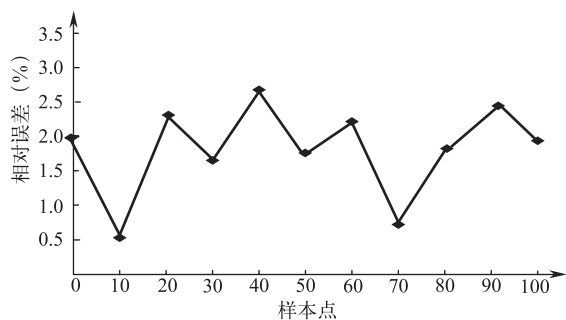


图3 监测值与实际值间的相对误差
Fig. 3 Relative error between monitored value and actual value

试验采用文献[4]基于双通道图像的智能监测方法、文献[5]全维度设备状态监测方法和本文方法，对发电厂内的发电机（编号1~6）进行状态监测，判断采用三种方法获取的发电机状态监测结果与实际状态是否一致，结果见表1。由表1可得，对其他两种方法的监测结果，采用本文方法对发电厂实际样本进行状态监测，其结果与样本实际状态结果高度一致。试验表明，采用本文方法监测发电厂设备状态，监测效果更优。

表1 实际样本状态监测结果						
Tab. 1 Actual sample condition monitoring results						
发电机编号	1	2	3	4	5	6
实际状态	低热	中热	正常	高热	局部放电	低能
文献[4]方法	正常	正常	高热	高热	正常	低能
文献[5]方法	正常	低热	中热	高热	无码	高能
本文方法	低热	中热	正常	高热	局部放电	低能

采用本文方法对该发电厂发电机、磨煤机和风机三种设备的运行状态进行10个月的监测，监测结果见表2。由表2可看出，该发电厂三种设备采用本文方法进行监测后，运行状态变化显著，可靠率、电压合格率均大幅提升，且均高达99.8%以上；客户投诉次数均不超过3次。试验表明，采用本文方法对发电厂设备进行监测，可极大地提高发电厂设备运行安全性及可靠性，提高客户满意度。

表2 发电厂不同设备状态监测分析			
Tab. 2 Condition monitoring and analysis of different equipment in power plant			
电厂设备	指标	应用前	应用后
发电机	可靠率（%）	84.1	99.8
	电压合格率（%）	89.2	99.9
	客户投诉次数 / 次	18	3
磨煤机	可靠率（%）	83.7	99.9
	电压合格率（%）	89.8	100
	客户投诉次数 / 次	26	1
风机	可靠率（%）	82.5	99.8
	电压合格率（%）	92.4	100
	客户投诉次数 / 次	15	0

4 结束语

由于发电厂设备状态多种多样,针对当前发电厂设备状态监测缺陷,本文研究基于大数据分析的智慧电厂设备状态监测方法。经试验表明,采用本文方法具有如下优点:

1) 本文方法采用的改进蚁群算法的迭代次数少且路径短,拐点平滑且数量少。

2) 采用本文方法监测发电机振动结果与其实际结果的相对误差不足3%。

3) 依据本文方法提取的振动信号频谱图可有效区分出设备异常情况。

4) 采用本文方法对发电厂实际样本进行状态监测结果与样本实际状态结果高度一致。

5) 采用本文方法监测发电机、磨煤机和风机三种设备状态,用户满意度大幅提高。

综上所述,采用本文方法监测发电厂设备状态,具有更快的收敛速度与更高的准确度,且可提高发电厂设备运行安全性及可靠性,监测效果更好,为发电厂设备的平稳运行提供保障。

参考文献

- [1] 吴天昊,刘韬,施海宁,等.基于核主元分析法的核电厂设备状态监测技术研究[J].核动力工程,2020,41(5):132-137.
- [2] 高旭,刘宏君,杜丽艳,等.基于综合评价识别法的智能变电站虚回路在线状态监测技术研究[J].电力系统保护与控制,2019,47(3):182-187.
- [3] 李刚,仇晨光,曹帅,等.基于稳健状态估计的设备状态预警研究[J].热力发电,2020,49(11):1-7.
- [4] 孙海铭,曹桐滔,代作晓,等.基于双通道图像的电力设备智能监测技术研究[J].激光与红外,2019,49(11):1338-1343.
- [5] 艾精文,党晓婧,吕启深,等.基于物联网的具有全景功能的全维度设备状态监测系统研究[J].

电力系统保护与控制,2019,47(16):122-128.

- [6] 燕凯,岳振宇,高嘉浩.基于大数据的电力计量设备运行状态评估与故障诊断方法研究[J].电子器件,2019,42(5):1095-1098,1121.
- [7] 叶康,冷喜武,肖飞,等.基于大数据标签技术的电网监控智能分析方法[J].电测与仪表,2019,56(4):75-79.
- [8] 朱海鹏,赵磊,秦昆,等.基于大数据分析的电力监控网络安全主动防护策略研究[J].电测与仪表,2020,57(21):133-139.
- [9] 刘庆连,王雪平.智能电网大数据异常状态实时监测仿真[J].计算机仿真,2019,36(3):364-367.
- [10] 任伟,徐子立,宋晓林,等.基于数据挖掘的配网数字化计量系统运行特性监测和评价方法研究[J].高压电器,2020,56(8):183-191.
- [11] 王俊波,李慧,邱太洪.基于数据挖掘的CVT电容元件击穿故障分析[J].电力电容器与无功补偿,2019,40(3):96-100.
- [12] 马莉莉,刘江平.基于数据挖掘的光纤通信网络异常数据检测研究[J].应用光学,2020,41(6):1305-1310.
- [13] 刘吉,刘炳含,张月,等.基于大数据技术的火电厂节能环保多目标负荷优化分配[J].工程热物理学报,2020,41(1):29-38.
- [14] 陈翔.基于计算机数据挖掘的农机设备状态智能检测研究[J].农机化研究,2020,42(4):18-22.
- [15] 陈江兴,梁良,付俊峰,等.基于大数据的智能电网数据调度与快速分发方法研究[J].电测与仪表,2020,57(6):88-93.

作者简介

王建设(1987—),男,工程师,主要研究方向为发电企业信息化、数字化建设及网络安全管理。

(收稿日期:2021-11-22)