

基于特征迭代优选的OPGW光缆振动事件分类方法

李博¹,许大鹏²,李浩东³,祝永坤²,尚鑫¹,丁仁杰²

(1. 国网内蒙古东部电力有限公司 电力科学研究院,呼和浩特 010020;

2. 国网内蒙古东部电力有限公司,呼和浩特 010020;3. 长春理工大学,长春 130013)

摘要:针对光纤复合架空地线(OPGW)异常振动事件识别中特征相似度高、易误分类的问题,提出一种基于特征迭代优选的识别方法。该方法在极端随机树(ET)模型基础上引入迭代机制,结合误分类权重与节点不纯度构建特征重要性动态评估策略,并融合特征信息熵改进ReliefF算法以优化特征子集权重。实验结果表明:该方法在58维初始特征基础上经迭代优选降至33维,分类准确率达89.54%,召回率、精确率与F1分数均保持在89.8%以上,相较于朴素贝叶斯(NB)、支持向量机(SVM)、K近邻算法(KNN)等模型,各项性能指标提升7%~20%。

关键词:相位敏感光时域反射仪;ReliefF算法;权重优化;特征选择;分类

中图分类号:TN256 **文献标志码:**A **文章编号:**1002-5561(2026)03-0102-05

DOI:10.13921/j.cnki.issn1002-5561.2026.03.017

Classification of OPGW cable vibration events based on iterative feature selection

LI Bo¹, XU Dapeng², LI Haodong³, ZHU Yongkun², SHANG Xin¹, DING Renjie²

(1. Electric Power Science Research Institute of State Grid Eastern Inner Mongolia Electric Power Co., Ltd., Huhehaote 010020, China;

2. State Grid Inner Mongolia East Power Co., Ltd., Huhehaote 010020, China;

3. Changchun University of Science and Technology, Changchun 130013, China)

Abstract: To address the challenges of high feature similarity and susceptibility to misclassification in abnormal vibration event identification for optical fiber composite overhead ground wire(OPGW), this paper proposes an identification method based on iterative feature optimization. Building upon the extremely randomized trees(ET)model, the method introduces an iterative mechanism that integrates misclassification weights and node impurity to construct a dynamic feature importance evaluation strategy. Furthermore, it incorporates feature information entropy to improve the ReliefF algorithm for optimizing feature subset weights. Experimental results demonstrate that, starting from 58 initial features, the method reduces the feature dimensionality to 33 through iterative optimization, achieving a classification accuracy of 89.54%, with recall, precision, and F1-score all maintained above 89.8%. Compared with baseline models including naive bayes(NB), Support vector machine(SVM), and K-nearest neighbors(KNN), the proposed method improves all performance metrics by 7%~20%.

Key words: phase-sensitive optical time-domain reflectometer, ReliefF algorithm, weight optimization, feature selection, classification

0 引言

光纤复合架空地线(OPGW)作为电力系统的关键基础设施,兼具高压输电线路接地保护与光通信数据

收稿日期:2024-09-14。

作者简介:李博(1990—),男,吉林省吉林市人,硕士,高级工程师,主要从事输电检修工作,研究方向为输电线路覆冰舞动监测与防治、雷击防护技术、人工智能等方向,参与完成省部级科技项目5项,发表论文10余篇。



传输的双重功能。随着智能电网建设的深入推进,OPGW的健康状态直接关乎电力系统的稳定运行及通信网络的可靠性。然而,由于OPGW长期暴露于野外复杂环境中,极易遭受鸟害、覆冰、舞动等自然灾害的侵扰,引发光缆异常振动,进而导致光纤损伤甚至断裂^[1]。为实现对这类潜在损害的早期预警与实时响应,基于相位敏感光时域反射仪(Φ -OTDR)的分布式振动监测技术应运而生,该技术能够实时捕捉并分析OPGW沿线的振动信号,从而实现了对光缆物理状态的精准监控。

目前,学术界主要依托特征提取技术与机器学习算法构建分类模型,以实现 Φ -OTDR 振动事件的识别。在特征提取方面,研究者提出了多种策略:Tabi FBM 等人^[2]组合时域与频域特征(TFC)进行识别;Jiang Lihui 等人^[3]利用小波分解提取信号特征;Min H K 等人^[4]采用高斯混合模型;Zhu Hui 等人^[5]选取干扰信号的交叉率作为特征指标;Jiang Fei 等人^[6]应用梅尔频率倒谱系数(MFCC);Sun Qian 等人^[7]则提出从时空数据矩阵中提取特征形态的方法。在分类模型构建上,Shi Yi 等人^[8]在卷积神经网络(CNN)架构中引入并行路径以缩短识别时间;Tian Manling 等人^[9]结合信道注意力机制、时空卷积网络与双向长短期记忆网络(Bi-LSTM)提升识别性能;Shi Yi 等人^[10]则将 CNN 与支持向量机(SVM)相结合进行分类。尽管上述方法在一定程度上提升了识别效果,但仍存在普遍局限性:一方面,现有算法的泛化能力与实时性有待提高;另一方面,针对异常事件发生早期特征微弱且相似度高的情况,传统模型易出现误分类现象,导致整体识别效率降低。

针对上述问题,本文提出一种基于特征迭代优选的光缆振动事件分类方法,旨在通过优化特征选择与分类器结构,提升对微弱及相似振动事件的识别精度与实时性。

1 基于特征迭代优选的 OPGW 振动事件分类方法

1.1 光缆振动数据特征提取

由于振动信号具有多样性与非平稳性,仅采用单一类型特征难以实现信号的准确识别。本文在朱海强等人^[11]研究基础上,从多维度提取信号特征,包括时域特征、频域特征以及经集合经验模态分解(EEMD)^[12]后各本征模态函数(IMF)分量的时频域特征,据此构建数据集。

针对 OPGW 光缆振动事件信号,本文选取的时域特征包括:最大值、最小值、峰峰值、方根幅值、均值、方差、整流平均值、均方根值、波形因子、脉冲因子、峰值因子、裕度因子、峰度和偏度,共计 14 维;频域特征包括:频谱质心、均方根频率、频率方差、频率平坦度、频谱滚降点和频谱超额峭度,共计 6 维;时频域特征方面,采用 EEMD 算法分解信号后,提取各 IMF 分量的能量熵、排列熵和模糊熵作为时频特征^[13]。

综合上述 3 类特征,本文构建了一个包含 58 维特征的数据集,作为后续基于迭代特征优选的 OPGW 光缆振动事件分类算法的输入。

1.2 基于误分类调节的特征重要性动态计算

为了实现对相似振动信号的精确分类,本文提出了一种基于误分类调节的特征重要性动态计算方法。该方法结合极端随机树(ET)模型的不纯度减少量与误分类惩罚项,计算特征重要性 $I(f_i)$,其计算公式为

$$I(f_i) = \sum_{t=1}^T G(m, f_i) - \lambda M(f_i) \quad (1)$$

$$G(m, f_i) = I_p - (p_l I_l + p_r I_r) \quad (2)$$

$$\lambda = 1 + \frac{\sum_{i=1}^n \Delta P(f_i)}{n} \quad (3)$$

$$M(f_i) = w(f_i) \sum_{j=1}^N R_j(f_i) \quad (4)$$

式中: f_i 为第 i 个特征, T 为所有决策树的数量, $G(m, f_i)$ 为第 m 棵树中特征 f_i 带来的不纯度减少量, I_p 、 I_l 和 I_r 分别为父节点、左子节点和右子节点的不纯度, p_l 和 p_r 分别为左右子节点的样本比例, λ 为针对误分类贡献度的惩罚因子。 $\Delta P(f_i)$ 表示特征对于模型性能的影响值,由 ET 模型计算得到,通过对所有特征的平均性能影响进行归一化处理,实现惩罚因子的动态调整。 $M(f_i)$ 为第 i 个特征的误分类贡献度, N 为样本总数。 $w(f_i)$ 为特征权重,初始值设为 1,在后续的迭代中更新为改进的 ReliefF 算法^[14]计算得到的权重值, $R_j(f_i)$ 是一个指示函数,若第 j 个样本在使用特征 f_i 参与分类时发生误分类,则取值为 1,否则为 0。

基于所计算得到的特征重要性 $I(f_i)$,本文设置了筛选阈值 $\theta = 0.9$,即在每次特征迭代循环中,仅保留综合重要性评分排名前 90% 的特征,以便提高分类效率和准确性。

1.3 基于特征重要性评估的特征选择

在获得初步特征重要性 $I(f_i)$ 并对特征进行筛选后,本文将采用改进的 ReliefF 算法对保留特征的权重 $w(f_i)$ 进行调整,增强模型对关键预特征的敏感度,减轻在数据集中分布广泛且可能引起分类决策不稳定的特征的权重。同时,引入每次迭代计算得到的特征重要性作为调节因子,可以在保持特征区分力的同时,减少由于特征不确定性带来的过拟合风险,使得模型更加鲁棒。其计算公式为

$$w'(f_i) = w(f_i) - \sum_{j=1}^k \frac{d_{\text{same}}}{tk_{\text{same}}} \frac{I(f_i)}{H(f_i)} + \sum_{l \notin l_{x_i}} \sum_{j=1}^k \frac{p(l)}{1 - p(l_{x_i})} \frac{d_{\text{diff}}}{tk_{\text{diff}}} \frac{I(f_i)}{H(f_i)} \quad (5)$$

式中: d_{same} 和 d_{diff} 分别为相同类别和不同类别在特征 f_i 上的差异, t 为迭代次数, k_{same} 和 k_{diff} 分别为同类和不同

类的样本数量, $p(l)$ 为标签或类别 l 的先验概率, $p(l_{x_i})$ 为给定样本 x_i 后的条件概率, $H(f_i)$ 为特征在同类样本中的样本熵, 其计算公式为

$$H(i) = - \sum_j p(f_{ij}) \log(p(f_{ij})) \quad (6)$$

式中: $p(f_{ij})$ 为第 i 个特征取 j 值的概率。

1.4 基于特征迭代优选的光缆振动事件分类

针对由 OPGW 光缆振动信号中提取到的复杂高维振动数据, 本文将改进的特征优选算法与 ET 模型相结合, 提出了一种基于特征迭代优化的光缆振动事件分类方法。该方法通过迭代地评估和选择最具影响力的特征, 旨在保留最关键信息的同时降低数据维度, 从而提高分类模型的准确性和计算效率; 还采用了一种动态的特征权重分配策略, 以进一步增强模型对复杂振动模式的识别能力, 确保了在面对相似度较高的振动事件时能够实现高精度的分类。分类原理图如图 1 所示: 1) 对采集到的震动数据进行预处理, 从时域、频域和时频域 3 个维度进行特征提取, 为后续分类构建数据集。2) 将数据集划分为测试集和训练集, 初始化特征权重, 并将训练集输入至 ET 模型中进行初次训练。3) 计算节点不纯度和误分类贡献度, 并使用集成误分类调节的特征重要性计算方法计算特征重要性, 根据预设阈值筛选特征。4) 计算样本熵并使用基于特征重要性的改进 ReliefF 算法对剩余特征进行权重分配。5) 将调整后的特征输入 ET 模型进行下一轮迭代。若特征集的改变显著影响分类性能且未达迭代上限, 则返回继续迭代, 否则退出循环并输出最终 ET 模型。6) 将测试集输入模型中, 输出分类结果并

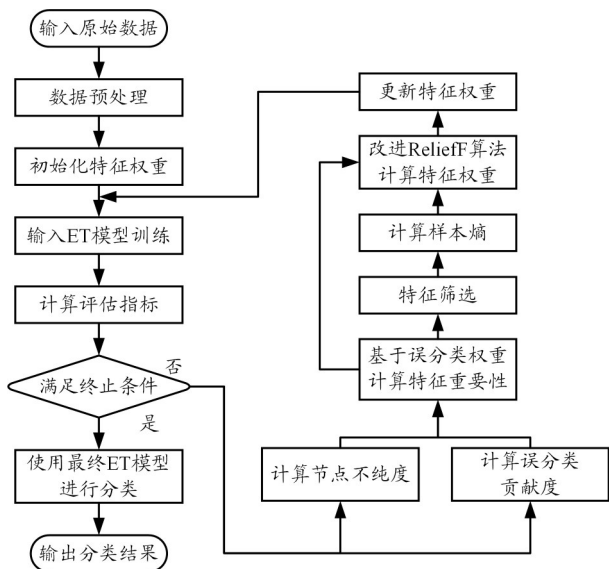


图1 基于特征迭代优选的光缆振动事件分类原理图

计算各类指标。

本文提出的模型不仅在保留最重要振动信号特征的同时, 降低了数据的复杂度, 并引入针对误分类的动态特征权重分配, 显著增强了模型对复杂振动模式的识别能力。此外, 通过多轮特征选择和权重更新, 能够有效地识别出关键的特征子集, 实现对振动事件的准确分类。

2 实验及结果分析

2.1 实验设置

本实验采用一根长 20 m、24 芯层绞式结构的 OPGW 光缆作为测试对象, 两端各接入 1 000 m 光纤以模拟长距离传输环境。通过人为模拟野外运行中常见的 4 类异常振动事件, 并利用 Φ -OTDR 系统采集振动信号。 Φ -OTDR 系统的主要参数如表 1 所示。实验现场光缆架设及设备连接如图 2 所示。

表1 Φ -OTDR系统的主要参数表

参数名称	参数值
探测脉冲宽度/ns	32
脉冲重复频率/Hz	1 000
中心波长/nm	1 550
探测总距离/m	2 000
空间分辨率/m	5
DAQ采样率/Hz	5 000

注: DAQ表示数据采集。



(a) 光缆架设现场图 (b) Φ -OTDR设备现场图

图2 实验现场图

本实验模拟了风舞、覆冰舞动、鸟害、异物覆盖 4 类 OPGW 光缆常见的异常振动事件。

由于每种类型的事件收集时间长度不同, 因此对数据进行裁剪以实现格式统一。裁剪后的样本数据由 2 000 个点组成, 得到事件数据中的样本数量如表 2 所示。

表2 4类样本数量

事件类型	样本数量/组	标签
风舞	279	label1
覆冰舞动	293	label2
鸟害	327	label3
异物覆盖	278	label4

在模型训练与评估策略上, 本文将数据集按 7:3 的比例随机划分为训练集和测试集。其中, 70% 的训练集用于模型的迭代训练、特征优选及权重调整, 以获取最优特征子集; 30% 的测试集用于最终评估分类器对未知数据的识别能力。此外, 为进一步消除单一数据划分带来的偶然性偏差, 确保评估结果的稳健性, 本实验采用 5 折交叉验证方法对模型性能进行综合评估。

2.2 实验结果分析

为全面量化评估分类效果, 本文选取准确率、召回率、精确率和 F1 分数作为主要评价指标。准确率反映模型整体分类正确的比例; 召回率和精确率分别从漏报和误报角度衡量模型对特定类别的识别能力; F1 分数则是两者的调和平均数, 用于综合平衡模型性能。通过上述多维指标的分析, 能够客观验证改进算法在 OPGW 振动事件分类任务中的有效性与鲁棒性。

2.2.1 对比实验结果分析

为评估本文提出方法的性能, 选取朴素贝叶斯(NB)、SVM、K 近邻算法(KNN)、决策树(DT)和随机森林(RF)5 种经典机器学习算法作为基线模型进行对比。所有模型均在同一数据集上进行验证, 实验结果如表 3 所示。可以看出, 本文所提方法在准确率、召回率、精确率和 F1 分数 4 项指标上均表现最优, 相比其他模型提升了 7%~20%。

表 3 对比试验结果

方法	准确率	召回率	精确率	F1 分数
NB	0.709 0	0.708 9	0.701 0	0.685 9
SVM	0.822 0	0.825 4	0.803 5	0.824 3
KNN	0.748 5	0.746 6	0.752 7	0.745 9
DT	0.740 1	0.743 6	0.742 8	0.742 7
RF	0.819 2	0.821 6	0.819 3	0.809 4
本文方法	0.895 4	0.899 6	0.899 1	0.898 9

表 4 迭代实验结果

迭代次数/次	特征维度/维	准确率	召回率	精确率	F1 分数
0	58	0.827 7	0.831 5	0.827 7	0.808 8
1	50	0.864 4	0.869 7	0.869 3	0.866 6
2	45	0.875 7	0.880 6	0.880 4	0.878 5
3	40	0.884 1	0.889 1	0.887 5	0.887 9
4	37	0.892 6	0.897 4	0.895 0	0.895 4
5	33	0.895 4	0.899 6	0.899 1	0.898 9

注: 初始模型的迭代次数为 0 次。

2.2.2 迭代实验结果分析

本文以初始 ET 模型为基准, 记录每次迭代后的准确率、召回率、精确率和 F1 分数, 并与初始结果进行对比。实验结果如表 4 所示。可以看出, 经迭代优选后特征维度从 58 维降至 33 维, 分类准确率最高达 89.54%, 召回率、精确率与 F1 分数均保持在 89.8% 以上。

3 结束语

本文针对 OPGW 光缆异常振动事件的震动数据相似度较高、易误分类等问题, 提出了一种基于特征迭代优选的光缆振动事件识别方法。该方法在 ET 模型基础上引入迭代优化机制, 构建了融合误分类权重与节点不纯度的动态特征重要性评估策略, 并结合特征信息熵对传统 ReliefF 算法进行改进, 实现了特征维度的自适应筛选与权重动态分配。实验结果表明, 该方法在 58 维初始特征基础上经迭代优选降至 33 维, 分类准确率最高达 89.54%, 召回率、精确率与 F1 分数均保持在 89.8% 以上, 较 NB、SVM、KNN、DT 及 RF 等 5 种模型提升 7%~20%, 有效克服了相似振动事件区分度低的难题。本文为 OPGW 光缆振动监测提供了高效、鲁棒的特征处理与识别方案, 对提升智能电网输电线路状态感知的实时性与可靠性具有明确的工程应用价值。

然而, 本文所提方法在迭代优选过程中仍存在特征评估与权重更新耗时较长的问题, 一定程度上影响了模型的整体训练效率。后续研究将致力于优化迭代过程的计算效率。

参考文献:

- [1] Liu Xiao, Bao Zhen, Luo Hong, et al. Vibration monitoring of optical fiber composited in overhead transmission line using Φ -OTDR[C]//IOP-Publishing. Proceedings of Journal of Physics. IOP Publishing. Guangzhou: IOP-Publishing, 2023: 012051.
- [2] Tabi F B M, Han Dezhi, An Bowen, et al. Research and software design of an Φ -OTDR-based optical fiber vibration recognition algorithm[J]. Journal of Electrical and Computer Engineering, 2020, 2020: 1-13.
- [3] Jiang Lihui, Liu Xiangming, Zhang Feng. Multi-target recognition used in airpoty fiber fence warning system[C]//IEEE. Proceedings of 2010 International Conference on Machine Learning and Cybernetics. Qingdao: IEEE, 2010: 1126-1129.
- [4] Min H K, Lee C Y, Lee J S, et al. Abnormal signal detection in gas pipes using neural networks[C]//IEEE. Proceedings of IECON 2007-33rd Annual Conference of the IEEE Industrial Electronics Society. Taipei: IEEE, 2007: 2503-2508.
- [5] Zhu Hui, Pan Chao, Sun Xiaohan. Vibration pattern recognition and

李博, 许大鹏, 李浩东, 等: 基于特征迭代优选的OPGW光缆振动事件分类方法

classification in OTDR based distributed optical-fiber vibration sensing system[C]//SPIE. Proceedings of Smart Sensor Phenomena, Technology, Networks, and Systems Integration 2014. San Diego: SPIE, 2014: 29–34.

[6] Jiang Fei, Li Honglang, Zhang Zhenhai, et al. An event recognition method for fiber distributed acoustic sensing systems based on the combination of MFCC and CNN[C]//SPIE. Proceedings of 2017 International Conference on Optical Instruments and Technology: Advanced Optical Sensors and Applications. Beijing: SPIE, 2018: 15–21.

[7] Sun Qian, Hao Feng, Yan Xueying, et al. Recognition of a phase-sensitivity OTDR sensing system based on morphologic feature extraction [J]. Sensors, 2015, 15(7): 15179–15197.

[8] Shi Yi, Wang Yu, Zhao Liang, et al. An event recognition method for Φ -OTDR sensing system based on deep learning[J]. Sensors, 2019, 19(15): 3421.

[9] Tian Manling, Dong Hui, Yu Kuanglu. Attention based Temporal convolutional network for Φ -OTDR event classification[C]//IEEE. Proceedings of 2021 19th International Conference on Optical Communications and Networks(ICOON). Qufu: IEEE, 2021: 1–3.

ings of 2021 19th International Conference on Optical Communications and Networks(ICOON). Qufu: IEEE, 2021: 1–3.

[10] Shi Yi, Wang Yuanye, Wang Liyuan, et al. Multi-event classification for Φ -OTDR distributed optical fiber sensing system using deep learning and support vector machine[J]. Optik, 2020, 221: 165373.

[11] 朱海强, 张志利, 高慧敏, 等. 基于两级SVM的 Φ -OTDR系统振动事件识别算法 [J]. 电子器件, 2023, 46(3): 783–789.

[12] Ali J B, Fnaiech N, Saidi L, et al. Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals[J]. Applied Acoustics, 2015, 89: 16–27.

[13] Zou Shuli, Qiu Taorong, Huang Peifan, et al. Constructing multi-scale entropy based on the empirical mode decomposition (EMD) and its application in recognizing driving fatigue[J]. Journal of Neuroscience Methods, 2020, 341: 108691.

[14] Geurts P, Ernst D, Wehenkel L. Extremely randomized trees[J]. Machine Learning, 2006, 63: 3–42.