

基于改进决策树算法的低压配电电缆故障测距研究

戈贞玲¹ 张丽¹ 张闯¹

GE Zhenling ZHANG Li ZHANG Chuang

摘要

在低压配电系统中, 电缆内部的绝缘气隙、突起等局部缺陷会造成电缆电气性能的变异, 具体表现为频变效应。这种频变效应不仅增加了击穿电弧的发生概率, 还严重干扰了对异常信号或孤立点的准确判断。传统测距方法在这种情况下难以有效评估电缆的故障状态, 甚至可能导致误判, 进而产生较大的测距误差。为此, 文章提出了一种基于改进决策树算法的低压配电电缆故障测距新方法。该方法采集了低压配电电缆的运行环境、绝缘状态及电气特性等多维度数据。为了处理数据中的孤立点和异常值问题, 引入了k-中心点聚类算法进行孤立点检测。通过聚类分析, 能够精准识别并剔除与正常数据存在明显偏差的孤立点, 以提升数据集的准确性和可靠性。接着, 利用改进后的决策树算法, 对低压配电电缆的故障状态进行评估。在电缆故障测距环节, 结合故障状态评估的结果, 精确计算故障位置, 实现故障测距。通过对比实验证明, 该方法相比现有测距方法准确性和稳定性更高, 为低压配电电缆故障的快速定位提供了有力支持。

关键词

改进决策树算法; 低压配电网; 电缆故障测距; 故障状态评估; 故障测距

doi: 10.3969/j.issn.1672-9528.2025.02.042

0 引言

随着电力网日新月异的发展态势, 低压配电电缆在电力传输与分配领域中已然成为举足轻重的关键角色, 其重要性愈发凸显, 无可替代。因电缆运行环境复杂、使用年限增长以及负荷不断增加, 电缆故障频发, 给电力系统和用户带来了巨大的经济损失和安全风险。因此, 对低压配电电缆故障测距的研究尤为重要。

低压配电电缆通常铺设在地下或电缆沟内, 一旦发生故障, 快速准确地定位故障点成为一项极具挑战性的任务。传统的故障测距方法, 如电桥法, 虽然在一定程度上能够解决部分问题, 但其测量精度和适用范围有限, 难以满足现代电力系统对故障测距快速、准确的要求。因此, 研究和开发新的电缆故障测距技术, 对于提高电力系统的可靠性和安全性具有重要意义。

近年来, 国内学者在电缆故障测距领域进行了大量研究, 提出了多种测距方法。周国威等人^[1]提出了一种基于变分模态分解(VMD)和改进双端行波法的电缆故障测距方法。该方法利用改进的双端行波法实现故障测距。在该方法中, 双端行波法依靠准确识别故障点反射波确定故障位置。但频变效应使击穿电弧发生概率增加, 电弧会干扰行波信号。行波

在电缆中传播时, 电弧可能改变其波形、幅值和传播速度等特性, 识别故障点反射波时易误判, 从而导致测距误差较大。夏向阳等人^[2]提出了一种基于S变换模能量分析的直流电缆故障测距方法。该方法利用S变换对故障信号进行时频分析, 通过计算模能量来确定故障位置。该方法用于低压配电电缆故障测距时, 频变效应增加击穿电弧发生概率并干扰异常信号判断。进行时频分析时, 电弧干扰信号与频变效应产生的异常信号混入故障信号, 使S变换所得时频图复杂且不准确, 测距效果有待进一步验证。

鉴于上述研究方法的不足, 本文提出了一种基于改进决策树算法的低压配电电缆故障测距方法。

1 低压配电电缆数据采集与聚类孤立点检测

低压配电电缆相关数据包含运行环境数据、电缆绝缘状态数据以及电缆电气特性数据等诸多方面, 这些数据对于反映电缆运行状态及潜在故障至关重要。在实际的数据采集环节, 受测量误差、输入错误及数据固有变异等因素影响, 数据集中可能混入孤立点^[3]。这些孤立点与正常数据存在显著偏差, 若不进行处理, 将严重干扰后续的测距工作, 甚至导致误判。为应对这一问题, 本文引入k-中心点聚类算法进行孤立点检测。该算法通过对数据集进行聚类分析, 精准识别并剔除与正常数据存在明显偏差的孤立点, 从而提升数据集的准确性和可靠性^[4]。

1. 郑州科技学院信息工程学院 河南郑州 450000

假设采集的数据集为 D ，包含 n 个数据点，每个数据点为 d 维向量。需要找到 k 个中心点 $C = \{c_1, c_2, \dots, c_k\}$ ，使得目标函数 J 最小，目标函数 J 定义为所有数据点到其最近中心点的距离之和：

$$J = \sum_{i=1}^n \sum_{j=1}^k \|d_i - c_j\|^2 \quad (1)$$

式中： d_i 表示数据集 D 中的第 i 个数据点； c_j 表示第 j 个中心点。

在孤立点检测中，设定一个距离阈值 T 。对于每个数据点 d_i ，计算其到其所属聚类中心点 c_j 的距离 $d(d_i, c_j)$ 。如果 $d(d_i, c_j) > T$ ，则认为 d_i 为孤立点。图 1 给出了预处理前后某一相区外故障的电压波形变化情况。

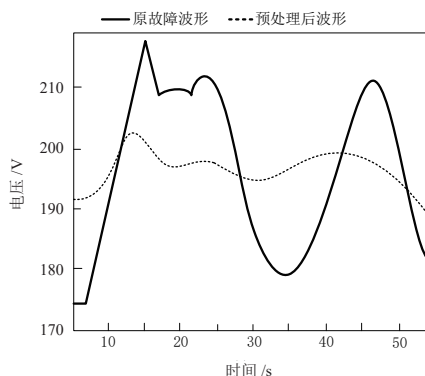


图 1 预处理前后故障电压波形变化情况

分析图 1 可以看出，在未进行预处理之前，由于谐波的影响，数据中存在明显的异常波动，这些波动点可能就是孤立点或者与正常数据偏差较大的数据点。而经过预处理后，这些异常波动得到了有效抑制，数据变得更加平滑，原本疑似孤立点的异常数据不再明显。这表明，通过本文通过采用 k -中心点聚类算法进行孤立点检测，确实能够精准地识别出与正常数据存在明显偏差的孤立点，并将其从数据集中剔除。经过这样的处理后，数据集的准确性和可靠性得到了极大提升，为后续故障测距及状态评估提供了有力支持。

2 基于改进决策树算法的低压配电电缆故障状态评估

在低压配电电缆故障测距的过程中，对电缆故障状态进行评估，能够精确判定故障的类型（像是短路、断路或者绝缘老化等情况）以及程度（例如轻微损伤或者严重破损等），有助于后续实现故障测距^[5]。经过孤立点检测后的数据中，和正常数据偏差较大的孤立点已被去除，数据变得更加纯净，这样就减少了异常数据对故障状态评估的干扰。基于此，构建了基于改进决策树算法的故障状态评估模型，并通过训练数据集确保其准确性和可靠性。

在完成低压配电电缆数据的处理之后，引入改进后的决策树算法，用于对低压配电电缆的故障状态进行评估。在构

建分别基于日常检修数据和在线监测数据的子树时，确定测试属性是一个核心挑战^[6]。若按照最小化信息熵的原则来选择测试属性，具体做法公式为：

$$T_z = \min \sum_{p=1}^v \frac{s_{1p} + \dots + s_{mp}}{s} I(s_{1p}, \dots, s_{mp}) \quad (2)$$

式中： T_z 表示子树生成结果； m 表示分类数量； s 表示数据样本总数 p 表示被选择属性可能输出值。将整个样本划分为 v 个子集，则 s_{pq} 表示子集 p 中分类 q 的样本数，此种划分的信息量可表示为：

$$I = - \sum_{q=1}^m g_{pq} \log_2(g_{pq}) \quad (3)$$

式中： g_{pq} 表示集合中样本属于类 q 的概率。当 s 一定时，选择不同的属性会导致 s_{pq} 的值不同。将每个候选属性分别按照上述公式计算，并选择使 s_{pq} 最小的属性作为测试属性。

在合并基于日常检修数据和在线监测数据的两个子树时，前提条件是它们必须拥有相同或相似的属性。若缺乏这样的直接属性，两棵子树将保持独立。为了实现子树的融合，需要寻找两者之间的相似属性。在此基础上实现在线监测数据中测得的时间属性向服役年限的合理转化^[7]，具体操作可用公式表示为：

$$F_n = C_s - T_s \quad (4)$$

式中： F_n 表示服役年限； C_s 表示测量时间； T_s 表示投入使用年限。在合并其他子树的过程中，若未能发现显而易见的相似属性，本文借助关联分析手段，从中挑选出关联性最强的属性作为连接点。面对庞大的数据样本集，首先对这些样本进行初步的分类处理，构建出多个子树，然后再根据找到的相似或关联属性将这些子树合并。

根据上述内容，完成了对故障决策树的构建，并针对其中标识为故障的结果进行后续的测距工作。这一流程不仅提高了决策树的分类准确率和泛化能力，还为低压配电电缆的故障测距提供了更为可靠的基础^[8]。

3 电缆故障测距

在基于决策树评估电缆故障时，当电缆的参数已知且电缆未发生故障时，电缆上任意一点的计算电压和理论电压是相等的。然而一旦电缆发生故障，故障点将会改变电缆的模型结构，这样从电缆各端测量得到的电压就会出现差异，进而导致计算所得的电压值存在明显误差。因此，在故障条件下，如果某个搜索点计算结果的误差很显著，那么这个点就不是故障点；反之，某个搜索点的计算结果为零，那么这个搜索点就是故障点。基于这样的原理，本文提出通过求解方程来找到能使电压差为零的未知数，进而实现电缆故障测距，从而准确定位故障点。

针对上述通过决策树评估存在故障的电缆，在电缆各项参数已知的情况下，根据电缆两端的电压和电流，得出公式为：

$$u = f_1(\lambda) \quad (5)$$

$$u = f_2(1-\lambda) \quad (6)$$

式中： u 表示电缆上任意一点的电压； λ 表示待求解的未知量。在电缆未发生故障的正常情况下，电缆中的任意一点根据 $f_1(\lambda)$ 计算的电压值与 $f_2(1-\lambda)$ 计算的电压值理论上是相等的^[9]。电缆故障后，假设故障点为 f 。在排除故障点 f 之外的任意位置 b 进行搜索时，由于电缆模型结构发生了改变，从各端测得电压不同。这种情况下，计算得到的 $f(\lambda, 1-\lambda)$ 值会有较大误差。因此，在故障已触发的情况下，若 $f(\lambda, 1-\lambda)$ 计算结果存在显著误差，则当前搜索点非故障点。反之，若搜寻点刚好是故障点，则因线路两边的线缆模型不发生变化，所以两端所求出的电压是相同的^[10]。此时， $f(\lambda, 1-\lambda)$ 理论上应等于零。因此，在故障条件下，若 $f(\lambda, 1-\lambda)$ 计算结果为零，则当前搜索点即为故障点。电缆发生故障后， $f(\lambda, 1-\lambda)$ 中通常仅含有一个未知数 λ 。通过求解该方程找到使 $f(\lambda, 1-\lambda) = 0$ 的 λ 值，即可实现电缆故障测距，定位故障点。根据以上论述，得到电缆故障的具体位置，完成测距。

4 对比实验

4.1 实验对象

为了验证本文提出的基于改进决策树算法的测距方法在实际应用中的可行性，以 10 kV 电缆线路为实例，进行了故障定位模拟实验。为确保实验结果的可对比性，设置了两组对照实验：对照 A 组采用基于 VMD 的测距方法，对照 B 组则采用基于 S 变换模能量分析的测距方法，而实验组则应用本文上述提出的测距方法。在相同的实验条件下，利用这 3 种方法对 10 kV 电缆线路进行了故障测距。图 2 为该 10 kV 电缆线路仿真模型。

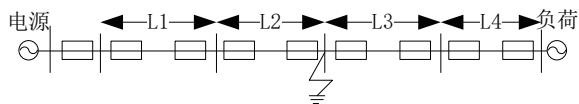


图 2 10 kV 电缆线路仿真模型

图 2 中，电源配置为 10 kV 的理想电压源，其初始相位角设定在 0°，用于模拟无限大母线的情况。实验中使用两个独立的电缆 L1 和 L2，分别代表在某一位置发生故障的电缆的两部分，通过调整这两段电缆的长度（保持总长度不变）来模拟电缆上不同位置的故障。

在内部故障场景下，设定了一个 100 Ω 的接地电阻，并引入一个 RLC 电路来模拟负载条件。为便于描述故障定位的具体区段，整条电缆线路被划分为 4 个部分：L1、L2、L3、L4。

4.2 实验结果

4.2.1 电缆故障状态判定准确性对比

在电缆故障测距之前，需要采用不同方法对电缆是否存在故障进行预判。为了验证这 3 种方法的准确性，将 3 种方法判定为电缆绝缘良好和绝缘不良的结果与实际状况进行了对比，结果如图 3~6 所示。通过对比图 3~6 可以清晰地看到，实验组的方法在判定电缆故障状态上与实际结果更为吻合，仅有极少数电缆的绝缘状态分布位置与实际存在些许偏差。而对照 A 组和对照 B 组的判定结果与实际情况存在较大差异。

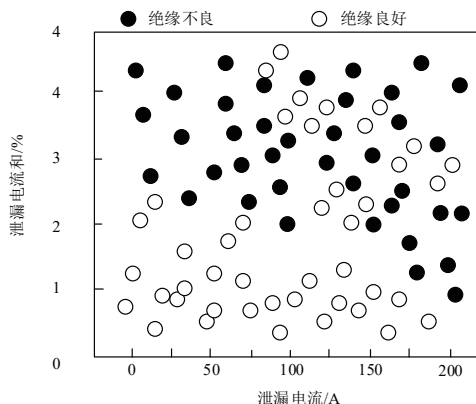


图 3 低压配电电缆实际绝缘状态分布图

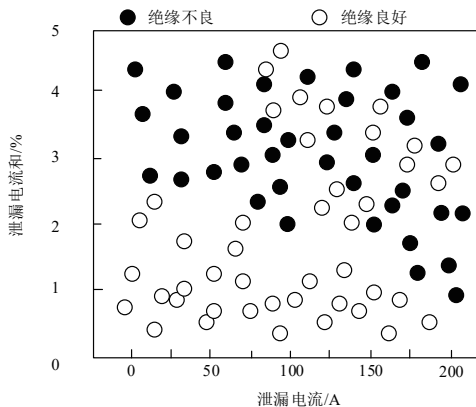


图 4 实验组电缆故障状态判定结果

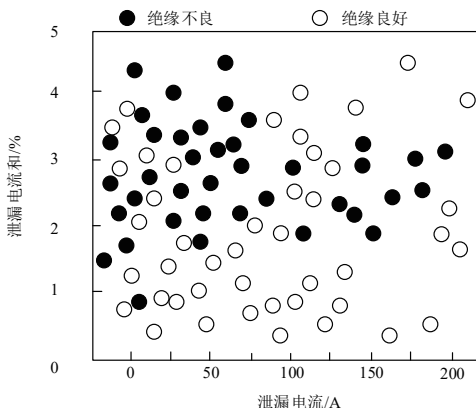


图 5 对照 A 组电缆故障状态判定结果

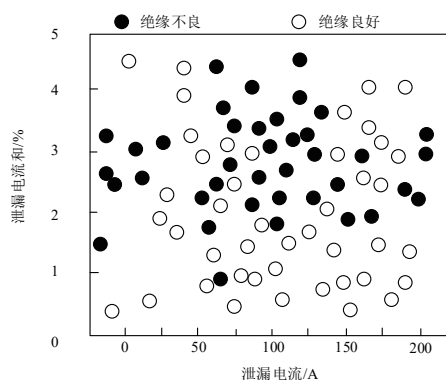


图6 对照B组电缆故障状态判定结果

通过图3~6对比,证明了实验组的方法能够准确判定低压配电电缆的故障状态,从而为后续的故障测距提供了可靠的前提和依据。通过引入k-中心点聚类算法,本文方法能够识别并处理数据中的孤立点和异常值,从而提升了数据集的准确性和可靠性。这种数据处理能力的提升,使得后续利用改进后的决策树算法对低压配电电缆的故障状态进行评估时,能够基于更为准确的数据进行,进而提高了故障状态判定的准确性。

4.2.2 测距精度对比

将电缆两相接地故障测距误差作为评价指标,对比3种测距方法的精度和测量准确性,实验结果如图7所示。

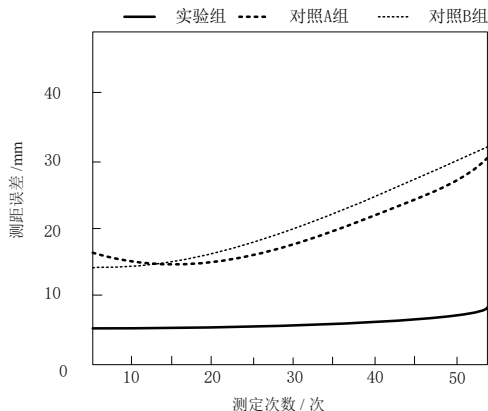


图7 三种方法电缆两相接地故障测距误差对比

分析图7可知,运用3种不同的测距方法时,针对电缆两相接地故障,实验组方法的测距误差始终处于最低水平。相较而言,在相同故障条件下,对照A组和对照B组的测距误差明显更大。结果表明,实验组所提出的距离测量方法在精度与可靠性等方面有较大提升。这意味着,实验组方法借助孤立点处理提高了数据集的准确性与可靠性,进而提升了整个测距方法的精度。

5 结语

本文提出了一种创新的低压配电电缆故障测距方法,该方法融合了改进后的决策树算法与k-中心点聚类算法。通

过多维度数据采集与孤立点检测,显著提升了数据集的准确性和可靠性。进而利用优化的决策树算法对电缆故障状态进行精确评估,并在此基础上实现了故障位置的快速定位。该方法相比传统测距方法展现出更高的准确性和稳定性,为低压配电电缆故障的快速诊断提供了有效手段。展望未来,将进一步探索其在复杂电网中的应用潜力,并与深度学习、大数据分析等先进技术结合,不断提升线缆故障定位的精度与效率。

参考文献:

- [1] 周国威,宋凯耀,刘海霞,等.基于VMD和改进双端行波法的电缆故障测距研究[J].机械与电子,2024,42(8):7-14.
- [2] 夏向阳,刘奕珏,刘雄,等.基于S变换模能量分析的直流电缆故障测距[J].高压电器,2024,60(7):201-209.
- [3] 杨明嘉,夏成军,池梓斌,等.基于沿线电流故障分量差值的交叉互联电缆故障测距[J].电力系统保护与控制,2023,51(19):54-66.
- [4] 武阳阳.基于分布式故障测距思路的管网电缆架空线路技术研究[J].电器工业,2023(8):17-23.
- [5] 杨明嘉,夏成军,赖胜杰,等.基于线芯-护层过渡电阻无功特性的交叉互联电缆故障测距[J].电工技术学报,2024,39(5):1372-1389.
- [6] 谈灏,魏星,姜静飞.输电缆故障测距发展综述[J].电气技术与经济,2023(4):214-216.
- [7] 赵昌鹏,王晓岩,庞丹,等.66 kV 电缆线路电桥故障测距法的应用[J].集成电路应用,2022,39(12):274-275.
- [8] 翁亚伟.基于脉冲电流法电缆故障测距仿真实验分析[J].电气传动自动化,2022,44(6):19-22.
- [9] 张俊,夏能弘,王节,等.考虑金属护层结构的电缆新型故障测距方法研究[J].电力科学与技术学报,2022,37(5):124-132.
- [10] 陶宇航,张熹,宫祥龙.10 kV 电缆故障测距及定位典型案例[J].电气技术,2022,23(2):88-93.

【作者简介】

戈贞玲(1997—),女,河北张家口人,硕士,助教,研究方向:教育信息化、分布式计算。

张丽(1986—),女,河南平顶山人,硕士,讲师,研究方向:物联网技术。

张闯(1982—),男,河南南阳人,硕士,讲师,研究方向:物联网、图像处理。

(收稿日期:2024-10-28)