

DOI:10.16356/j.1005-2615.2024.05.012

基于 K-means 聚类 and 随机森林的电缆风险评估及 修复决策

杨帆¹, 王红斌¹, 方健¹, 何嘉兴¹, 黄柏¹, 王莉²

(1. 南方电网广东广州供电局电力试验研究院, 广州 510000; 2. 南京航空航天大学自动化学院, 南京 211106)

摘要: 交联聚乙烯电缆是 10 kV 配电系统中的重要设备, 其安全性至关重要。对电缆的修复决策做出科学判断, 有助于提高配电系统的安全性并降低经济成本。鉴于此, 本文提出了一种基于 K-means 聚类和随机森林 (Random forest, RF) 分类模型的电缆风险评估及修复决策方法。该方法首先根据电缆的绝缘状态, 定义电缆的风险等级和风险程度; 然后利用 K-means 聚类算法对多个老化指标进行聚类以实现风险等级区间的划分, 从而建立多老化指标风险矩阵; 基于多老化指标风险矩阵, 利用综合权重法确定多维老化指标所对应的分类标签; 最后基于 RF 算法建立并训练电缆的修复决策分类模型, 输出电缆的修复决策结果。所提方法的平均正确率达到 99.70%, 实现了电缆快速且可靠的修复决策。

关键词: 老化指标; 风险矩阵; 电缆; 随机森林; K-means 聚类

中图分类号: TM75

文献标志码: A

文章编号: 1005-2615(2024)05-0892-08

Cable Risk Assessment and Repair Decision Based on K-means Clustering and Random Forest

YANG Fan¹, WANG Hongbin¹, FANG Jian¹, HE Jiaying¹, HUANG Bai¹, WANG Li²

(1. Power Test and Research Institute of China Southern Power Grid Guangdong Guangzhou Power Supply Bureau, Guangzhou 510000, China; 2. College of Automation Engineering, Nanjing University of Aeronautics & Astronautics, Nanjing 211106, China)

Abstract: Crosslinked polyethylene cable is an important equipment in 10 kV distribution system, and its safety is very important. Making scientific judgments on cable repair decisions can help improve the safety and reduce the economic cost. In view of this, a cable risk assessment and repair decision method based on K-means clustering and random forest (RF) classification model is proposed. The method first defines the risk level and risk degree of the cable based on the insulation status of the cable. Then the K-means clustering algorithm is used to cluster the multi-aging index dataset and classify the risk level intervals to build a multi-aging index risk matrix. Based on the risk matrix of the multi-aging index, the classification labels corresponding to the multi-aging index are determined by using the comprehensive weight method. Finally, the classification model of the repair ways of the cables is established and trained based on the RF algorithm, and the selection results of the repair ways are output. The average accuracy of the proposed method reaches 99.70%, achieving rapid and reliable repair decisions for cables.

Key words: aging index; risk matrix; cable; random forest; K-means clustering

基金项目: 广州供电局配网类科技项目 (GZHKJXM20200028)。

收稿日期: 2023-04-27; 修订日期: 2024-01-08

通信作者: 杨帆, 男, 工程师, E-mail: yangfansy@126.com。

引用格式: 杨帆, 王红斌, 方健, 等. 基于 K-means 聚类和随机森林的电缆风险评估及修复决策 [J]. 南京航空航天大学学报, 2024, 56(5): 892-899. YANG Fan, WANG Hongbin, FANG Jian, et al. Cable risk assessment and repair decision based on K-means clustering and random forest [J]. Journal of Nanjing University of Aeronautics & Astronautics, 2024, 56(5): 892-899.

10 kV 配电系统中交联聚乙烯绝缘电缆的使用越来越多,且工作环境日趋复杂^[1-2]。在热、潮湿、化学腐蚀和机械损伤等作用下,电缆的绝缘性能会快速下降,导致风险性不断增大^[3]。近年来因电缆故障而导致的线路跳闸甚至火灾事件层出不穷,造成了巨大的经济损失^[4]。因此,亟须对电缆进行及时、可靠的修复,制定科学的修复决策方案。目前,国内外学者针对电缆风险评估的研究较少。文献[5]利用模糊数学理论对电缆线路中各类组件进行风险综合评估,包括电缆本体、中接头和过电压保护器等。通过各类电缆线路组件的状态评价数据构建模糊综合评判矩阵,再与给定的等级参数矩阵点乘,计算出最终的风险评估值。文献[6]同样将电缆线路中各个组件作为分析对象,利用失效危险指数计算各组件的故障率,再考虑各组件的修复时间和线路的重要度进行最终的风险评估。利用不同线路中各组件的风险评估结果可以制定最优检修策略。文献[7]根据电缆线路的地区位置、工作环境和经济发展重要性等,构建了层次结构模型,并利用判断矩阵计算影响电缆线路资产损失的各风险因素的权重值,最终求取综合风险评估值,实现对电缆线路的评估。上述风险评估研究主要是针对电缆线路及其组件,利用了模糊综合评判、失效危险指标和层次分析法等方法来计算评估值,然后通过评估值进行风险评估。然而,上述文献并未对实际评估指标的退化数据进行风险等级划分且存在风险评估指标单一的问题。K-means 聚类算法具有聚类效果好且收敛速度快的优点,可以实现对老化指标的区间划分。因此,本文借鉴上述文献中风险评估的研究思路,对断裂伸长率、断裂强度、介电常数、介质损耗因数和电气强度五个老化指标进行分析,构建出电缆多老化指标风险矩阵,为电缆修复决策提供理论和数据支撑。

电缆的修复方式主要包括绕包修复和截断换新。采用截断换新的修复方式将造成停电,而采用绕包修复又会导致供电可靠性差的问题。因此,根据电缆的风险等级制定科学合理的修复决策,在保证安全性的前提下完成修复成为一种研究趋势。将电缆的绝缘状态与修复方式的选择建立一一对应关系,便可利用分类模型实现电缆修复方式的选择。目前分类模型主要包括支持向量机、随机森林(Random forest, RF)和反向传播(Back propagation, BP)神经网络等。文献[8]将改进鲸鱼优化算法与支持向量机分类模型相结合,实现蒸馏装置的微小故障快速分类。文献[9]将慢特征分析与 RF 分类模型相结合,实现风机叶片结冰故障的判断。文献[10]将混沌天牛群算法与 BP

神经网络分类模型相结合,以提高分类模型的正确率。文献[11]将模拟电路的故障特征输入朴素贝叶斯分类器中进行训练和验证,该分类器具有很好的诊断精度。文献[12]针对核电站中执行器和传感器故障问题,采用 K 近邻算法对故障进行分类。该方法能够对多种执行器故障、传感器故障以及多个致动器和传感器同时发生的故障进行分类。上述参考文献中所采用的支持向量机分类模型的分类准确度受到超参数的影响程度大,需要结合启发式优化算法进行超参数选择,其寻优时间较长,进而导致分类效率降低。此外,采用神经网络分类模型对已知数据进行训练,随着隐藏层数量和层数的增加,其训练时间也将增加。RF 算法通过决策树和 Bagging 抽样可以快速解决分类问题,是一种高效且可靠的分类算法。因此,本文利用多老化指标风险矩阵和层次分析法确定修复决策分类标签,并通过构建 RF 分类模型实现电缆修复决策的快速判断。

1 多老化指标风险矩阵的构建

1.1 电缆风险等级和风险程度的定义

首先需要获取电缆的绝缘状态,然后根据绝缘状态定义风险等级和风险程度,如表 1 所示。其中风险等级被分为 I 类、II 类、III 类和 IV 类,分别对应绝缘状态为差、中、良和优。各老化指标的数值均可以从不同角度反映电缆的绝缘状态,这些数据偏离初始值越大,所对应的电缆绝缘性能越差。

表 1 风险等级定义
Table 1 Definition of risk level

| 风险等级 | 风险程度定义 |
|---------------|------------------------------------|
| I 类(绝缘状态为差) | 引起绝缘击穿、火灾和停电等重大事故 |
| II 类(绝缘状态为中) | 电缆绝缘性能退化严重,但不足以导致击穿事故,会导致非计划性维护或修理 |
| III 类(绝缘状态为良) | 电缆绝缘性能略微下降,各项绝缘指标均处于安全状态 |
| IV 类(绝缘状态为优) | 电缆绝缘性能未发生退化,各项绝缘指标均处于健康状态 |

1.2 老化指标数据集

采用与电缆风险评估相关的老化指标构建数据集,具体包括断裂伸长率、断裂强度、介电常数、介质损耗因数和电气强度。这些数据集来自天津电力公司 10 kV 配网变电站中的同一批电缆,在 13 年期间通过不同试验采集到^[13]。所有数据均在离线条件下获取,力学性能测试是通过将电缆绝缘层压切为哑铃试样,采用万能拉力机进行试验测试,获得断裂伸长率和断裂强度的数据。介电性能测试是使用德国 Novo Control 公司生产的 Concept 40 型宽频介电阻抗谱分析仪执行的,这项测试在室温

和真空条件下进行,获取电缆试样的介电常数和介质损耗因数的数据。电气性能测试是通过短时快速升压的方法进行的,在测试过程中,以每秒500 V的速度升压,直到发生击穿,从而获得电缆试样的电气强度数据。老化指标的散点如图1所示。从图中可以看出,相邻老化指标间的间距较大,数据量较小。因此,为了获得足够的的数据,本文采用三次样条插值^[14]的方式补充数据,具体为:将已知 n 个数据点作为插值节点,插值节点的自变量 x 为老化时间,因变量 y 为多维老化指标。其中老化时间符合 $x_1 < x_2 < x_3 < \dots < x_n$,老化指标符合 $y_1 > y_2 > y_3 > \dots > y_n$,对于任意区间上的插值函数如下

$$y_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i \quad (1)$$

式中: y_i 为第 i 个区间的老化指标, a_i 、 b_i 、 c_i 、 d_i 均为第 i 个区间上插值函数的待求参数。

最终构建 $1\,000 \times 6$ 的数据集,数组的前5列为老化指标,第6列为分类标签。图1中的平滑退化曲线为三次样条插值所扩充的结果。

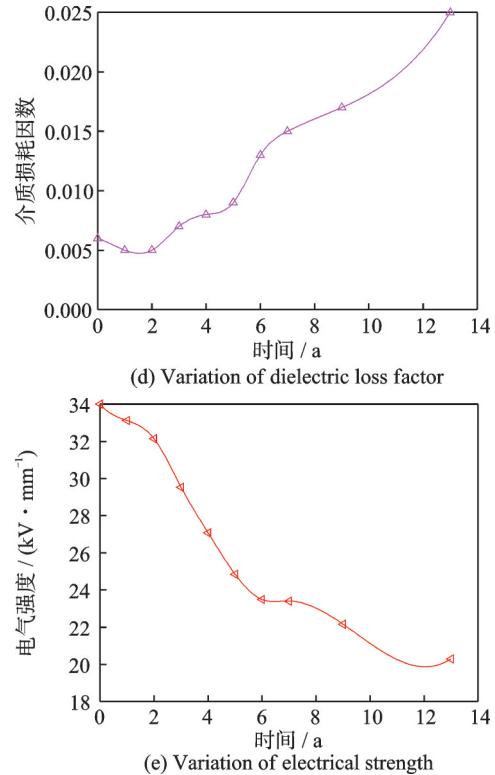
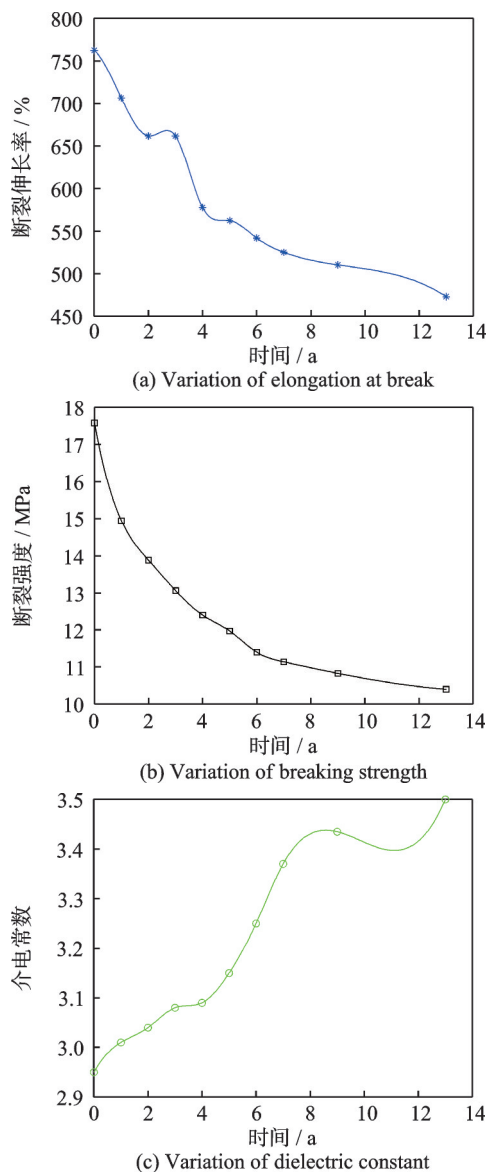


图1 不同老化指标的变化曲线

Fig.1 Variation curves of different aging indicators

1.3 基于K-means聚类算法的性能参数分类

为了构建上述 $1\,000 \times 6$ 的数据集第6列的分类标签,本文采用K-means聚类算法^[15]对多维老化指标进行优、良、中、差4个等级划分。K-means聚类算法的基本步骤如下:

(1) 在每个指标所在数据集中,随机选择 k 个指标值作为最初的聚类中心点。

(2) 根据各指标数据点所在区间属性进行聚类。如果某个指标值 X_j 被分配至第 i 个聚类,则所对应的权重 $w_{ji}=1$,否则将其置为零值,具体如下

$$\begin{cases} \sum_{i=1}^k w_{ji} = 1 & \forall j = 1, 2, \dots, n \\ \sum_{i=1}^k \sum_{j=1}^n w_{ji} = n \end{cases} \quad (2)$$

$$w_{ji} = \begin{cases} 1 & \|X_j - C_i\| \leq \|X_j - C_m\|, \forall m \neq i \\ 0 & \text{其他} \end{cases} \quad (3)$$

(3) 建立聚类过程中的目标函数 J ,如式(4)所示。根据目标函数值是否保持不变,判断是否继续进行迭代计算。如果目标函数值达到收敛,表明聚类结果已达到要求。如果目标函数值未达到收敛,则继续进行迭代计算,直至收敛,从而完成整个聚类过程。

$$J = \sum_{i=1}^k J_i = \sum_{i=1}^k \sum_{j=1}^n w_{ji} \|X_j - C_i\|^2 \quad (4)$$

(4) 由式(5)计算出新的聚类中心点,并返回至上述步骤(2)中。

$$C_j = \frac{\sum_{j=1}^n w_{ji} X_j}{\sum_{j=1}^n w_{ji}} \tag{5}$$

基于 K-means 聚类算法对反映电缆绝缘性能的 5 个老化指标进行 4 个区间聚类。在老化指标变化曲线中随机选择 4 个聚类中心点,通过迭代计算得出 4 组数据,每组数据被认为具有相同的区间属性。断裂伸长率、断裂强度、介电常数、介质损耗因数和电气强度等指标的聚类结果分别如图 2(a~e)

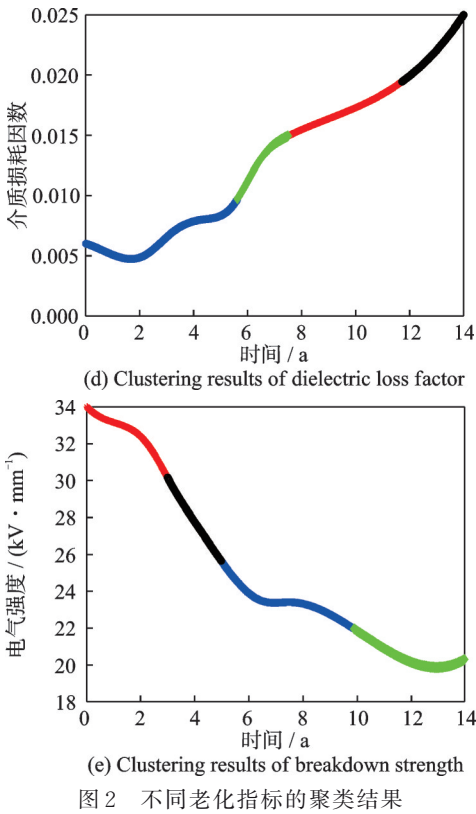
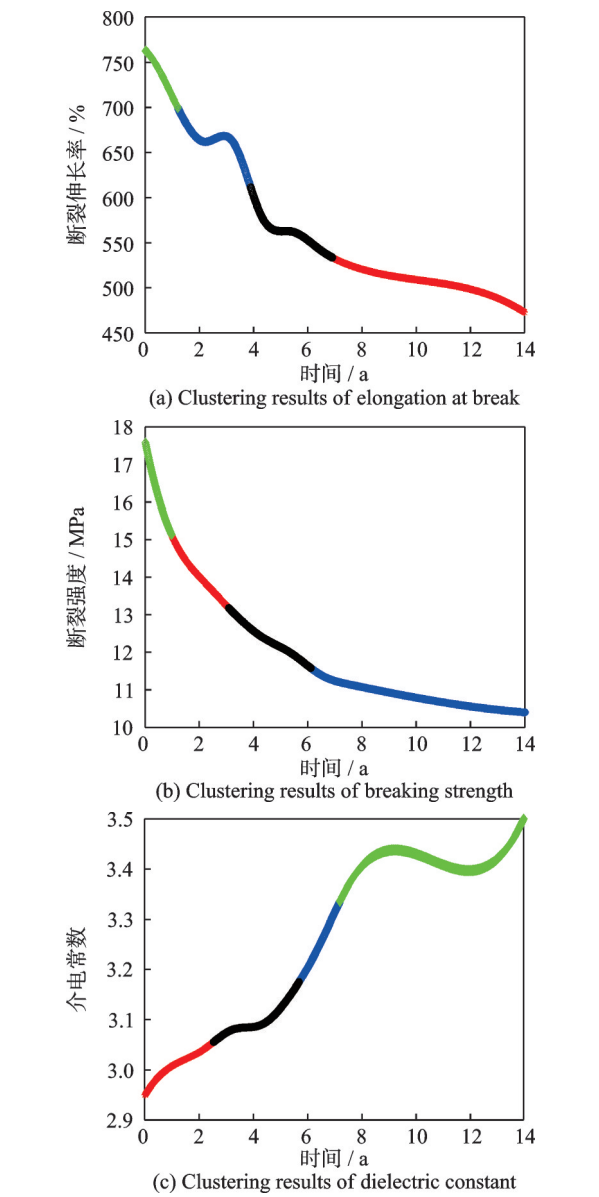


图 2 不同老化指标的聚类结果

Fig.2 Clustering results of different aging indicators

所示。按照各老化指标的聚类结果即可实现优、良、中、差 4 个等级划分。

1.4 多老化指标风险矩阵

根据上述 5 个老化指标的聚类结果建立多老化指标风险矩阵如表 2 所示。断裂伸长率指标在“优”“良”“中”“差”区间的变化范围分别为 698.52~762.50、612.61~698.52、533.45~612.61、473.00~533.45;断裂强度指标在“优”“良”“中”“差”区间的变化范围分别为 15.09~17.58、13.16~15.09、11.57~13.16、10.40~11.57;介电常数指标在“优”“良”“中”“差”区间的变化范围分别为 2.95~3.05、3.05~3.17、3.17~3.33、3.33~3.50;介质损耗因数指标在“优”“良”“中”“差”区间的变化范围分别为 0.004 7~0.007 1、0.007 1~0.012 4、0.012 4~0.018 7、0.018 7~0.025 0;电气强度指标在“优”“良”“中”“差”区间的变化范围分别为 30.25~34.00、25.68~30.25、22.00~25.68、19.87~22.00。

表 2 老化指标风险区间
Table 2 Aging index risk interval

| 老化指标 | 优 | 良 | 中 | 差 |
|-----------------------------|-----------------|-----------------|-----------------|-----------------|
| 断裂伸长率/% | 698.52~762.50 | 612.61~698.52 | 533.45~612.61 | 473.00~533.45 |
| 断裂强度/MPa | 15.09~17.58 | 13.16~15.09 | 11.57~13.16 | 10.40~11.57 |
| 介电常数 | 2.95~3.05 | 3.05~3.17 | 3.17~3.33 | 3.33~3.50 |
| 介质损耗因数 | 0.004 7~0.007 1 | 0.007 1~0.012 4 | 0.012 4~0.018 7 | 0.018 7~0.025 0 |
| 电气强度/(kV·mm ⁻¹) | 30.25~34.00 | 25.68~30.25 | 22.00~25.68 | 19.87~22.00 |

1.5 电缆风险评价指标体系

从电缆的介电性能、导热性能、力学性能和运行工况4个方面构建风险评估指标体系,如表3所示。其中电缆介电性能指标包括介质损耗因数、介电常数和电气强度;导热性能指标包括载流量和表面温度;力学性能指标包括断裂伸长率、断裂强度和硬度;运行工况指标包括电压、电流、环境温度、环境湿度和运行年限。载流量根据电缆结构参数和最大工作温度计算得到,环境温度和表面温度通过铂热电阻或热成像仪测量得到,硬度通过邵氏硬度计测量得到,电压通过电阻分压型电压传感器测量得到,电流通过电流互感器测量得到,环境湿度通过电容式湿度传感器和电阻式湿度传感器测量得到,运行年限通过工作人员的统计数据得到。

表3 电缆风险评估指标体系

Table 3 Cable risk assessment index system

| A层次 | B层次 | C层次 |
|------|------|--------|
| 综合评估 | 介电性能 | 介质损耗因数 |
| | | 介电常数 |
| | | 电气强度 |
| | 导热性能 | 载流量 |
| | | 表面温度 |
| | 力学性能 | 断裂伸长率 |
| | | 断裂强度 |
| | | 硬度 |
| | 运行工况 | 电压 |
| | | 电流 |
| | | 环境温度 |
| | | 环境湿度 |
| | | 运行年限 |

2 修复决策方法

2.1 随机森林原理

RF算法是Breiman等在1995年提出一种基于决策树和Bagging方法的智能算法^[16]。该算法的功能主要包括分类和回归。其中,决策树是以基尼指数作为属性选择量度,表达式如下

$$G_D = 1 - \sum_{j=1}^m P_i^2 \tag{6}$$

式中: G_D 为基尼指数, P_i 为概率, m 为类别总数。基尼指数反映了从数据集 D 中随机抽取两个样本,其类别标记不一致的概率。属性 A 的基尼指数如式(7)所示,以在属性 A 下最小基尼指数的子集 D_1 和 D_2 作为相应的分裂子集。

$$G_{D,A} = \frac{|D_1|}{D} G_{D_1}(D_1) + \frac{|D_2|}{D} G_{D_2}(D_2) \tag{7}$$

Bagging方法通过对样本数据集进行部分随机

抽取生成决策树。该方法能够有效提高分类器的泛化能力。RF分类功能的实现流程如图3所示。构建RF分类模型的具体步骤如下:

(1) 利用Bootstrap采样方法从多维老化指标及其对应分类标签所组成的训练集中抽取 n 个子集,其中 n 为决策树的个数。

(2) 步骤(1)中随机抽取的每个子集均用于训练出相应的决策树。在该训练过程中,每个节点分裂选用的特征都是在特征向量中随机抽取所得,且基于基尼系数选择最佳的分裂所用特征。

(3) 集成 n 棵决策树从而构成RF分类模型。

将测试数据集中的一组变量(断裂伸长率、断裂强度、介电常数、介质损耗因数和电气强度)输入至构建好的RF分类模型中,每个决策树同时处理此组变量并得出相应的决策结果。然后RF分类模型的最终决策结果通过所有决策树结果的多数投票决定,即获得票数最多的类别为RF分类模型的最终决策结果。得益于Bootstrap抽样和输入变量的随机选择,RF算法中决策树之间的相关性较低,可以获得较高的分类精度。与BP神经网络、支持向量机和极限学习机等分类算法相比,RF算法可以提供优越的鲁棒性和泛化能力。因此,本文将RF算法用于电缆修复决策中。

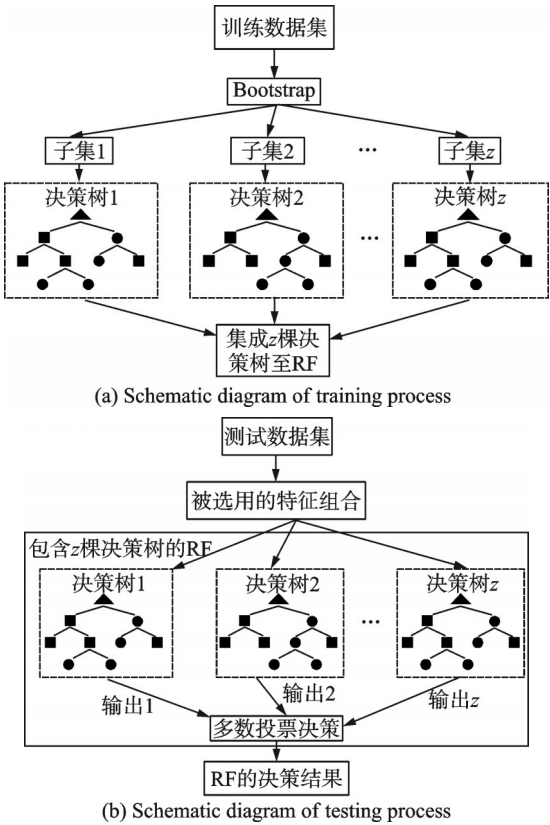


图3 RF分类功能的实现流程

Fig.3 Implementation flow of RF classification function

2.2 修复决策对应的分类标签设置

由于上述 5 个老化指标的风险等级区间不同,难以确定以某一个风险等级区间设置统一的分类标签。因此,需要求取 5 个老化指标的权重值,并以权重值最大的老化指标所对应风险等级区间进行分类标签设置。每个表征电缆绝缘性能的老化指标的权重值是通过层次分析法得到^[17-18],该方法是将电缆风险评估中的老化指标作为一个系统,通过构建上下级层次结构、构建判断矩阵和重要度排序 3 个方面进行分析。具体步骤如下:

(1) 划分电缆风险评估老化指标体系的层次结构,确定各层次之间的关系及主要元素。

(2) 根据表 4 中任意两个老化指标对比结果的标号定义,对各老化指标进行相互比较,从而构建出判断矩阵。

表 4 对比结果的标号定义
Table 4 Definition of scale

| 标号 | 定义 |
|----------|---------------------------------|
| 1 | 重要 |
| 3 | 略显重要 |
| 5 | 明显重要 |
| 7 | 非常重要 |
| 9 | 极为重要 |
| 2,4,6,8 | 邻近的中间值 |
| a_{ij} | 老化指数 <i>i</i> 与 <i>j</i> 的重要性之比 |
| a_{ji} | 老化指数 <i>j</i> 与 <i>i</i> 的重要性之比 |

(3) 利用特征向量法计算出不同老化指标的权重值,以实现重要度排序。

(4) 根据一致性指标和一致性比率对判断矩阵进行一致性检验。其中一致性指标的计算如式(8)所示,一致性比率的计算如式(9)所示。当 CR 值小于 0.1 时,所构建的判断矩阵满足要求。

$$CI=(L_{\max}-n)/(n-1) \tag{8}$$

$$CR=CI/RI \tag{9}$$

式中:CI 为一致性指数,CR 为一致性比率, L_{\max} 为最大特征值, n 为老化指标个数,RI 为平均随机一致性指数。

本节将电缆风险评估综合值作为第一层次,将 5 个老化指标作为第二层次。本文由 12 位专家根据相关论证和自身知识给出判断矩阵中的各元素值,如式(10)所示。根据式(8,9)计算得到该判断矩阵的一致性比率为 0.031 1,该值小于 0.1,即表明所构建的 5×5 的判断矩阵是符合要求的。利用特征向量法求得各老化指标的主观权重,为了保证客观性,利用熵权法计算出各老化指标的客观权重,最后根据式(11)计算出各老化指标的综合

权重。

$$A=\begin{bmatrix} 1 & 2 & 5 & 4 & 1/3 \\ 1/2 & 1 & 4 & 3 & 1/4 \\ 1/5 & 1/4 & 1 & 1/2 & 1/7 \\ 1/4 & 1/3 & 2 & 1 & 1/6 \\ 3 & 4 & 7 & 6 & 1 \end{bmatrix} \tag{10}$$

$$W_k=\frac{\sqrt{\alpha_k\beta_k}}{\sum_{k=1}^n\sqrt{\alpha_k\beta_k}} \tag{11}$$

式中: W_k 为综合权重, α_k 为主观权重, β_k 为客观权重。各权重值如表 5 所示,电气强度所占综合权重最大,最大值为 0.349 4。因此,以电气强度的风险等级区间划分结果对老化指标数据进行编号,其中落在“优”和“良”区间的老化指标对应的标签设为 0,表示可以不对电缆进行修复。落在“中”区间的指标对应的标签设为 1,表示建议对电缆进行修复。落在“差”区间的指标对应的标签设为 2,表示必须对电缆进行修复。风险等级、修复决策和分类标签三者对应关系如表 6 所示。将其余 4 个老化指标按照相同的时间序列区间设定标签,构建出每个时刻下具有 5 个输入特征、1 个输出标签的数据集,从而训练出受 5 个老化指标影响的分类模型。

表 5 不同老化指标对应权重值
Table 5 Corresponding weight values of different aging indicators

| 老化指标 | 主观权重 | 客观权重 | 综合权重 |
|--------|---------|---------|---------|
| 断裂伸长率 | 0.235 9 | 0.223 8 | 0.243 6 |
| 断裂强度 | 0.152 9 | 0.303 0 | 0.228 1 |
| 介电常数 | 0.046 1 | 0.141 5 | 0.085 6 |
| 介质损耗因数 | 0.068 8 | 0.112 8 | 0.093 4 |
| 电气强度 | 0.496 3 | 0.219 0 | 0.349 4 |

表 6 不同修复决策对应的分类标签
Table 6 Classification labels corresponding to different repair decisions

| 风险等级 | 修复决策 | 分类标签 |
|------|-------|------|
| 优、良 | 可以不修复 | 0 |
| 中 | 建议修复 | 1 |
| 差 | 必须修复 | 2 |

2.3 修复决策的分类结果

以 1 000 组数据对 RF 分类模型进行训练和验证,每组数据中包括 5 个老化指标和 1 个分类标签。首先将 1 000 组数据进行随机排序,使不同风险等级下的老化指标数据集随机分布,以更好地验证 RF 分类模型的准确性。然后将前 600 组数据集作为训练集,后 400 组数据集作为测试集。最后利用训练集得到的 RF 分类模型对测试集数据进行修复决策的分类。取 400 次分类结果中的最后 11

次修复决策分类结果进行显示,如图4所示。从图中可以看出,训练好的RF分类模型根据输入的5个老化指标,可以准确输出相应的分类标签,实现了电缆修复决策的最终判断。该次分类结果的400组数据集中仅有1组数据的分类结果出现错误,正确率为99.75%。为了避免一次分类结果的偶然性,重复上述分类步骤10次,得到10次分类结果中的正确个数及正确率如表7所示,10次分类结果的平均正确率为99.70%。

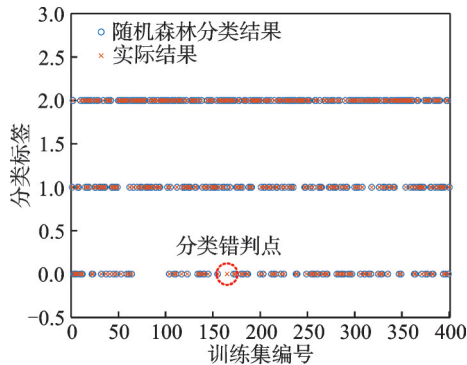


图4 修复决策的分类结果

Fig.4 Classification results of repair decisions

表7 10次分类结果

Table 7 Classification results of 10 times

| 分类次数 | 正确个数 | 正确率/% | 运算时间/s |
|------|------|--------|--------|
| 1 | 399 | 99.75 | 0.21 |
| 2 | 398 | 99.50 | 0.15 |
| 3 | 400 | 100.00 | 0.13 |
| 4 | 398 | 99.50 | 0.11 |
| 5 | 397 | 99.25 | 0.13 |
| 6 | 399 | 99.75 | 0.11 |
| 7 | 400 | 100.00 | 0.11 |
| 8 | 399 | 99.75 | 0.12 |
| 9 | 399 | 99.75 | 0.12 |
| 10 | 399 | 99.75 | 0.12 |

为了进行算法对比,分别利用RF算法和长短期记忆(Long short-term memory, LSTM)算法对电缆修复决策进行10次分类,统计两种算法10次分类结果的平均正确率和运算时间。RF算法运行过程中决策树的数量为12,每次分裂决策树节点时,所有可用的特征中能随机选择的特征数量为2。LSTM算法中包括1个序列输入层、1个LSTM层、1个全连接层、1个Softmax层和1个分类层,训练过程最大迭代次数为100次,梯度阈值为1,初始学习率为0.000 5,学习率下降因子为0.05。表8显示了两种分类算法的正确率及运算时间的对比,从表中可以看出,RF算法的分类正确率要高于LSTM,且运算时间更短。

表8 两种分类算法的正确率及运算时间

Table 8 Correct rate and computing time of two classification algorithms

| 预测方法 | 正确率/% | 运算时间/s |
|------|-------|--------|
| RF | 99.70 | 0.13 |
| LSTM | 92.58 | 19.23 |

3 结 论

针对配电系统中电缆绝缘退化相对严重的问题,根据电缆老化指标建立风险评估矩阵,并制定出可靠且准确的电缆修复决策。本文创新性地建立了多老化指标风险评估矩阵,并提出了电缆修复决策方法,得出如下结论:

(1)将断裂伸长率、断裂强度、介电常数、介质损耗因数和电气强度5个老化指标进行风险等级划分,构造出多老化指标风险矩阵。

(2)根据权重值最大指标的风险等级区间设定3段时间序列,并对应3种分类标签,从而为修复决策分类模型提供具有5个输入特征和一个输出标签的数据集。

(3)所提电缆修复决策方法的平均正确率达到99.70%,平均正确率高于LSTM,且具有更快的运算速度。

在未来的工作中,将在线采集更多表征电缆绝缘状态的特征量,构建完备的风险评估体系。并在此基础上提出更加科学的风险评估方法和修复决策方法。

参考文献:

- [1] 时珊珊,崔正达,陈颖,等. 电气化交通和城市电网协同韧性提升方法综述[J]. 电工电能新技术, 2022, 41(3): 43-54.
SHI Shanshan, CUI Zhengda, CHEN Ying, et al. Summary of methods for improving the collaborative resilience of electrified transportation and urban power grid[J]. Advanced Technology of Electrical Engineering and Energy, 2022, 41(3): 43-54.
- [2] ZHANG Xi, LIU Youbo, DUAN Jiajun, et al. DDPG-based multi-agent framework for SVC tuning in urban power grid with renewable energy resources [J]. IEEE Transactions on Power Systems, 2021, 36(6): 5465-5475.
- [3] 冯庆燎,陈申宇,张攀. 基于配网管廊的电力、通信两张网运营环境改善初探[C]//2017年度中国电力企业管理创新实践优秀论文大赛论文集(上册). 北京: 中国电力企业管理, 2018: 306-311.
FENG Qingliao, CHEN Shenyu, ZHANG Pan. Preliminary exploration of improving the operation environment of power and communication dual networks

- based on distribution network corridors[C]//Proceedings of the 2017 Annual China Electric Power Enterprise Management Innovation Practice Excellent Paper Competition(Volume I). Beijing: China Electric Power Enterprise Management, 2018: 306-311.
- [4] 张文涛. 电缆及电缆沟实时综合防控系统初探[J]. 企业技术开发, 2015, 34(29): 15-16.
ZHANG Wentao. Preliminary study on real-time comprehensive prevention and control system of cable and cable trench[J]. Technology Development of Enterprise, 2015, 34(29): 15-16.
- [5] 黄新波, 吴明松, 朱永灿, 等. 基于模糊数学的电缆线路风险评估模型研究[J]. 高压电器, 2021, 57(9): 19-25.
HUANG Xinbo, WU Mingsong, ZHU Yongcan, et al. Research on the risk assessment model of cable lines based on fuzzy mathematics[J]. High Voltage Apparatus, 2021, 57(9): 19-25.
- [6] 廖雁群, 袁婧, 徐小冰, 等. 基于风险评估的高压电缆巡检周期优化模型[J]. 高电压技术, 2021, 47(1): 305-314.
LIAO Yanqun, YUAN Jing, XU Xiaobing, et al. Optimization model for high-voltage cable inspection cycle based on risk assessment[J]. High Voltage Engineering, 2021, 47(1): 305-314.
- [7] 吴明松, 黄新波, 朱永灿, 等. 基于层次分析法的电缆线路风险评估[J]. 西安工程大学学报, 2019, 33(6): 637-642.
WU Mingsong, HUANG Xinbo, ZHU Yongcan, et al. Risk assessment of cable lines based on analytic hierarchy process[J]. Journal of Xi'an Polytechnic University, 2019, 33(6): 637-642.
- [8] 李慧, 徐海亮, 王浩, 等. 基于改进鲸鱼算法优化支持向量机的故障诊断的研究与应用[J]. 科学技术与工程, 2022, 22(13): 5284-5290.
LI Hui, XU Hailiang, WANG Hao, et al. Research and application of fault diagnosis based on support vector machine optimized by improved whale algorithm[J]. Science and Technology and Engineering, 2022, 22(13): 5284-5290.
- [9] 恽钰蛟, 贾鹏, 伍兴达, 等. 基于慢特征分析和随机森林的风叶结冰故障监测[J]. 工业控制计算机, 2022, 35(6): 103-105.
YUN Yujiao, JIA Peng, WU Xingda, et al. Wind blade icing fault monitoring based on slow feature analysis and random forest[J]. Industrial Control Computer, 2022, 35(6): 103-105.
- [10] 王丽, 陈基漓, 谢晓兰, 等. 基于混沌天牛群算法优化的神经网络分类模型[J]. 科学技术与工程, 2022, 22(12): 4854-4863.
WANG Li, CHEN Jili, XIE Xiaolan, et al. Neural network classification model based on chaotic beetle swarm optimization[J]. Science and Technology and Engineering, 2022, 22(12): 4854-4863.
- [11] HE W, HE Y G, LI B, et al. A naive-bayes-based fault diagnosis approach for analog circuit by using image-oriented feature extraction and selection technique[J]. IEEE Access, 2020, 8: 5065-5079.
- [12] NAIMI A, DENG J, SHIMJITH S, et al. Fault detection and isolation of a pressurized water reactor based on neural network and K-nearest neighbor[J]. IEEE Access, 2022, 10: 17113-17121.
- [13] 宋鹏先, 王浩鸣, 房晟辰, 等. 基于综合健康指数的热缩材料寿命评估模型[J]. 绝缘材料, 2020, 53(11): 45-49.
SONG Pengxian, WANG Haoming, FANG Shengchen, et al. A life assessment model for heat shrinkable materials based on comprehensive health index[J]. Insulation Materials, 2020, 53(11): 45-49.
- [14] 许小勇, 钟太勇. 三次样条插值函数的构造与Matlab实现[J]. 兵工自动化, 2006, 25(11): 76-78.
XU Xiaoyong, ZHONG Taiyong. Construction and Matlab implementation of cubic spline interpolation function[J]. Army Automation, 2006, 25(11): 76-78.
- [15] 刘亮, 许灵, 刘斯文. 基于K-Means聚类的高校困难学生贫困等级划分研究: 以蚌埠学院为例[J]. 白城师范学院学报, 2017, 31(8): 32-35, 58.
LIU Liang, XU Ling, LIU Siwen. Poverty grading of poor college students based on K-means clustering: A case study of bengbu college[J]. Journal of Baicheng Normal University, 2017, 31(8): 32-35, 58.
- [16] Breiman. Random forests[J]. Machine Learning, 2001, 45(1): 5-32.
- [17] 邓雪, 李家铭, 曾浩健, 等. 层次分析法权重计算方法分析及其应用研究[J]. 数学的实践与认识, 2012, 42(7): 93-100.
DENG Xue, LI Jiaming, ZENG Haojian, et al. Analytic hierarchy process weight calculation method analysis and its application research[J]. Mathematics in Practice and Understanding, 2012, 42(7): 93-100.
- [18] 席禹, 于力, 蒋文辉, 等. 基于修正层次分析法的多阶段PMU最优配置[J]. 广东电力, 2022, 35(9): 87-93.
XI Yu, YU Li, JIANG Wenhui, et al. Multi-stage PMU optimal configuration based on modified analytic hierarchy process[J]. Guangdong Electric Power, 2022, 35(9): 87-93.