

基于 ELM 和迁移学习的电网设备多因素综合故障率分析方法

陈丽惠, 李哲, 周键宇, 马立熠, 杨梓萌

(云南电网有限责任公司保山供电局, 云南 保山 678000)

摘要: 电网设备的故障率分析是电网风险预警和运行评估的基础和关键, 而电网设备多种故障因素与综合故障率的非线性权重关系的评估决定了故障率分析的准确度, 为充分利用历史故障信息数据以准确评估多故障因素与综合故障率的权重关系, 提出一种基于极限学习机 (ELM) 和迁移学习的电网设备多因素综合故障率分析方法; 首先, 设计了多源历史故障信息统计体系, 构建电网设备故障分析样本数据库; 将自身、过载、恶劣天气等作为主要故障因素, 通过机器学习 ELM 算法, 分析故障因素与综合故障率的权重关系, 并利用多个强相关样本集的有效迁移提高机器学习的准确度, 通过 ELM 结合迁移学习算法, 克服了电网设备综合故障率分析中样本数量少且统计周期长的难题, 充分利用历史故障信息, 实现多因素设备综合故障率的准确评估, 为风险评估和检修安排提高可靠的数据支撑; 最后通过实际电网运行实例, 验证了所提方法的可靠性和准确性。

关键词: 故障率; 故障信息; 极限学习机; 多因素; 迁移学习

Multi Factor Comprehensive Fault Rate Analysis Method of Power Grid Equipment Based on ELM and Transfer Learning

CHEN Lihui, LI Zhe, ZHOU Jianyu, MA Liyi, YANG Zimeng

(Baoshan Power Supply Bureau of Yunnan Power Grid Co., Ltd., Baoshan 678000, China)

Abstract: The fault rate analysis of power grid equipment is the basis and key on risk assessment and operation assessment of power grid, the evaluation of nonlinear weight relationship with multiple failure factors and comprehensive failure rate of power grid equipment determines the accuracy of failure rate analysis. In order to make full use of historical fault factors and accurately assess the synergistic relationship of multiple failure factors and comprehensive failure rate, a multifactor comprehensive failure rate analysis method based on extreme learning machine (ELM) is proposed. Firstly, a multi-source historical fault information statistics system is designed, the sample database for fault analysis of power equipment is constructed. The relationship between the comprehensive fault rate and fault factors, such as equipment itself, overload, bad weather, is analyzed by ELM. The effective migration of multiple sample sets is proposed to improve the accuracy of machine learning. The analysis and calculation of comprehensive failure rate of equipment based on historical fault information can be realized. The algorithm based on ELM and transfer learning overcomes the difficult problem of small sample quantity and long statistical period in comprehensive failure rate analysis for power grid equipment. And it can improve the reliable data support for risk assessment and maintenance arrangement. Finally, the reliability and accuracy of the method are verified by the actual power grid operation data.

Keywords: fault rate; fault information; extreme learning machine; multivariate; transfer learning

0 引言

电力设备作为电力系统的基本单元, 通过电网设备故障率的可靠评估实现电力系统各种越限可能进行评估, 可以有效预防对系统运行风险, 是电力系统可靠运行评估的关键所在。随着电网的日益发展, 电网设备的数目和关联程度呈指数级快速增长, 由此带来的设备检修任务及其重要性也逐渐增长^[1-3]。同时引起电网设备发生故障的因素很多, 例如大风、雷雨、冰雹、过载、超期服役、小动物等,

各种内部和外部的故障因素之间也相互影响, 这种关联性显然是非线性的, 缺乏统一的描述方式, 而且这些故障因素与设备综合故障率的关联性也会因地域、季节等条件的变化而变化; 针对数量庞大、关系繁杂的电网相关主要电网设备, 完全依靠人工经验去分析是无法实现的, 因而采用智能处理方法对其历史运行数据进行分析是提高电网检修效率, 进而提高电网整体运行安全水平的有效途径; 另一方面, 电网故障信息及其故障信息诊断结果是电网运行的重要经验, 也是广大运行人员运行经验的宝贵总结, 利

收稿日期: 2022-08-29; 修回日期: 2022-09-27。

基金项目: 南方电网科技项目(051200HA42210001)。

作者简介: 陈丽惠(1982-), 女, 云南曲靖人, 工程师, 主要从事电力系统监控、电力调度自动化方向的研究。

引用格式: 陈丽惠, 李哲, 周键宇, 等. 基于 ELM 和迁移学习的电网设备多因素综合故障率分析方法[J]. 计算机测量与控制, 2023, 31(4): 30-35, 48.

用故障相关历史数据的进一步提升设备故障率的计算分析准确度, 是利用电网运行经验总结加强电网运行稳定水平的必要之举, 也是电网运行人员可靠及时掌握电网运行状态的高效途径^[5-8]。

当前设备故障的相关研究集中在故障诊断领域, 文献 [9-10] 通过关联算法对故障进行了诊断及分类; 文献 [11] 对故障诊断的算法进行了研究; 文献 [12] 对故障文本进行分析和分类, 对故障信息进行抽取和分析; 而在设备故障率分析计算方面以基于历史数据的平均值计算实现; 文献 [13] 通过分故障因素的概率分布方式计算综合故障率; 文献 [14] 基于故障历史信息对实时故障率计算进行了改进; 综上, 当前对故障智能诊断的研究一定程度上解决了复杂电网设备风险评估与有限人力之间的矛盾, 但对于历史故障信息的集成分析利用对设备故障率分析的相关研究仍较少, 充分利用电网设备的历史故障信息数据, 建立故障因素与综合故障率的准确模型, 提高故障率的计算分析准确度, 对设备的运行状态评估及检修有着积极的指导意义和实用价值, 同时也是供电区域运行状态的评估的关键依据。在故障因素与综合故障率关联关系分析方面, 当前主要是通过统计概率模型及其优化的方式实现, 这种分析方式对多样故障因素与综合故障率关系的准确程度仍有待加强, 同时由于相互关系也会因地区、运行条件等而有所不同^[15-17], 基于故障概率的方式也会因样本的稀少而进一步降低故障率的评估准确度。随着人工智能学习技术的发展, 其非线性关系分析方面的应用也越来越多, 与电网设备多因素综合故障率分析的需求比较契合, 但设备故障率分析的相关样本的显著特点是统计周期长且数量少, 极限学习机 (ELM) 作为一种人工智能学习算法^[18-22], 具有泛化能力强, 计算速度快, 对小样本的适应性较好的特点, 与电网设备故障率分析的特点相适应。

本文针对多因素的综合故障率分析计算存在的因素协同的高度非线性问题, 提出一种基于 ELM 和迁移学习的电网设备多因素综合故障率分析方法。对 ELM 和迁移学习算法与电网设备多因素综合故障率机器学习的适应性进行分析, 进而设计了历史故障信息统计体系, 按照构建故障信息分析样本所需信息对不同来源的故障信息进行统计; 将设备自身、过载、恶劣天气 3 个方面作为故障的主要因素, 采用不同运行时间的历史故障率数据作为训练样本, 通过极限学习机 (ELM) 对综合故障率与各个单故障因素故障率的关系进行拟合, 为进一步丰富样本数量提高拟合关系的准确性, 对多个训练样本集进行迁移组合, 最后通过应用实例对提出的分析方法进行了验证。

1 ELM 及迁移学习算法

电网设备故障与多因素相关联, 各个故障因素之间也存在一定的相关性, 显然这之间的关联呈现很强的非线性, 与设备本身、所处的运行环境都密切相关, 作为人工智能学习的样本, 故障数据样本相对较少, 对设备综合故障率

的机器学习属于小样本的神经网络计算; 另一方面对于日益复杂的电力系统, 对设备综合故障率的计算要求兼顾可靠性和快速性, 以及时有效地对电网运行提供参考信息。针对故障样本的特点和故障结果分析的需求, ELM 是较为合适的一种机器学习算法, 实现对影响电网设备故障的主要因素的相关性质的分析和计算, ELM 具有结构简单、训练快速、泛化能力强的特点, 非常适合于小样本的神经网络训练, ELM 也是机器学习的一个重要研究领域。ELM 的结构是单隐层前馈神经网络, 由输入层、隐层和输出层组成, 与传统 BP 算法相比, 输入层到隐层的权值、隐层的阈值随机设定, 设定无需进行基于梯度的误差算法, 隐层到输出层的权值通过解方程组方式进行确定, 隐层的输出矩阵按照输出的误差最小为原则进行求解, 当输入层和输出层各个节点的权值和偏差得到后即完成了对输入与输出关系的非线性拟合。按照机器学习所得的非线性关系, 以计算样本作为输入层的输入, 进而得到所需的输出。

对于 D 个不相同的样本 (x_i, y_i) , 其中 x_i 为输入, y_i 为输出, 极限学习机的一般模型如式 (1) 所示:

$$y_i = \sum_{i=1}^D \beta_i h(\omega_i x_i + b_i) \quad (1)$$

其中: ω_i 为输入层节点到隐层第 i 节点权值; b_i 为隐层第 i 神经元的偏置值; h 为激励函数; β_i 为隐层到输出层的权值。

ELM 的学习过程等价于求解式 (1) 中的隐层到输出层的权值矩阵最小二乘解, 如式 (2) 所示:

$$\beta^* = H^+ Y \quad (2)$$

其中: β^* 为所求的隐层到输出层的输出权值最优解矩阵, H^+ 为 Moore-penrose 广义逆矩阵, Y 为输出矩阵。

迁移学习可视为对人工智能算法的一种准确度提升方法^[23], 将对一组样本数据机器学习时获得的模型相关知识应用到另一个机器学习中, 以实现在目标机器学习中获得更准确的模型, 在统计学习领域迁移学习已被证明能够改善人工智能学习模型的泛化性能。迁移学习的基本框架如图 1 所示, 其中源域为提取关联知识的学习, 要学习的新知识称为目标域, 源任务和目标任务可以不相同但需要相关。迁移学习主要针对样本获取不足的问题, 充分利用历史源域数据, 为目标任务提供有价值的信息。将迁移学习与人工智能算法相结合, 可以实现对大数据的充分利用, 提高人工智能算法的泛化能力, 还可进一步提高提高机器学习的训练效率, 是举一反三的发散思维的数学体现, 与电网设备多因素综合故障率分析的机器学习需求相适应, 通过找寻有内在关联性的数据样本进一步提高机器学习的可靠性。在迁移学习中有效地使用源域数据实现数据价值最大化, 防止负迁移, 增强机器学习可靠性, 是建立迁移学习的关键问题。

2 多源电网设备故障信息统计体系

当前的电网的运行监测系统中, 状态信息已被分类, 广泛存在于生产调度系统 (EMS)、设备资产管理系统

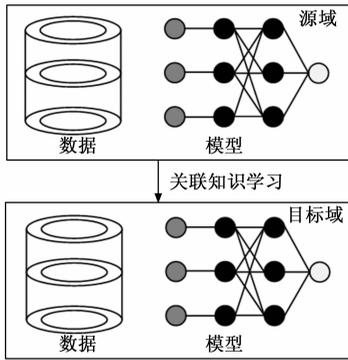


图 1 迁移学习框架

(PMS)、电力故障抢修管理系统 (TCM) 中, 具有跨专业、跨部门、数据格式多变、设备门类众多、数量庞大的特点^[24-25], 各统计系统既相互独立, 也有信息的重叠, 为确保故障信息样本的可靠确实, 需要从各个统计数据库对故障信息进行收集统计的系统, 本文的故障信息统计系统架构如图 2 所示, 信息的来源主要选取 SCADA 和 TCM 中的故障信息, 将 SCADA 系统中事故分闸、通信变位、辅助信息等作为触发信号, 针对事故分闸信息进行分析, 提取所需的故障数据, 将 TCM 系统中故障上报、辅助信息等作为触发信号, 针对故障上报信息进行分析, 提取所需的故障数据。同时由于电网中存在对日、周和月级别的统计信息, 通过设备编号和故障信息时间标签避免统计的故障信息中存在大量重复信息。

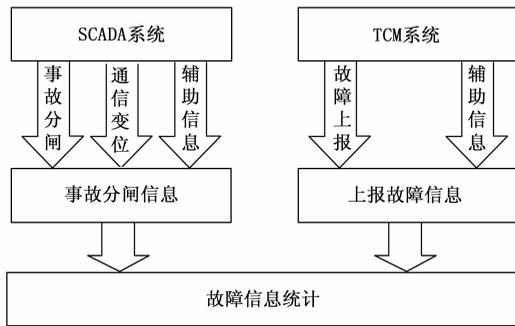


图 2 多源故障信息统计架构

电力设备的故障信息特点是专业性强, 需要进行互通交流, 而且故障信息以富文本的方式存在^[26-27], 为实现对故障信息的抽取和管理, 首先需要制定统一的词典, 对其专业性词汇进行识别, 便于大规模电网设备数据的使用。词典建立的依据主要是电网公司发布的设备故障案例文本、相关的导则和标准, 统计故障信息主要组成如图 3 所示, 包括设备信息、气候信息、故障信息、标签信息及其细分内容。通过预设词典的词语抽取案例文本中的关键描述, 获得确定结构化的故障信息数据, 形成故障信息统一知识库。涉及的电网设备主要包括变压器、电力电缆、断路器、隔离开关、负荷开关等对供电有较大影响的设备。故障信息统一知识库采用 MySQL 数据库对故障时间进行存储和管

理。知识库可以实现结构化故障信息的导入、查阅、修改、导出、统计等功能。

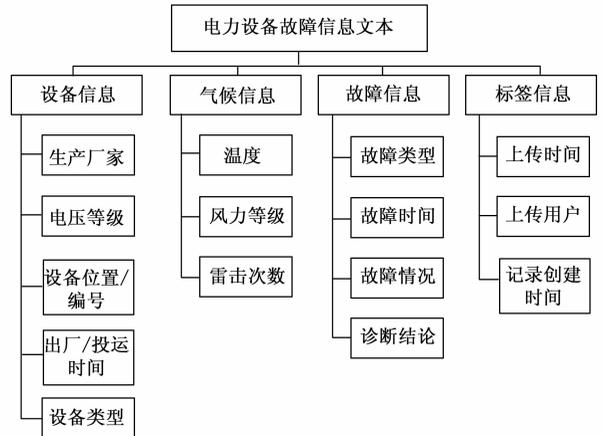


图 3 统计故障信息组成

依据已有的统计结果, 设备自身服役期限长导致的老化、自然灾害、过载保护动作、树木施工等外界破坏为故障因素。其中设备老化、过载和自然灾害的故障因素占比最大, 另一方面电网设备发生故障往往是多种因素综合的结果, 即各种故障因素之间存在一定的关联因素, 设备状态的內因与外因对设备的故障结果均有一定的影响; 设备自身导致的故障率基本由于使用年限较长后因设备老化、质量下降、辅助机构可靠性下降造成, 一般认为其故障概率在到达一定年限后与使用时间呈指数上升的关系; 设备过载不一定是设备出现了损坏, 当设备过负荷后继电保护动作也会造成供电时间的损失, 设备过载造成的故障概率因此也与继电保护动作的误差和准确性相关, 老化的设备可能造成保护测量装置的误差变大; 恶劣天气主要是雷击和大风大雨对电网设备造成的影响, 在设备服役期长及过载期间的影响也会较大; 另一方面树木施工等偶然性因素较大, 因此本文将设备自身、过载和恶劣天气作为主要的故障因素分析对象。

3 基于 ELM 的综合故障率分析

3.1 历史设备故障率统计计算

人工智能学习算法需要以一定量的样本数据为基础, 基于 ELM 的综合故障率的分析样本即为设备的历史故障率, 依据历史故障率数据中对综合故障率和各个故障因素之间的关系进行挖掘。本文的电网设备主要考虑变压器、电力电缆、断路器、隔离开关、负荷开关等, 依据上节所述的设备故障信息统计体系, 从统计的故障信息知识库中统计设备相关的故障信息, 从故障信息中按照设备自身、过载和恶劣天气的故障因素进行分类, 可得综合故障率及各种故障因素占总故障的权重, 计算公式如式 (3) 所示, 从公式中可以看出, 依据 ELM 分析所需的不同时间尺度的样本数据, 通过对应时间尺度的统计数据, 可以对应计算出不同统计时间长度对应的合故障率及各种故障因素占总

故障的权重。

$$\begin{cases} \lambda_m = \sum_{n=1}^3 N_{nm} / N_m \\ \omega_{mn} = N_{nm} / \sum_{n=1}^3 N_{nm} \end{cases} \quad (3)$$

其中: λ_m 为第 m 类设备的综合故障率; n 为对应的故障因素类别; N_{nm} 为第 m 类设备发生第 n 类故障的天数; N_m 为第 m 类设备的运行天数; ω_{mn} 为第 m 类设备第 n 类故障因素所占的权重。

3.2 基于 ELM 算法的单测试样本集计算

由于设备故障是小概率事件, 过载因素和天气因素造成的故障偶然性较强, 在短时间内的分布并不均匀, 因此相对较长的统计时间的准确度大于较短的统计时间, 比如以月平均统计数据相较于星期平均数据分析的规律更可靠, 另一方面设备自身原因发生故障的概率随着设备服役年限的增长而增大, 为使得样本能够包含时间的因素, 假设样本选取年份数为 Y , 以月数递增的方式根据式 (3) 进行故障率计算来构建样本, 即以 1 个月、2 个月、...、12 个月的故障率作为训练样本。

如前所述, 综合故障率的 ELM 神经网络包括输入层、隐层和输出层, 其中输入信号为依据统计数据计算出的权重与综合故障率的乘积, 输入信号分别对应的是本文所主要依据的设备自身、过载和恶劣天气 3 个故障因素, ELM 网络的输出层为拟合后的综合故障率, 隐层的神经元个数对提高拟合模型准确度有很大影响, 隐层节数目过少会影响 ELM 训练的精度, 过多则会加长训练时间增长, 也可能导致过拟合, 在本文对综合故障因素的分析中所注重的主要因素数目并不大, 因此隐层神经元个数通过测试结果比较来确定, 本文选取隐层神经元个数为 $2A+1$, 其中 A 为 ELM 输入个数, 综上, 电网设备多因素综合故障率 ELM 算法模型如图 4 所示。

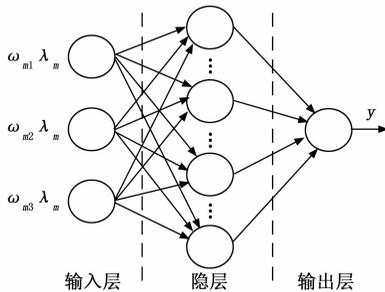


图 4 ELM 算法框图

ELM 的输入层和隐层之间的权值、隐层的偏置值通过随机设定, 其取值可参考文献 [9]。本文选择 Sigmoid 函数作为激励函数, 隐层的输出通过式 (4) 进行计算:

$$\begin{cases} h_i = \sigma(\sum_{k=1}^3 \omega_{k,i} x_k + b_i) \\ \sigma(x) = \frac{1}{1 + e^{-x}} \end{cases} \quad (4)$$

其中: h_i 为隐层第 i 个节点的输出, σ 为激励函数, $\omega_{k,i}$ 为输入层节点到隐层第 i 节点权值, b_i 为隐层第 i 神经元的偏置值, x_k 为输入层输入值, 即表 2 中所示数值, D 为输入层个数。

输出层输出与其输入, 即隐层的输出, 之间的关系如式 (5) 所示:

$$y = \sum_{j=1}^7 \beta_j h_j \quad (5)$$

其中: β_j 为隐层到输出层的权值, y 为输出层的输出。

在 ELM 算法中, 仅需对隐层到输出层的权值进行计算, 即可完成整个学习过程。对 β 的计算通过基于目标的优化方程求解得出。将样本与计算结果的近似平方差为最小的原则对方程进行求解。目标函数如式 (6) 所示:

$$\begin{cases} \min \| \mathbf{H}\boldsymbol{\beta} - \mathbf{S} \|^2 \\ \mathbf{H} = (h_1, h_2, \dots, h_7) \\ \boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_7) \end{cases} \quad (6)$$

其中: $\boldsymbol{\beta}$ 为隐层到输出层的输出权值矩阵, \mathbf{H} 为隐层输出矩阵, \mathbf{S} 为样本中的实际综合故障率。

通过调整权值使得输出结果与样本间的误差为最小, 尽可能对综合故障率与单个故障因素之间的非线性特性进行拟合。隐层到输出层的权值最优解可按式 (2) 进行计算。

综上, 基于 ELM 的单样本集的综合故障率模型计算主要步骤为:

- 1) 初始化阶段: 输入训练样本, 确定输入数据和样本输出值, 设定初始输入层到隐层权值 ω 和偏置值 b 。
- 2) 完成所有样本的计算, 根据式 (2)、(6) 计算出隐层到输出层的输出权值最优解矩阵, 完成对所有样本数据的学习, 计算出如式 (7) 所示的综合故障率计算模型:

$$\lambda'_{mc} = \sum_{n=1}^3 \beta_{mn}^* \omega_{mn} \lambda_{mc} \quad (7)$$

其中: λ_{mc} 待计算样本的综合故障率; ω_{mn} 为待计算样本中的第 m 类设备第 n 类故障因素所占的权重; λ_{mc} 为待计算样本的如式 (3) 计算的综合故障率。

3.3 多学习样本集的迁移计算

单区域的设备故障事件属于小概率随机事件, 训练样本数量有助于提高分析相互协同关系的精度, 在迁移学习的应用方面需要考虑样本数据集之间的关联性, 对于同型号的产品, 其受设备自身老化、过载和恶劣天气的影响可认为基本相同, 因此不同区域但是相同型号产品的故障信息数据存在内在的本质相关性, 相同型号产品的故障信息可以作为当前区域设备故障率的强相关补充学习样本, 另一方面, 综合多学习样本集的故障率计算模型也扩大了使用范围, 可靠地扩展至同型号产品的故障率计算。

尽管同型号产品存在内在的本质关联性, 但设备本身以及所处的环境都存在一定的差距, 即目标域和源域的规律分布必然存在一定的差距, 因此需要评估两者的差别并在此基础上进行适当的融合。对于 ELM 算法, 隐层到输

出层的权值 β 为主要的学习计算数值,也是综合故障率的最终计算依据,本文的迁移算法将隐层到输出层的权值作为目标域和源域差异评估的主要依据。按照 ELM 学习所得的综合故障率计算模型,对于同型号产品,可以用各个样本集对应的隐层到输出层的输出权值 β 对样本集之间的相关性进行判断。本文提出的多学习样本集迁移计算方法步骤如下:

1) 依据多源电网设备故障信息统计的数据库,对主区域和设定的迁移学习相关的其他区域的电网设备的故障信息进行提取,并按设备自身、过载和天气因素进行分类,以月为统计时间尺度,对设备历史综合故障率和设备故障因素权重进行计算,组成主区域和迁移学习相关的其他区域的 ELM 计算样本;

2) 针对主区域形成的样本集,根据当前设备的单故障样本集如 3.2 节所述进行 ELM 机器学习,拟合综合故障率与各个故障因素之间的关联关系,得到单样本集的隐层到输出层的输出权值 β_s ;

3) 针对迁移学习相关的其他区域形成的样本集,根据与当前待计算设备的同型号其他设备相关的单故障样本集如 3.2 节所述进行 ELM 机器学习,得到其他样本集相关隐层到输出层的输出权值 $(\beta_{r1}, \beta_{r2}, \dots, \beta_{rt})$,其中 t 为其他测试样本集数目;

4) 进一步分析目标域和源域规律分布的相关性,通过评估其他区域设备与主区域设备 ELM 学习所得隐层到输出层的输出权值,即计算 $\|\beta_s - \beta_{rd}\| (d=1, \dots, t)$,以此数值作为评估其相关性的依据,该数值越大则两样本的差距越大,数值越小则两样本的差距越小,按照此原则将计算数值从小到大进行排序,形成评估序列 PS ;

5) 选取评估序列 PS 中前 60%~70% 计算数据对应的隐层到输出层的输出权值,将其作为可靠迁移的样本数据,进一步与当前设备对应主区域故障集得到的隐层到输出层的输出权值进行平均值计算,作为最终的 ELM 隐层到输出层的输出权值。

综上,基于 ELM 和迁移学习的综合故障概率计算流程如图 5 所示。

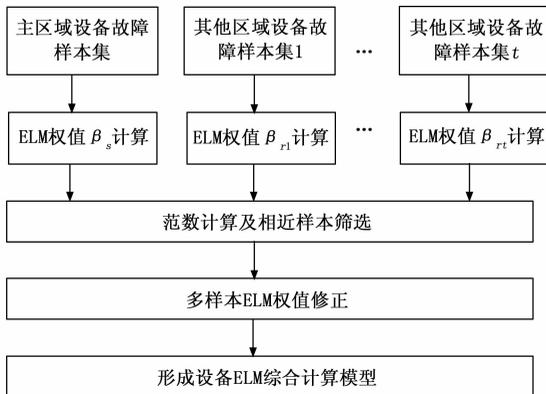


图 5 基于 ELM 和迁移学习的综合故障概率计算流程

4 实例验证

为对本文所提方法进行测试和分析,统计南方电网保山地区 100 个故障点 2014—2019 年主要电网设备变压器、电力电缆、断路器、隔离开关、负荷开关发生故障的文本信息,并对其中的次数及原因通过文本分析进行统计,将其作为多测试样本集,录入故障信息统一知识库,之后进一步对其中相同品牌的电网设备的故障次数及原因进行筛选。每个学习样本集由 1—60 个月累计的综合故障率及各个因素的权重组成。按照以上数据得到的机器学习模型计算出的设备故障率的云平台故障信息统计结果如图 6 所示。

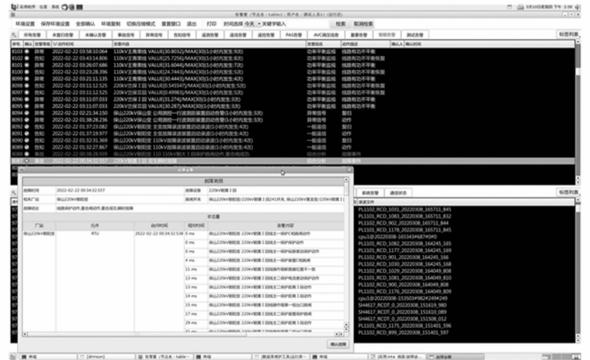


图 6 故障信息统计

本文的故障信息挖掘系统涉及到大量设备的故障信息统计和分析,对设备信息的需求往往涉及到多个区域,在电网现有采用的本地集中服务器方式下,多区域间信息的传递需要多个调度层级的周转和网络区域的转换,非常不利于信息的及时采集和反馈,因此本系统采用云平台的方式实现对信息的采集、分析,在云服务器中通过文中的极限学习机方法完成机器学习的模型计算,并将更新的模型数据参数下发至各本地服务器,本文采用的实验平台架构如图 7 所示,其中 ECS 服务器采用 16 核 CPU,内存 64 G,服务器调度采用弹性扩展管理,对系统资源的调度由接入的服务区域、设备分析数量以及待分析的文本数量决定。

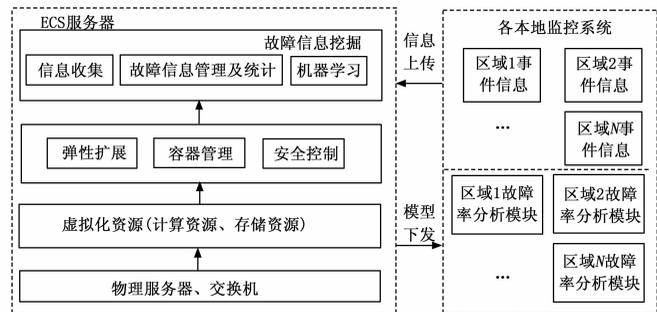


图 7 实验平台架构

为分析本文所提方法的可靠性,将历史月故障率平均值(算法 1)、单样本集 ELM 算法(算法 2)、激励函数为正弦函数多样本集 ELM 算法(算法 3)与本文激励函数为 Sigmoid 函数多样本集 ELM 算法(算法 4)的分析准确性

进行比较,采用 4 种算法对上述 100 个故障点根据 2019—2020 的单故障因素的故障率计算其综合故障率,准确性评价标准以 2019—2020 年故障率统计的偏差不超过 15%,多学习样本集 ELM 算法中选取与其最接近的 70% 作为综合权重计算目标。表 1、表 2 分别为各种算法下准确度和计算时间对比结果。从对比结果可以看出,采用月平均故障率与实际值的偏差较大,反映在准确率较低,采用不同的激活函数的多学习样本集数据迁移 ELM 算法比单样本集 ELM 准确率要高。算法 1 因其算法简单而计算速度较快,算法 2 因样本数目较少计算速度也较快,由于采用 ELM 的输出层较少,在运算过程中考察运算速度仍然是较快的,同时考虑几种算法计算时间的差距以及对应的准确率,本文所提算法适合电力故障数据大规模数据的智能快速处理所需的准确性和快速性。

表 1 不同算法的准确度比较

| 设备类型 | 准确度/% | | | |
|------|-------|------|------|------|
| | 算法 1 | 算法 2 | 算法 3 | 算法 4 |
| 变压器 | 46.6 | 69.5 | 81.5 | 83.2 |
| 电力电缆 | 47.2 | 71.3 | 81.5 | 84.3 |
| 断路器 | 45.6 | 65.2 | 78.9 | 79.8 |
| 隔离开关 | 46.7 | 69.5 | 80.3 | 81.2 |
| 负荷开关 | 50.2 | 74.3 | 84.5 | 85.6 |

表 2 不同算法的运算时间比较

| 设备类型 | 计算时间/s | | | |
|------|--------|------|------|------|
| | 算法 1 | 算法 2 | 算法 3 | 算法 4 |
| 变压器 | 0.05 | 0.21 | 0.32 | 0.33 |
| 电力电缆 | 0.05 | 0.22 | 0.43 | 0.42 |
| 断路器 | 0.05 | 0.25 | 0.36 | 0.35 |
| 隔离开关 | 0.05 | 0.22 | 0.35 | 0.37 |
| 负荷开关 | 0.05 | 0.23 | 0.39 | 0.38 |

综合实例运行的结果,所设计的历史故障信息统计体系及云平台的使用实现了对不同监控系统、不同区域间的历史故障信息数据进行有效统计,在此基础上构建了机器学习所需的故障信息样本,通过对电网设备典型设备,变压器、电力电缆、断路器等综合故障率分析,将本文所提算法与传统的月平均故障率和单样本集机器学习的方法在准确度和计算时间方面进行了比较,并进一步将不同计算参数下的本文算法也进行了比较,可以看出本文所提算法在准确度和计算时间方面均有明显优势,迁移学习提高了机器学习的准确性,ELM 本身的计算特点使得综合故障率计算具有快速性,因此本文所提基于 ELM 和迁移学习算法兼顾了计算可靠性和计算时效性,满足了多因素下设备综合故障率计算的可靠性和快速性需求,可以及时有效地为电网运行评估提供依据。

5 结束语

为实现对多故障因素下电网设备故障率的高准确度分

析以及设备历史故障信息的进一步挖掘利用,更好地指导运行人员对设备的检修及电网运行风险评估,本文提出基于 ELM 和迁移学习的多因素综合故障率分析方法。设计了多源历史故障信息统计体系对故障信息机器学习样本进行构建和管理,对其中的主要故障因素:设备自身、过载和恶劣天气,通过机器学习 ELM 算法根据历史数据分析故障因素之间以及故障因素与综合故障率之间的关系,为进一步提高机器学习的准确性,通过多学习样本集的迁移,对 ELM 拟合的输出权重矩阵进行修正。最后通过应用实例显示本文所提的基于 ELM 和迁移学习的电网设备多因素综合故障率分析方法有较高的可靠性,根据综合故障率提示可以有效地为运行人员提供检修计划和风险评估的数据支撑。

参考文献:

- [1] 孙元章,程林,何剑. 电力系统运行可靠性理论 [M]. 北京:清华大学出版社,2012.
- [2] FAN J M, WANG F, SUN Q Q, et al. Hybrid RVM-ANFIS algorithm for transformer fault diagnosis [J]. IET Generation Transmission & Distribution, 2017, 11 (14): 3637-3643.
- [3] 郭兴,徐武,唐文权,等. BPSO 优化算法在含 DG 配电网中的故障诊断研究 [J]. 计算机测量与控制, 2021, 29 (3): 88-92.
- [4] 陈锦铭,焦昊,郭雅娟,等. 基于多源数据融合的配网故障停电损失评估及热点区域可视化 [J]. 电力信息与通信技术, 2019, 17 (11): 7-13.
- [5] 孙铁军,曲丽萍,关海爽,等. 基于智能优化的模糊 Petri 网的电网故障诊断 [J]. 控制工程, 2021, 28 (1): 127-134.
- [6] 徐香香. 极端天气下电网故障在线预警及风险评估技术研究 [D]. 南京:东南大学,2020.
- [7] 郑洁云,胡梦月,胡志坚,等. 考虑可靠性及需求响应的配电网规划模型 [J]. 电力科学与技术学报, 2019, 34 (3): 173-182.
- [8] 闪鑫,戴则梅,张哲,等. 智能电网调度控制系统综合智能告警研究及应用 [J]. 电力系统自动化, 2015, 12 (1): 65-72.
- [9] 杨志淳,沈煜,杨帆,等. 考虑多元因素态势演变的配电变压器迁移学习故障诊断模型 [J]. 电工技术学报, 2019, 34 (7): 1505-1515.
- [10] 唐瑞佳,杨帆,李肇卿. 基于统计学验证的中低压用户供电可靠率关联性研究 [J]. 供用电, 2021, 38 (8): 40-46.
- [11] 苑津莎,张利伟,李中,等. 基于互补免疫算法的变压器故障诊断 [J]. 电工技术学报, 2015, 30 (24): 67-75.
- [12] 宇文梦柯. 基于领域知识的电网设备故障缺陷案例分析方法的研究与实现 [D]. 北京:北京邮电大学,2019.
- [13] 赵洪山,赵航宇. 考虑元件故障率变化的配电网可靠性评估 [J]. 电力系统保护与控制, 2015, 43 (11): 56-62.
- [14] 方健,林翔,王红斌,等. 基于设备实时故障率的配电网停电概率计算方法 [J]. 电测与仪表, 2023, 60 (1): 116-123.
- [15] 孙莉,李静,李继云,等. 基于稀疏贝叶斯极限学习机的光伏电站设备故障诊断研究 [J]. 太阳能学报, 2020, 41 (8): 221-226.

(下转第 48 页)