

电气设备监测中大数据分析技术的优化策略

龙 毓

(成都市技师学院(成都工贸职业技术学院), 四川 成都 610000)

摘 要 本文分析了电气设备监测大数据应用中的现状与挑战,重点探讨了数据采集渠道的分散性与整合难题、数据质量把控的不足以及数据分析技术的滞后性,阐述了数据采集、存储、处理与分析的核心技术,包括无线网络、分布式存储、深度学习算法等,提出了完善数据采集与存储体系、强化数据清洗与预处理、构建实时监测与预警智能系统等优化实施策略,强调了大数据技术在故障预判、资源调配和运维决策中的重要作用,以期提升电气设备的管理效率与安全性提供参考。

关键词 电气设备监测; 大数据技术; 数据采集; 故障预判; 智能运维

中图分类号: TM7; TP3

文献标志码: A

DOI: 10.3969/j.issn.2097-3365.2025.04.003

0 引言

随着电力系统和工业设备的复杂性不断增加,电气设备的实时监测与故障预警变得尤为重要。大数据技术的应用为电气设备监测提供了新的解决方案,通过多维度数据采集、存储和分析,实现了对设备状态的精准掌握。然而,数据采集渠道的分散性、数据质量的控制以及分析技术的滞后性仍是目前应用中的主要挑战。本文将探讨电气设备监测大数据应用中的现状、关键技术及优化策略,以期提升设备管理效率和安全性提供理论支持与实践指导。

1 电气设备监测大数据应用现状剖析

1.1 数据采集渠道与整合难题

各类电气设备的监测数据来源繁杂多样,部分源于设备内置传感器,部分则依靠外部加装的监测装置获取,这些数据在传输协定、数据格式以及采集频率层面均未形成统一规范。以电力系统中的不同设备为例,电力变压器的监测数据可能采用独特的格式与传输路径,而与之相关联的配电箱监测数据所遵循的标准却大相径庭,致使数据整合时难以实现无缝对接与融合^[1]。另外,一些陈旧的电气设备未预留适配现代大数据采集系统的数据采集端口,或者其端口与现有系统存在兼容性问题,造成部分设备数据采集缺失,使得数据完整性受损,进而干扰对整个电气系统运行状况的全面评判与精确剖析。

1.2 数据质量把控与校验缺失

在电气设备监测过程中,传感器作为数据采集的关键部件,极易受到多种不利因素的干扰。其自身精度偏差、老化问题以及复杂恶劣的环境因素,如高温、

强电磁干扰等,常常致使所采集的数据出现噪声、异常值与数据漂移现象。例如,在高温且强电磁干扰的环境下,电气设备的电流互感器所监测的数据就容易产生偏差。并且在数据传输阶段,信号衰减与电磁干扰等因素也不容忽视,可能导致数据丢失或错误的情况发生。

当前的数据处理流程在面对这类数据质量问题时显得力不从心。现有的技术手段缺乏高效的自动检测、甄别与校正机制,大多只能依靠人工经验来进行甄别。这种人工方式效率极为低下,而且误差风险很高,不仅会严重削弱基于大数据分析结果的可靠性,还极有可能造成对电气设备运行状态的错误判断,最终对设备的正常维护以及安全运行决策产生极为不利的影 响,可能导致维护不及时或错误维护,增加设备故障风险与运行成本,影响整个电气系统的稳定性与安全性。

1.3 数据分析技术适配性滞后

当下电气设备运行数据具备多变量、非线性、强耦合及时序性等复杂特征,而现有的数据分析手段难以充分契合。传统的统计分析方法在提取电气设备故障特征时无法精准捕捉不同运行参数间的动态关联,难以构建精准的故障预测模型。新兴的深度学习算法虽在理论层面展现出强大的特征学习能力,但在电气设备监测范畴的应用尚处于探索阶段,面临数据标注困难、模型训练复杂以及可解释性欠佳等难题。此外,针对发电设备、输电设备、配电设备等不同类型电气设备的个性化数据分析技术开发不足,缺乏针对性的分析模型与算法,导致在实际应用中难以充分挖掘大数据对各类电气设备监测与管理的价值,无法满足多

样化的电气设备运维需求。同时，不同工况与运行环境下的数据差异未被充分考量，致使数据分析结果的适应性受限，难以有效指导实际运维操作。

2 电气设备监测关键大数据技术解析

2.1 数据采集技术要点

数据采集技术是电气设备监测大数据应用的起始点，其核心在于精准、全面且高效地获取电气设备运行数据。多种传感器被广泛应用，如用于电流测量的电流互感器，基于电磁感应定律 $e = -M \frac{di}{dt}$ （其中 e 为感应电动势， M 为互感系数， i 为电流变化量， t 为时间变化量），将大电流转换为可处理的小电流信号，以便数据采集设备读取。对于电压监测，电压互感器起着关键作用。在传感器布局上，需综合考量电气设备的复杂结构与运行工况，例如在大型电机的监测中，要在定子绕组、转子轴等关键部位安置温度传感器、振动传感器等，以获取多维度数据。此外，无线传感网络技术（WSN）也逐渐兴起，其节点由传感器、微处理器和无线通信模块组成，通过自组织网络将采集到的数据传输至汇聚节点，实现数据的初步整合与远程传输，有效提高了数据采集的灵活性与扩展性，为后续的大数据分析奠定坚实的数据基础。

2.2 数据存储技术原理

分布式存储架构是主流选择，以 Ceph 存储系统为例，它采用基于对象的存储方式，将数据存储在多个存储节点上，并通过 CRUSH 算法进行数据分布，确保数据的均匀存储与高可靠性。其数据冗余策略可有效应对节点故障，保障数据的完整性。在数据存储格式方面，Avro 格式较为常用，它具有良好的跨语言兼容性与高效的序列化性能。Avro 模式定义数据结构，数据以二进制形式存储，减少了存储空间与数据传输开销^[2]。同时，为了满足快速查询与分析需求，基于列存储的存储引擎如 ClickHouse 被广泛应用，它按列存储数据，在数据读取时仅需加载所需列数据，相较于传统行存储方式，大幅提高了数据查询效率，尤其是在大规模数据统计分析场景下，能够快速响应数据检索请求，提升整个电气设备监测大数据系统的运行效率。

2.3 数据处理与分析技术核心

深度学习算法展现出强大的能力。如循环神经网络（RNN）及其变体长短期记忆网络（LSTM），在处理电气设备运行参数的时间序列数据时表现卓越。LSTM 通过独特的门结构（输入门 $i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$ ，遗忘门 $f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$ ，输出门 $o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$ ，其中 W 为权重矩阵， h_{t-1} 为上一时刻隐藏状态， x_t 为当前输入， b_o 为偏置项， σ 为 sigmoid 函数）来控制信息

的传递与保存，有效捕捉数据中的长期依赖关系，可用于预测电气设备的性能衰退趋势。另外，聚类分析算法，如 K-Means 算法（通过不断迭代更新聚类中心 $u_j = \frac{1}{|C_j|} \sum_{x_i \in C_j} x_i$ ，其中 μ 为第 j 个聚类中心， C_j 为第 j 个聚类， x_i 为数据点）可对电气设备的运行状态进行分类，识别出正常运行、潜在故障等不同状态集群，为精准的设备维护与故障预警提供有力支持，从而提升电气设备运行的可靠性与安全性。

3 电气设备监测大数据优化实施策略

3.1 科学规划传感器的选型与布局

在数据采集环节，应依据电气设备的类型、运行特性及关键监测点，科学规划传感器的选型与布局。对于高压电气设备，选用高耐压、高精度且具备抗电磁干扰能力的传感器，如采用罗氏线圈电流互感器测量大电流，其具有良好的线性度和宽频特性。在布局上，围绕设备的进线端、出线端、绕组等核心部位设置传感器阵列，确保全面捕捉电流、电压、温度、湿度等多参数数据。同时，构建数据采集网络，采用工业以太网、现场总线等技术，实现传感器与数据采集终端的稳定连接，遵循特定的通信协议，如 Modbus、Profinet 等进行数据传输，保证数据传输的高效性与准确性。在存储方面，选用合适的分布式存储架构，如基于 Hadoop 分布式文件系统（HDFS）构建存储集群。HDFS 采用分块存储机制，将大文件分割成多个数据块，存储在不同的节点上，通过冗余备份策略保障数据的可靠性。设置合适的数据副本数量，一般为 3 份，分别存储在不同机架的节点上，以应对节点故障。同时，对存储的数据进行分类管理，按照电气设备类型、监测时间等维度建立目录结构，便于数据的检索与调用。定期对存储系统进行性能监测与优化，调整数据块大小、节点负载均衡等参数，确保存储系统能够高效处理海量电气设备监测数据。

3.2 强化数据清洗与预处理流程

建立数据清洗规则库，针对电气设备监测数据中常见的噪声、异常值等问题制定处理规则。对于噪声数据，采用滤波算法，如基于小波变换的去噪方法，通过对信号进行多尺度分解，将高频噪声分量滤除，保留有用的信号成分。对于异常值，利用统计学方法，如 3σ 原则，计算数据的均值和标准差，将超出均值 ± 3 倍标准差范围的数据判定为异常值，然后根据数据的历史趋势、相邻数据关系等采用插值法或均值替代法进行修正。在数据预处理阶段，进行数据标准化处理^[3]。对于不同量程、不同物理量纲的数据，如电流

以安培为单位, 温度以摄氏度为单位, 采用归一化或标准化变换公式, 将数据映射到特定的区间或使其服从标准正态分布, 消除数据之间量纲的影响, 便于后续的数据分析与模型构建。同时, 对数据进行特征提取与选择, 运用主成分分析 (PCA) 等方法^[4]。

3.3 构建实时监测与预警智能系统

构建实时监测系统时, 集成各类传感器数据采集接口与数据处理模块, 采用流式数据处理技术, 如 Apache Flink, 对实时传入的电气设备监测数据进行快速处理。设置数据缓冲区, 当数据流量过大时, 将数据暂存于缓冲区, 按照先进先出的原则进行处理, 避免数据丢失或处理延迟。在预警系统方面, 建立基于机器学习算法的异常检测模型。采用孤立森林算法, 通过构建随机树森林, 计算每个数据点在森林中的路径长度, 将路径长度较短的数据点判定为异常点, 从而识别电气设备运行中的异常状态。根据不同电气设备的运行特性与安全阈值, 设定个性化的预警阈值, 如变压器油温超过特定温度、断路器动作次数超过设定次数等。当监测数据触发预警阈值时, 通过多种通信方式, 如短信、邮件、声光报警等及时通知运维人员, 并提供详细的异常数据报告与可能的故障原因分析。

4 电气设备监测大数据技术成效作用

4.1 精准故障预判与预防

在电气设备监测领域, 机器学习算法尤其是随机森林算法发挥着极为关键的作用。它能够对海量的历史数据以及实时产生的数据展开深度训练学习, 从而精准地识别出电气设备故障的早期微妙特征。以电力电缆监测为例, 通过对电缆的温度数据进行持续追踪与分析, 可洞察其是否存在异常升温趋势; 对局部放电量的监测, 能及时发现绝缘层内部可能出现的微小破损; 而负载电流的波动情况, 则能反映出电缆的承载状态以及是否有接触不良等问题。一旦这些参数呈现出异常变化趋势, 便意味着可能存在绝缘老化、接头松动等潜在故障隐患。基于精准的预判结果, 运维团队能够提前布局, 安排详尽的计划性维护工作, 合理地调整设备运行参数, 从源头上有效规避故障的突发, 极大程度地减少因突发故障引发的设备严重损坏以及长时间停机所带来的巨大损失, 为电气设备的持续、稳定、高效运行筑牢坚实防线, 有力保障电力系统的安全可靠供电^[5]。

4.2 优化设备资源调配

通过对多台电气设备的运行数据进行关联分析, 如不同变电站内变压器的负载率、功率因数等数据对

比分析, 能精准掌握各设备的运行状态与负载能力。基于此, 可在电网负荷高峰时合理调配电力资源, 将部分负载转移至运行状态良好且负载较轻的设备上, 提高整个电气系统的资源利用率; 同时也能根据设备的实际运行需求, 科学安排设备的检修与维护时间, 避免因过度检修或检修不足造成的资源浪费, 从而降低设备运维成本, 提升电气系统的整体运行效率。

4.3 助力运维决策科学化

大数据技术通过对电气设备全生命周期的数据收集与分析, 包括设备采购、安装调试、运行维护、更新改造等各个阶段的数据, 建立起全面的设备信息库。运用数据挖掘技术, 如聚类分析, 对设备故障类型、故障频率、维修成本等数据进行分类处理, 清晰呈现设备的运维状况与风险程度。运维人员可据此制定个性化的运维策略, 对于高风险设备增加巡检频次、提前储备关键备件; 对于低风险设备适当延长维护周期, 实现运维资源的精准投入, 提高运维决策的科学性与合理性, 确保电气设备在高效、安全的状态下运行。

5 结束语

大数据分析技术在电气设备监测领域已然崭露头角并收获了初步成效, 经过全面且深入地剖析其应用现状, 细致入微地解析核心关键技术, 精心制定切实可行的优化策略, 使它在故障诊断方面得以精准定位隐患, 于成本控制方面实现资源的高效调配, 在安全保障上为电气设备构筑起坚固防线, 作用斐然。但不可忽视的是, 数据复杂性等诸多挑战依旧横亘在前, 亟待持续不懈地探索创新, 如此方能深度挖掘并释放大数据于电气设备监测中的无穷潜力, 为电气系统的持久稳定运行奠定坚实的基础, 确保电力供应的安全、可靠与高效。

参考文献:

- [1] 武阳, 王倩. 基于大数据分析的变电设备健康监测与预测模型研究 [J]. 自动化应用, 2024, 65(12): 212-214.
- [2] 陈欣. 基于物联网技术的电气自动化设备监测与故障诊断 [C]// 中国智慧工程研究会. 2024 智慧施工与规划设计学术交流会议论文集. 广西正远电力工程建设监理有限责任公司, 2024.
- [3] 钟庆, 陈伟坤, 许中, 等. 设备故障统计数据与电能质量监测数据的关联分析 [J]. 电力电容器与无功补偿, 2016, 37(06): 29-34, 45.
- [4] 吴春锐. 基于大数据分析的电气系统供电负荷预测与调度优化 [J]. 电器工业, 2024(08): 86-90.
- [5] 袁天霖, 田斌, 朱刚, 等. 基于大数据分析的电力设备状态检测技术研究 [J]. 电子设计工程, 2024, 32(08): 106-110.