

# 继电保护设备智能运维数据的一体化挖掘方法

陈 婧

中国能源建设集团新疆电力设计院有限公司 新疆 乌鲁木齐 830000

**摘要：**本文围绕继电保护设备智能运维数据一体化挖掘展开研究，先阐述运维数据类型与特征、一体化挖掘需求及原则，再详解数据清洗、整合标准化、降维筛选的预处理方法。核心挖掘部分涵盖多数据关联分析、设备状态评估预测、故障诊断预警及多模态数据融合挖掘，还提出技术架构、模型性能、人员能力三方面保障措施。研究为继电保护设备智能运维提供完整数据挖掘方案，助力提升设备运维效率与电力系统安全稳定性。

**关键词：**继电保护设备；智能运维数据；一体化挖掘；多模态融合；保障措施

引言：在电力系统中，继电保护设备至关重要。随着智能电网发展，其产生的运维数据海量且复杂。传统分析方法难以全面挖掘数据价值，无法满足对设备状态精准评估、故障提前预警等需求。一体化挖掘方法能整合多源数据，深入分析内在关联，为运维决策提供有力支持，对保障电力系统安全稳定运行意义重大。

## 1 继电保护设备智能运维数据相关基础认知

### 1.1 运维数据的类型与特征

继电保护设备智能运维涉及多类核心数据。设备实时状态数据涵盖电流、电压、温度等参数，这类数据需持续采集以反映设备运行动态；历史故障记录数据包含故障类型、发生时间、处理结果等信息，是追溯故障规律的关键；环境影响数据则包括湿度、振动、电磁干扰等内容，直接关联设备运行稳定性。各类数据呈现鲜明特征<sup>[1]</sup>。时间连续性体现在实时状态数据需按固定间隔采集，形成连续时序序列；多源异构性表现为不同数据来自传感器、故障录波器、环境监测设备等不同源头，格式与维度存在差异；时空关联性指设备状态数据与环境数据在时间维度同步变化，且不同设备的状态数据存在空间上的相互影响。这些特征直接影响挖掘方法选择，例如多源异构性要求挖掘方法具备数据格式适配能力，时间连续性则需挖掘方法支持时序数据的分析处理。

### 1.2 一体化挖掘的核心需求

一体化挖掘对数据整合有明确需求。不同数据源常因采集设备、存储方式差异形成格式壁垒与存储孤岛，需通过统一的数据接入标准与存储架构，实现实时状态、历史故障、环境影响等多类型数据的集中汇聚，避免数据分散导致的挖掘片面性。对挖掘深度的需求体现在需突破单一数据表层分析的局限。传统分析多聚焦单一数据的数值变化，一体化挖掘需深入挖掘不同数据间的内在关联，例如通过设备电流变化与环境湿度数据的

关联分析，探究环境因素对设备故障的影响机制。对实时性与前瞻性的需求同样关键。实时性要求挖掘方法快速处理新增的实时数据，及时识别设备运行异常；前瞻性则需基于历史与实时数据的融合挖掘，预判设备性能劣化趋势与潜在故障，为运维人员提前制定应对策略提供数据支撑，避免故障突发性导致的系统风险。

### 1.3 一体化挖掘的基本原则

数据兼容性原则要求挖掘方法具备灵活适配能力。面对多源异构数据，需通过格式转换、维度统一等技术手段，确保方法可处理不同类型、不同格式的数据，避免因数据不兼容导致挖掘中断或结果偏差。关联性原则强调挖掘过程中对数据关联的重视。需通过时序匹配、参数映射等方式，挖掘设备状态、历史故障、环境数据间的潜在联系，而非孤立分析单一数据，确保挖掘结果能全面反映设备运行规律。实用性原则聚焦挖掘结果的运维价值。挖掘输出需转化为具体的运维参考信息，例如明确的故障风险等级、可操作的参数调整建议，而非抽象的数据模型或分析结论，确保能直接服务于运维决策。安全性原则贯穿挖掘全流程。需通过数据加密传输、访问权限管控、挖掘过程数据脱敏等措施，防止电力系统核心运维数据泄露或被篡改，保障电力系统运行安全。

## 2 继电保护设备智能运维数据预处理方法

### 2.1 数据清洗技术

运维数据中常存在噪声、缺失值、异常值，需针对性采用处理技术。针对传感器采集的实时状态数据（如电流、电压）中混杂的随机噪声，可采用平滑滤波技术，通过滑动窗口平均或小波变换等方式，削弱高频干扰信号，保留数据真实变化趋势，避免噪声导致的挖掘误判。面对数据采集中断或设备故障引发的缺失值，若缺失数据量较小且时序连续，可采用线性插值或邻近值

填充法；若缺失数据与其他参数存在关联（如温度数据缺失可关联环境湿度），则通过基于数据关联性的填充法，利用相关参数推算缺失值，确保数据序列完整性。异常值识别需结合继电保护设备运行阈值与统计分析，先依据设备额定参数设定合理范围（如电流超出额定值一定范围判定为异常），再通过标准差或箱线图等统计方法筛选偏离正常分布的数据，确认异常后结合历史相似工况数据修正，或剔除无参考价值的极端异常值，保障数据质量满足后续挖掘需求。

## 2.2 数据整合与标准化

多源异构数据的整合需建立统一数据模型，针对结构化的故障记录（如表格形式的故障类型、时间）与非结构化的设备日志（如文本形式的检修记录），通过数据格式转换将非结构化数据转化为结构化格式，再按设备编号、时间戳等关键字段，将实时状态、历史故障、环境数据等多类型数据关联存储，打破数据孤岛<sup>[2]</sup>。数据标准化需消除维度差异影响，对不同量级的状态数据（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算法，将数据映射至统一数值区间，避免因量级差异导致挖掘模型偏向某类数据。对时间序列数据，需以统一时间间隔进行时间对齐，例如将不同采集频率的电压数据、环境数据均调整为每分钟一条记录，确保数据在时间维度上同步，为后续关联挖掘奠定基础。

## 2.3 数据降维与特征筛选

高维度运维数据（如包含数十项参数的设备状态数据）需通过降维减少冗余，主成分分析是常用方法，通过线性变换将多个相关维度转化为少数不相关的主成分，保留数据核心信息，例如将电压波动、电流变化、功率参数等维度合并为反映设备电能状态的主成分，降低数据处理复杂度。特征筛选需结合继电保护设备运行机理，优先保留与设备状态、故障风险强相关的特征变量，例如设备绝缘性能相关的参数、与短路故障关联的电流突变特征等。通过相关性分析或互信息计算，剔除与运维目标无关的特征（如与设备故障无关联的环境光照数据），减少无关信息对挖掘模型的干扰，既提升后续挖掘效率，又增强挖掘结果的准确性。

## 3 继电保护设备智能运维数据一体化挖掘核心方法

### 3.1 多数据关联分析方法

基于关联规则的挖掘方法通过分析海量运维数据，挖掘不同数据间的潜在关联关系。例如在处理环境数据与设备故障数据时，可通过频繁项集挖掘算法，识别环境湿度超出特定范围与设备绝缘故障发生频率间的关联

模式，明确环境因素对设备故障的影响程度，为针对性防护提供依据。基于图神经网络的关联挖掘方法需先构建数据关联图谱，将设备状态参数、环境因素、故障类型作为图谱节点，以数据间的关联强度作为边的权重。通过图神经网络的节点嵌入与消息传递机制，捕捉节点间的复杂关联，例如直观呈现“温度升高-电阻增大-绝缘故障”的链式关联，清晰追溯故障发生的源头因素，帮助运维人员快速定位影响设备运行的关键变量。

### 3.2 设备状态评估与趋势预测方法

基于时序分析的状态评估方法聚焦历史状态数据的变化规律，通过提取数据的时序特征（如均值、波动幅度、周期性变化），建立设备正常运行状态的特征模型。将实时采集的状态数据与模型对比，计算偏差程度，以此评估设备当前健康状态，区分正常波动与异常变化，确保状态评估的准确性。基于机器学习的趋势预测方法中，长短期记忆网络（LSTM）因能捕捉时序数据的长期依赖关系，适用于设备状态预测。通过将历史状态数据（如每日电流峰值、温度变化曲线）输入LSTM模型，训练模型学习状态变化规律，进而预测未来一段时间内设备状态参数的变化趋势。若预测结果显示参数将超出正常范围，可提前识别潜在异常，为预防性运维提供时间窗口。

### 3.3 故障诊断与预警方法

基于故障特征匹配的诊断方法需先挖掘历史故障数据，提取不同故障类型对应的关键特征（如短路故障时的电流突变幅度、接地故障时的电压异常模式），构建故障特征库。当设备出现异常时，从实时数据中提取特征，与特征库中的故障特征进行相似度匹配，快速确定故障类型，减少故障排查时间。基于深度学习的预警方法中，卷积神经网络（CNN）可有效提取运维数据中的故障前兆特征。通过将多维运维数据（如电流、电压、温度的时序序列）转化为数据矩阵，输入CNN模型进行特征提取，捕捉易被忽略的微弱故障信号（如绝缘性能劣化初期的细微电流变化）。结合设备运行阈值，当提取的特征超出安全范围时触发预警，为运维人员争取故障处置时间，避免故障扩大影响。

### 3.4 多模态数据融合挖掘方法

继电保护设备运维数据包含数值型（如电流、电压）、文本型（如设备检修日志、故障报告）、波形型（如故障录波波形）等多模态数据，需通过融合挖掘实现信息互补。针对文本型数据，采用自然语言处理技术（如词嵌入、文本分类）提取关键信息（如检修中发现的设备缺陷类型、处理措施），将非结构化文本转化为

结构化特征向量；针对波形型数据，通过小波变换、傅里叶分析提取波形的频率、幅值等特征参数，形成波形特征矩阵<sup>[1]</sup>。构建多模态数据融合模型，将数值特征、文本特征、波形特征按权重融合为统一特征空间，权重分配依据各模态数据与运维目标的关联强度确定（如故障诊断中波形特征权重高于文本特征）。融合后的数据输入集成学习模型（如随机森林、梯度提升树），结合各模态数据的优势提升挖掘精度，例如通过数值数据反映实时状态、文本数据补充历史检修信息、波形数据捕捉故障瞬时特征，实现对设备运行状态更全面的判定与故障更精准的诊断，避免单一模态数据信息片面导致的挖掘偏差。

#### 4 继电保护设备智能运维数据一体化挖掘的保障

##### 4.1 技术架构保障

支撑一体化挖掘的技术架构需兼顾数据处理效率与安全性。分布式数据存储与计算架构通过将大规模运维数据分散存储于多个节点，结合并行计算技术，可同时处理实时状态数据、历史故障数据等多类型数据，避免单一节点负载过高导致的处理延迟，满足挖掘过程中对数据读取与计算速度的需求。数据安全需贯穿挖掘全流程。数据传输阶段采用专用加密协议，对传输的运维数据进行加密处理，防止数据在传输过程中被截取；数据存储阶段实施访问控制，按角色划分数据访问权限，仅允许授权人员查看或操作核心运维数据；挖掘过程中采用数据脱敏技术，对涉及电力系统隐私的关键信息（如设备核心参数、运维调度记录）进行处理，避免数据泄露或被非法篡改，保障电力系统运行安全。

##### 4.2 模型性能保障

挖掘模型的验证与优化需依托模拟运维场景。构建涵盖正常运行、轻微异常、故障状态等多场景的模拟环境，将历史运维数据按场景分类输入模型，对比模型输出结果与实际场景的匹配度，验证模型在不同工况下的准确性与稳定性。若模型在某类场景中误差较大，需调整模型结构或优化算法参数，提升模型适配能力。模型的实时更新机制需定期执行。按固定周期收集新产生的运维数据与故障案例，将其整合为模型更新数据集，重

新训练模型以更新参数；同时补充新故障类型对应的特征至特征库，确保模型能识别新型故障。对于数据变化频繁的场景（如设备检修后参数调整期），可缩短更新周期，避免模型因数据滞后出现性能衰减，始终保持良好的挖掘效果。

##### 4.3 人员能力保障

针对运维人员的技术培训需分层设计。基础培训聚焦一体化挖掘方法的原理，帮助运维人员理解关联分析、趋势预测等方法的核心逻辑；实操培训涵盖挖掘工具的操作流程，包括数据导入、模型调用、结果查看等步骤，确保运维人员能独立完成基础挖掘操作；进阶培训侧重挖掘结果的解读与应用，指导运维人员根据故障预警结果制定运维方案，将挖掘成果转化为实际运维行动。跨领域协作机制需建立常态化沟通渠道。定期组织数据挖掘技术人员与电力运维专业人员开展研讨，技术人员可了解运维场景中的实际需求（如特定故障的诊断优先级），优化挖掘方法；运维人员可反馈模型应用中的问题（如某类预警误判率高），帮助技术人员调整模型。通过协同配合，确保挖掘方法贴合继电保护设备运维的实际需求，避免技术与业务脱节。

##### 结束语

继电保护设备智能运维数据一体化挖掘，是提升电力系统运维水平的关键。通过有效的数据预处理、核心挖掘方法应用以及多方面保障措施，能充分挖掘数据价值，实现设备状态精准评估、故障提前预警。未来，需持续优化挖掘方法与技术，加强人员培训与跨领域协作，以更好地适应电力系统发展需求，保障电网安全稳定运行。

##### 参考文献

- [1]赵超阳,张锋,康郁堃.继电保护设备智能运维数据的一体化挖掘方法[J].微型电脑应用,2023,39(2):127-130.
- [2]钱晓峰.发电厂继电保护运维防误技术研究及应用[J].自动化应用,2021(01):93-94+98.
- [3]朱忠义.发电厂继电保护运维防误技术研究及应用探讨[J].电子测试,2021(23):120-121+111.