

# 计算机大数据助力水电厂监控自动化故障预警研究

潘 亮

五凌电力有限公司凌津滩水电厂 湖南 常德 415720

**【摘 要】**：水电厂传统的依靠固定阈值的监控方式容易造成大量的误报，不能提前发现设备的潜伏性故障隐患，利用大数据全域采集、时序特征挖掘、多模态融合建模等手段，层层分解故障预警底层运行机理，构建起“边缘—厂站—云端”三级协同诊断架构，配套数据治理、自适应算法、分级闭环管控成套落地流程，经过水电机组、主变、排水泵等多设备工况实试验证，形成适合水电场景、支持运维向预测性状态检修转型的数字化复用技术范式。

**【关键词】**：水电厂监控自动化；计算机大数据；多源数据融合

DOI:10.12417/2705-0998.26.11.038

在传统的水电厂电气自动化监控模式中，机械式阈值判别手段已不能满足现代水电工程高效稳定运行需求。传感采集、PLC 测控、运维台账多类数据存在孤岛问题，难以挖掘设备多参量耦合劣化规律。依托多座大中型水电厂智能化升级实践，可从应用机理、分层架构、落地实施策略三个维度搭建完整的电气设备数字化预控体系，解决传统监控预警滞后、故障溯源困难的行业痛点。

## 1 大数据赋能水电厂故障预警核心机理

### 1.1 多源异构数据全域采集机理

水电厂监控自动化原始数据具有多模态、异质性、时序错位等特点，数据来源可以分为电气运行类、设备传感类、运维业务类三大类。电气类数据用 PLC 测控分站实时采集电压、电流、功率、绕组温度等结构化时序信号，采集周期为 100ms 到 1min。传感类包含振动、局放超声、油溶解气体、水位等非均衡物理量，不同的传感器采样频率相差很大。运维类属于半结构化文本数据，有检修工单、缺陷记录、试验报告等，时间维度是离散的、无序的。

大数据采集机理的核心就是统一时空基准来完成数据的对齐，使用 MQTT 工业传输协议建立多终端的数据接入通道，用统一的时间戳作为关联锚点，对不同采样频次的数据进行插值补全和时序配准处理，消除由于采集时差造成的数据断层，给后续特征提取提供完整的基底。

### 1.2 时序数据特征深度挖掘机理

水电设备故障演化是连续时序的过程，劣化征兆隐藏在参数的长期波动趋势里，传统的单点阈值判别只能发现瞬时参数越限，不能捕捉到设备长期缓慢漂移所造成的隐性劣化征兆<sup>[1]</sup>。大数据时序挖掘机理利用滑动窗口分段解析时序序列，用 CEEMDAN 分解算法去除水力脉动、电网负荷波动等无干扰

分量，得到趋势项、突变项、周期性特征三种主要指标。

对排水泵、主变等典型设备建立不同的特征集，排水泵采集水井水位、单次运行时长、启停间隔时序特征，主变取油中多组分气体时序变化梯度。对发电机筛选出的轴承振动频谱进行分析，克服了传统分析只看瞬时值而忽略长期演变趋势的不足，把隐性的劣化信号剥离出来识别出来。

### 1.3 多维特征耦合故障辨识机理

水电厂复合型故障比例逐年增加，单个参数异常不能准确判断出故障种类，多物理量耦合关联是故障识别的主要思路。多维特征耦合机理依靠注意力融合网络建立跨模态特征映射关系，自动给温度、振动、电气量、介质化学量这些不同特征赋予权重，定量体现各种参数对于故障的影响程度。

以变压器内部放电故障为例，油中乙炔浓度小幅上升不能单独定性，耦合绕组温升、铁芯接地电流、局放超声多维度特征可以区分内部局部放电和外部负荷波动干扰。模型依靠历史故障样本来训练耦合关系，创建起故障现象和诸多特征一一对应的图谱，摒弃了单测点单独判定的模式，降低了工况变动造成的误判概率。

## 2 大数据驱动监控自动化预警分层架构构建

### 2.1 边缘层：就地数据轻量化预处理模块

边缘网关部署在机组、配电室就地 PLC 控制柜内，负责原始传感数据前置预处理，缩减云端传输带宽，实现毫秒级就地异常初判<sup>[2]</sup>。硬件主控单元采用西门子 S7-1200 系列 PLC，内置嵌入式滤波、数据归一化处理单元，现场采集的电气、传感原始数据就地完成去噪、粗大异常值剔除、浅层特征提取，只上传压缩特征数据包到厂站，原始完整数据本地循环缓存 72h 供故障回溯溯源使用。

边缘端采用轻量化的三层 BP 神经网络模型, 激活函数使用 sigmoid, 隐含层节点设为 100, 用统计学  $3\sigma$  阈值规则做一级故障预警。发生过流、绕组超温等突发性故障时, 就地 PLC 可以直接下发设备联锁保护指令, 同时把初级异常告警推送到运维终端。

## 2.2 厂站层: 多模态数据融合研判模块

在厂站层部署时序数据库以及融合分析引擎, 接收到边缘标准化特征之后再由其对设备之间、不同模态之间进行综合判断, 成为三级架构的中枢核心。在厂站层部署 TDengine 时序数据库来存储中长期设备的特征序列, 创建起双层融合流水线, 浅层融合使用规则引擎做秒级的异常提示<sup>[3]</sup>。中层融合依靠 LSTM-TCN 模型来剖析长时间的劣化趋向, 给出故障出现的概率以及还剩多久可以维持健康的状态。

厂站内嵌入故障特征知识图谱, 将发电机匝间短路、变压器油劣化、泵体磨损等故障模板匹配进去, 自动生成故障定位和处置的可视化看板, 按照隐患轻重分为三级告警, 分层推送给相应的运维终端, 解决海量无关告警掩盖关键隐患的问题。

## 2.3 云端层: 全域模型迭代优化调度模块

云端负责多电站集群的统筹管理, 不会进行实时预警, 主要做全局样本的训练以及跨站知识迁移。云端搭建联邦学习框架, 各水电厂只上传模型梯度, 本地留存不外泄, 本地原始数据也不外泄, 实现了样本复用与数据隐私的平衡。云端生成对抗网络 (GAN) 扩充罕见故障样本, 改善样本不平衡状况, 减少由于样本失衡造成的过拟合, 定时下发优化后的推理模型, 按照季节、服役年限自动更新判定标准, 消除固定阈值不能适应的缺点<sup>[4]</sup>。

# 3 大数据预警体系工程化落地实施路径

## 3.1 标准化数据治理体系搭建举措

数据治理给预警系统提供底层支撑, 覆盖采集传输、多级清洗、分层存储、元数据标准化全工程链路, 加上统一的执行标准来消除异构数据失真、时序错位、存储低效等落地问题。全域采集使用的是双协议统一标准, 厂区 PLC、油气在线监测、红外成像终端等设备都可以支持 MQTT、OPC UA 工业传输协议, 在厂房强电磁干扰的环境下也能实现低时延的交互。统一全站测点编码、物理量计量单位、采样精度标准, 采集链路采用以太网和工业 4G 双冗余通道, 链路切换时间控制在 100 毫秒内, 避免由于单个通道断流造成数据断层<sup>[5]</sup>。

清洗任务分层部署的运算逻辑, 边缘终端用 Z-score 三倍标准差规则就地剔除瞬时粗大跳变数据。厂站层用孤立森林算法来识别长期参数漂移的样本, 增加空载、检修工况标签来隔离无效时序。缺失值分级补全, 短时空白用分段线性插值法, 长时段空缺调取同水文、同负荷历史序列拟合法。系统对于数据的完整性、时序连续性实行 7\*24 小时自动核算, 创建异常

传感、采集分站缺陷台账, 把台账发送给运维岗位, 并且规定整改时限。存储使用湖仓分层结构, 根据访问频率把数据分层存储到不同的介质上, 高频实时特征保存到 TDengine 时序库, 满足模型毫秒级调用需求, 长期历史时序数据存放在压缩 HDFS 集群里。检修工单、试验文本存到 MySQL 数据库中。配套生命周期管控, 实时数据保存 30 天, 年度数据保存 10 年, 机组、主变原始故障样本永久锁定。

## 3.2 自适应智能预警算法适配方案

抛弃全域设备通用单一模型粗放式的布置方式, 根据水电厂一次主设备、二次辅机不同种类的故障演化机理进行分类匹配混合算法组合, 同时建立月度模型自适应迭代运维流程, 用动态残差预警代替传统的静态固定限值, 适应汛期负荷变化、季节温差、设备年久失修导致工况漂移的场景。

分设备算法组合差异化配置细则落实到具体设备上, 对发电机、主变压器、GIS 高压组合电器等中长期渐进劣化类核心设备采用 LSTM-TCN 时序混合推理模型, TCN 模块从振动、油色谱、温度短周期突变特征入手, LSTM 模块从跨月度缓慢参数漂移趋势出发, 两方面特征相结合得出设备劣化概率分值。母线短路、瞬时过负荷、开关瞬时分合闸等短期突发性故障, 用孤立森林无监督异常检测算法和浅层 BP 神经网络, 用单窗口采样的数据做毫秒级的异常识别。渗漏、防洪、检修三类集水井排水泵、冷却风机等辅机设备使用轻量化三层 BP 神经网络架构, 模型输入特征只选取集水井时序水位、设备累计启停运行时长、电机三相电流三个低维监测参数, 在保证故障识别精度的基础上, 大幅度降低厂站服务器的算力消耗<sup>[6]</sup>。对局部放电红外图谱、超声波形等空间特征数据进行轻量化 CNN 模型处理, 得到放电热点区域、脉冲频谱二维特征, 解决深层网络推理延迟问题。

模型自适应迭代设置标准化月度运维流程, 每月从厂站新增标注故障样本中扩充训练数据集, 数据集按照 8:2 的比例随机分成训练子集和验证子集, 使用 Adam 梯度下降优化器, 基础动态学习率固定在 0.001 的范围内, 同时加上 L2 正则约束以及 0.4 概率丢弃层, 双向抑制模型过拟合的现象, 防止少量故障样本训练之后泛化能力下降的问题。迭代结束后自动输出验证集误报率、漏报率、故障识别准确率这三个评价指标, 指标没有达到设定的阈值时就继续增加训练轮次, 直到模型的辨识性能达到要求为止。

动态残差预警判断完全代替人工设置的固定阈值, 模型会自动获取最近 30 天内同一负荷区间、相同环境温湿度、相同水库水位的设备健康运行数据集, 建立不同工况的基准参数区间, 实时计算当前监测参数和基准区间残差数值, 增加多连续窗口校验约束规则, 只有当残差在连续 5 个采样窗口都超过允许范围的时候才会触发预警推送, 单个瞬时参数偏离直接判断

为随机噪声进行过滤,从根本上减少由于季节温差、昼夜负荷波动造成的批量虚警问题。对于投运初期没有历史故障样本的新设备,采用迁移学习适配方案,调取同型号成熟机组预训练权重进行模型初始化,只需要采集15天正常工况数据就可以完成基准区间的拟合,大幅缩短了新设备预警系统的投产适配周期。预留人工样本标注交互入口,运维人员对复合型故障现场消缺处置完毕后,可以将故障完整的时序特征、故障根因判定、处置措施记录一起上传到系统中,系统会把标注好的样本归入训练数据集中,不断丰富故障特征样本库,逐步扩大模型的故障识别范围,适应水电厂逐年增加的未知耦合故障类型。

### 3.3 预警分级处置闭环管控机制

构建“预警识别—分级推送—现场巡检核查—消缺工单归档—模型迭代优化”全流程闭环管控业务链路,打通智能预警分析平台与电厂运维工单管理系统双向数据交互接口,按照故障设备重要程度、停机损失规模划分三级标准化预警管控等级,为不同等级告警匹配差异化推送渠道、现场处置时限、设备联锁保护逻辑,同步构建预警效能月度复盘与模型自动调优子流程,实现预警处置经验反向赋能算法迭代。

三级预警层级划分及标准化处置规则固化落地,一级预警为设备前置潜伏隐患,参数残差微小偏离基准区间,无停机风险,系统只保留全量特征日志,每天汇总生成设备轻微劣化巡检清单推送给班组日常运维终端,不触发实时弹窗、短信告警。二级预警指的是设备性能持续性的劣化趋势,多维特征同时出现偏移的情况,存在短期降出力的风险,系统自动发出计划性检修的电子工单,工单里包含故障的疑似位置、相关的监测参数、以往同类故障的处理参照方案,限定运维班组在24小时内对现场进行核查,核查后的记录和设备复测数据一起上传到平台上存档。三级预警指的是突发的永久性设备故障,存在机组非计划停机、高压设备绝缘击穿的风险,系统同时启动多个渠道的告警推送,中控大屏红色高亮弹窗、运维人员移动端APP实时推送、应急联络短信并行下发,就地PLC同步联动设备联锁保护程序,执行故障区段隔离、备用设备投运、机组

降负荷等自动防护操作,限定值班管理人员4h内到达故障点位开展消缺作业。

系统完整保存单条预警的全部生命周期结构化档案,档案字段包含预警触发准确时间、相关多维特征时序曲线、设备资产编号、预警等级、接收处置人员、现场消缺耗时、故障根源定性、处置后复测参数、备品备件使用记录。每月自动汇总全域预警档案数据,对各级预警误报频次、漏报案例、平均处置时长三个主要的效能指标进行量化统计,生成可视化告警效能分析报表,在报表中标注出高频误报对应的设备、误报触发工况、特征异常类型。对于报表中的高频虚警故障点位,系统会自动调取故障时段的全部原始特征数据集,创建出算法优化任务单推送给技术运维人员,任务单里包含模型特征权重调整的区间以及残差预警阈值的修正范围,参数优化之后再推理模型下发到边缘终端,从而达成预警识别效果不断自我改进的闭环。

增设人机协同复合故障研判交互模块,对发电机定子绝缘耦合劣化、主变油色谱协同超标等多物理场复合型故障,系统自动整合温度、振动、气体、电流多模态特征研判报告,报告从各个角度展示各个特征的故障贡献权重、故障演化预测曲线、同类故障的历史停机损失数据。所有的故障处理完毕之后,系统会自动把工单里的根因标注文本、复测特征数据以及处置流程记录,标准化成带有标签的训练样本存入样本库,不需要人工再做二次的数据整理工作,从而不断充实模型对于复合故障特征的掌握范围,使模型对复杂的耦合故障识别更加准确。

## 4 结论

大数据借助多源数据融合、时序特性解析、多维故障识别冲破水电静态预警的限制,三级分层架构把算力划分到不同的层次上去,既能保证即时反应又能实现整体模型的改良。标准化数据治理、分设备适应算法、闭环管控流程可以减少误报,提高发现隐患的速度。之后可以联动数字孪生、联邦学习技术来提高罕见故障识别的精度,并且可以拓展到机组负荷优化、设备寿命评价等场景的应用上。

## 参考文献:

- [1] 高志东,李琦,梁浩欣,等.水电厂故障智能报警系统设计与研究[J].机电产品开发与创新,2026,39(3):4-7.
- [2] 杨在鑫.人工智能在水电厂电气设备故障诊断中的应用探索[J].信息化研究,2026,52(2):198-202.
- [3] 付万珊,周杰.PLC技术在水电厂电气自动化监控系统中的应用[J].轻工科技,2025,41(4):119-121+129.
- [4] 刘卫国.智能技术在水电厂电气自动化中的应用[J].中国高新科技,2025,(1):128-130.
- [5] 马成伟.智能技术应用在水电厂电气自动化中的研究[J].水上安全,2023,(16):7-9.
- [6] 唐孝舟,孙长兰,张玉彬,等.基于BP神经网络的水电厂排水泵故障预警研究[J].水电与新能源,2022,36(9):22-26.