

基于 DeepSeekR1 与 RPA 双引擎的电力运维智能决策方法及系统实现

刘振屹 罗正 廖培期

广西电网有限责任公司贵港供电局

DOI:10.32629/etd.v6i5.16914

[摘要] 在全球能源革命与我国“双碳”目标驱动下,电力行业向新型电力系统转型过程中,多源异构故障数据的高效处理与动态决策成为关键技术瓶颈。为此,融合深度求索1(DeepSeekR1)模型与机器人流程自动化(Robotic Process Automation,RPA)技术,构建“数据深加工+智能决策大脑”双引擎架构,通过“智能判据+动态过滤”机制与全流程自动化闭环,解决传统模式中信息处理低效、决策支撑碎片化及执行断层问题。该架构设计哈希指纹比对算法(重复数据剔除率98%)、动态时间窗口机制(漏采率 $\leq 0.1\%$)及跨时段故障推演算法,实现多源数据的时空关联分析与智能研判。实验显示,故障研判准确率提升至63%(较传统规则引擎提升18个百分点),决策延迟压缩至1分钟内,为电力运维提供兼具理论创新性与工程实用性的智能决策方法。

[关键词] DeepSeekR1; RPA; 电力运维; 智能决策算法; 双引擎协同; 全流程自动化

中图分类号: TM727 文献标识码: A

Research on Intelligent Decision-Making Method and System Implementation for Power Operation and Maintenance Based on DeepSeekR1 and RPA Dual Engines

Zhenyi Liu Zheng Luo Peiqi Liao

Guigang Power Supply Bureau of Guangxi Power Grid Co., LTD

[Abstract] Amid the global energy revolution and driven by China's "dual-carbon" goals, the power industry's transition to a new-type power system faces a key technical bottleneck: efficient processing of multi-source heterogeneous fault data and dynamic decision-making. To address this, a dual-engine architecture integrating the DeepSeekR1 model and Robotic Process Automation (RPA) technology, namely "data deep processing plant + intelligent decision-making brain", is constructed. Through the "intelligent criterion + dynamic filtering" mechanism and full-process automated closed-loop, it solves problems in traditional modes such as inefficient information processing, fragmented decision support and execution discontinuity. The architecture designs hash fingerprint comparison algorithm (98% duplicate data removal rate), dynamic time window mechanism (missing collection rate $\leq 0.1\%$) and cross-time fault deduction algorithm to realize spatiotemporal correlation analysis and intelligent judgment of multi-source data. Experiments show that fault judgment accuracy increases to 63% (18 percentage points higher than traditional rule engines), and decision delay is compressed to within 1 minute, providing an intelligent decision-making solution with both theoretical innovation and engineering practicability for power operation and maintenance.

[Key words] DeepSeekR1; RPA; power operation and maintenance; intelligent decision-making algorithm; dual-engine collaboration; full-process automation

在全球能源革命与“双碳”目标驱动下,新型电力系统“源网荷储”多要素协同,配网故障数据呈现“多源异构、时空关联、动态演化”特征,传统“人力主导+静态规则”运维模式面临三

重瓶颈:一是多源数据融合时空关联性丢失,SCADA系统非结构化数据关键信息嵌套致80%重复/无效数据难过滤;二是决策模型动态适应性不足,固定规则引擎无法适配电网拓扑实时变化,

连锁故障风险识别滞后; 三是“感知-决策-执行”闭环断裂, 人工介入使关键信息报送延迟率达60%^[1,2,3]。

近年人工智能在电力运维应用有进展: 文献[1]提出多模态AI大模型赋能设备健康评估, 但未解决跨时段数据动态关联; 文献[4]综述AI故障诊断技术, 缺乏与自动化执行协同机制; 文献[9]探索RPA+AI应用, 决策模块依赖静态规则难以应对复杂故障^[5,6,7,8]。现有方法在多源异构数据动态过滤与时空关联分析上存在短板, 且未形成“信息处理-决策支撑-执行闭环”全链条方案, 缺乏配网运维领域实践性成果^[1,7,9]。

本文针对上述问题, 提出基于DeepSeekR1与RPA双引擎的智能决策方法: 通过哈希指纹比对与动态时间窗口实现数据精准过滤, 依托DeepSeekR1知识图谱构建跨时段故障推演模型, 结合RPA机器人集群形成全流程自动化闭环。

1 系统研发架构

在推进机器人流程自动化(Robotic Process Automation, RPA)与DeepSeek R1模型融合的电力运维数字员工系统研发前, 传统运维模式面临信息处理低效、决策支撑碎片化、执行闭环断层三大核心挑战。为此, 文中构建以DeepSeek R1 671B模型为“决策大脑”、RPA技术为“执行手臂”的双引擎架构, 实现配网故障处理全流程智能化升级。

1.1 传统运维模式的三大核心挑战

(1) 数据处理的时空一致性缺失: 现有生产数字员工(PDE)多采用故障信息“直采直发”模式, 极端天气(如台风、暴雨)期间, 80%的重复数据源于缺乏增量识别机制^[4,8]。更为传统的故障信息报送模式则依赖调度员人工分析, 极端天气下人工筛选耗时累计2-4小时, 关键信息报送延迟率最高达60%。难以应对日均14.4万条的数据流^[1,7]。

(2) 决策模型的动态适应性不足: 单一设备的分合闸信息无法构建故障关联图谱, 静态规则引擎对“自动化开关拒动-保护定值冲突”等连锁风险的识别准确率仅45%^[7,9]。人工研判依赖运维经验, 难以识别自动化开关拒动、保护定值冲突等潜在连锁风险, 导致综合运维成本高(人均年成本16万元)、故障信息传递效率低(平均耗时0.5小时/次)。

(3) 执行闭环的自动化断层: 现有相似系统多与AI缺乏联动机制, “感知-思考-执行”链条断裂, “数据采集-分析-报告生成-分发”的6-8个环节依赖人工, 单流程耗时超30分钟, 且15%的格式错误率源于人机交互中的信息失真, 此类“半智能化”(semiintelligent, SI)状态与新型电力系统“全流程智能化”(fullprocessintelligent, FPI)要求存在显著差距, 同时严重制约运维闭环效率^[4,8]。

1.2 双引擎驱动架构设计

架构采用四层分布式设计(图1), 实现“数据清洗-知识推理-决策执行”的全流程智能化^[4,10]。

(1) 数据采集层: 基于SCADA系统实时抓取设备状态数据, 覆盖贵港网区98%的配网设备(10kV线路823条、自动化开关2250台、配变28830台), 日均处理14.4万条多源异构数据。

(2) 数据深加工: RPA构建三级清洗流水线(去重→标准化→关联匹配), 引用哈希指纹比对算法(式(1)、(图2))实现重复数据识别: $H(D_n) = SHA-256(D_n)$

式中: D_n 为第n轮采集数据集, $H(x)$ 为哈希函数, 若 $H(D_n) = H(D_{n-1})$ 则判定为重复数据。

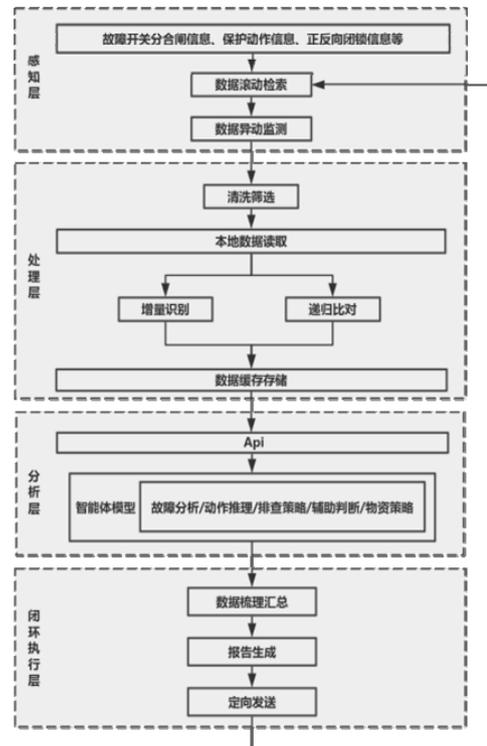


图1 PDE 双引擎架构分层示意图

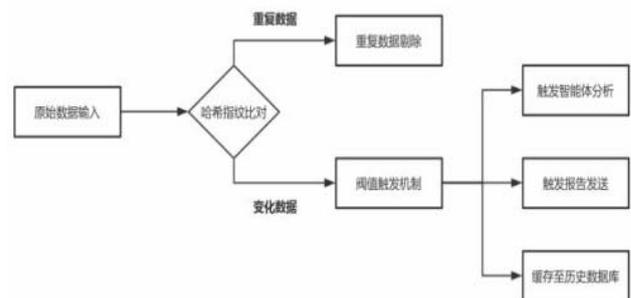


图2 PDE哈希指纹比对与阈值触发机制程序流程图

(3) 智能决策大脑: RPA清洗后的标准化故障信息输入智能体后, 通过“设备-故障-策略”逻辑链实现初步研判, 例如研判多时段自动化开关故障分合闸、保护、正反向闭锁等动作信息, 形成跨时段故障(Cross-temporal-level failure, CF)及风险研判结论。

(4) 自动化执行层: 通过RPA机器人集群执行数据录入、深层处理、报告生成、报告分发, 单流程耗时从人工30分钟压缩至1分钟, 无人化率达100%^[9]。

2 故障信息精准处理关键算法

2.1 多维度数据过滤与动态时间窗口机制

针对信息重复性难题,设计多维度过滤策略:

(1) 增量识别算法:采用哈希指纹比对技术(见式(1)),实现98%的重复数据剔除率,较传统关键词匹配算法(剔除率75%)提升23个百分点^[7]。

(2) 阈值触发机制:设置新增数据量阈值(如单次检索新增信息长度 ≥ 9),如相应变量触发阈值条件,则同步触发下一阶段执行程序启动。

(3) 动态时间窗口算法(式(2)):确保相邻检索周期重叠率 $\geq 99\%$,理想数据源条件下漏采率控制在0.1%以下:

$$t_{start}(n) = t_{end}(n-1), t_{end}(n) = now_{time}$$

式中: $t_{start}(n)$ 为第 n 轮检索起始时间, $t_{end}(n-1)$ 为第 $n-1$ 轮检索结束时间, $t_{end}(n)$ 为第 n 轮检索结束时间,默认将系统实时检索时间赋值给 $t_{end}(n)$ 。

2.2 递归比对算法机制

建立“配变-线路-时间”三维台账(存储格式为JSON),通过递归比对算法(式(3))逐轮校验未复电状态:

(1) 每轮次检索生成两个临时集合:

停电数据集: $A_n = \{line : C_i | S_i(t) = 1, t \in [t_{start}(n), t_{end}(n)]\}$, 复电数据集 $B_n = \{line : C_i | S_i(t) = 0, t \in [t_{start}(n), t_{end}(n)]\}$;

(2) 第一次比对校验: $A_{n2} = A_n - (A_n \cap B_n)$, $B_{n2} = B_n - (A_n \cap B_n)$

(3) 第二次比对校验:

若 $A_{n2} \neq \emptyset$ (仅有新停电集合): $U_{curr} = \{line | U_{prev} \cup A_{n2} | S_i(t) = 1, t > t_{end}(n-1)\}$ (将上轮未复电集合并入本轮未复电集合)

若 $B_{n2} \neq \emptyset$ (仅有新复电集合): $U_{curr} = \{line | U_{prev} \cap B_{n2} | S_i(t) = 0, t > t_{end}(n-1)\}$ (将上轮未复电集合中剔除本轮已复电集合)

式中: $line$ 为线路ID(如10kV贵阳线), C_i 为配变ID(如秦河公变), $S_i(t)$ 为 t 时刻状态(1表示停电,0表示复电), $t_{start}(n)$ 为本轮次检索起始时间, $t_{end}(n)$ 为本轮次检索结束时间, A_{n2} 为新停电集合, B_{n2} 为新复电集合, U_{curr} 为当前轮次未复电集合(存储结构为“线路ID:配变ID”键值对), $t_{end}(n-1)$ 为上一轮检索结束时间(确保仅纳入最新一轮数据周期内的状态更新)。

2.3 责任部门归属映射机制

建立包含5个运维部门、823条线路的双向动态映射库,支持多部门协同(如“10kV跨区1线”同时归属A、B部门),并允许手动更新映射库数据,当匹配到多个部门时,可触发多报告生成及发送机制(式(4))(每份报告标注责任边界),通过Dijkstra算法优化分发路径,确保每轮检索周期均完成多部门推送,分发遗漏率降至0.05%^[7]: $D = \{d_1 : line_{d1}, d_2 : line_{d2}, d_3 : line_{d3}, d_4 : line_{d4}, d_5 : line_{d5}\}$

若 $l \in D(d_i)$ (故障线路属于归属部门集合的映射): $\delta_{l,d_i} = 1$

若 $l \notin D(d_i)$ (故障线路不属于归属部门集合的映射): $\delta_{l,d_i} = 0$

式中: D 为部门映射数据集,其中 d_1 、 d_2 、 d_3 、 d_4 、 d_5 均为不同责任部门, $line_{d1}$ 、 $line_{d2}$ 、 $line_{d3}$ 、 $line_{d4}$ 、 $line_{d5}$ 均为不同

责任部门管辖范围的线路ID数据集, l 为本轮次检索周期的其中一条故障线路ID, δ_{l,d_i} 为表示线路 l 归属于各责任部门的映射值。

3 DeepSeek R1模型构建

3.1 电力运维知识图谱构建

基于DeepSeek R1模型构建故障研判智能体,嵌入保护动作逻辑(如三段式电流保护配合规则)、自动化开关动作策略(如自动化开关故障响应机制)等专业机理规则。初步整合历史故障报告(267份)、保护配置规则(1份)、自动化开关动作逻辑(2份),初步构建可实际应用并具备配网故障领域专业逻辑推理的知识图谱,覆盖“设备故障-保护动作-处理策略”全链条关联^[1]。

3.2 跨时段故障与风险推演算法赋能“长期记忆”

针对传统系统仅能处理单一时段数据的缺陷,PDE设计时间窗滑动机制(窗长48小时)^[6],故障综合分析研判持续提取故障发生后48小时内的开关分合闸动作、保护动作信号、正反向闭锁信号、开关及公变掉线信号等故障关联数据作为输入特征,通过Transformer架构实现跨时段关联分析,较传统时序模型(如LSTM)能成功实时研判出因故障处理过程(如试送)而实时变化的实际故障停电区域范围进而可针对特定故障事件得出更具全面性的故障研判结论。

4 全流程自动化执行体系构建

4.1 RPA机器人集群部署

(1) 数据采集模块:采用动态循环策略(每轮检索周期结束后自动化执行下一轮数据检索),单次采集及汇总耗时从人工10分钟缩短至10秒,准确率100%(对比人工操作失误率5%)。

(2) 报告生成模块:基于“线路-时间”关联表(主键为线路ID),自动生成包含故障动作序列、停电范围的结构化报告(Word格式),生成耗时15秒/份。

(3) 任务分发模块:通过“线路-部门”双向映射库(bidirectional mapping library, BML)(支持多对多关系,含5个运维部门、823条线路的协同映射关系),自动匹配责任部门并推送至企信群组,分发准确率99.9%,人工干预率0%。全流程程序流程图见图3。

4.2 系统集成与交互机制

设计遵循标准化JSON接口(包含至少14个必填字段:运维部门ID、线路ID、公变ID、故障时间戳、周期检索起始时间戳、周期检索截止时间戳等),通过API接口实现RPA与DeepSeek R1的实时交互:RPA将清洗后数据(日均处理量1.2GB)输入模型,模型返回的策略建议(含7类处理意见:故障区间定位、故障类型分析、潜在风险分析、设备动作序列、故障抢修建议、报告置信度判断、未复电设备提醒)经RPA封装后推送至elink,单条故障线路报告全流程延迟 $\leq 1\text{min}$ 。

5 实际应用后的程序短板及优化成果

实际应用非理想性数据源导致的连锁性缺陷问题

5.1 数据源延迟上送(连锁问题1)

(1)问题根源分析: 程序实际应用中,原“检索时间戳强制更新机制(RTFUM)”基于理想数据源(按实际产生时间秒级上送)设计,但实际电力生产环境中,常因自动化开关受损、调度主站端数据堵塞、通信通道中断等导致数据源延迟上送。此时程序仍按固定周期更新“本轮检索截止时间戳”,造成下一轮检索起始时间跳变至未来时刻,形成“检索时间窗口断层”;即便后续数据补传,也会因时间窗口错位导致漏采,即“检索二次真空(RSV)”。

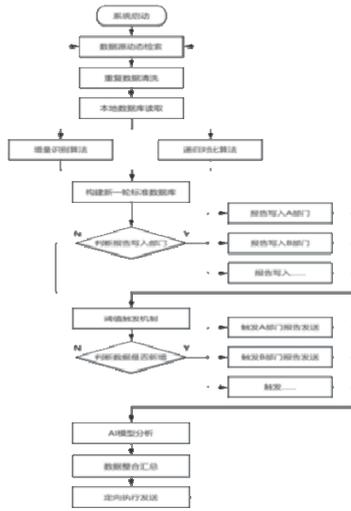


图3 PDE全流程程序流程图

(2)技术优化方案: 引入数据新增状态感知机制(data insertion state awareness mechanism, DISAM) (式(5)), 将检索时间戳的更新与数据实际上送到达强绑定, 有新增数据时, 正常更新检索截止时间戳, 维持时间窗口连续性; 无新增数据时, 冻结检索起始时间, 持续等待数据补传, 通过这种“按需更新”机制, 确保检索时间窗口始终覆盖实际数据产生区间, 彻底消除因数据源延迟导致的检索真空。

数据增量判断: $Pd_{i,t} = 1 | G_{i,t} \geq T, Pd_{i,t} = 0 | G_{i,t} < T$

时间戳条件更新: $F_t = Now_t | Pd_{i,t} = 1, F_t = prev(F(t)) | Pd_{i,t} = 0$, 其中 $prev(F(t))$ 表示保留上一轮存储的时间戳。

跨周期容错检索示例: 设第n轮检索的起始时间为 S_t , 截止时间为 E_t , 第n轮检索的起始时间为 S_{t+1} 。

$$S_{t+1} = E_t | Pd_{i,E_t} = 1, S_{t+1} = S_t | Pd_{i,E_t} = 0$$

即当有数据更新时, 下一轮检索从当前截止时间 E_t 开始; 若无数据更新时, 下一轮检索维持原起始时间 S_t , 直至数据补传。

式中: $Pd_{i,t}$ 为线路 l 在时刻 t 的数据更新标志(0/1), $G_{i,t}$ 为线路 l 在时刻 t 的检索结果集合, T 为数据阈值上限, F_t 为线路 l 的时间戳存储文件, Now_t 为时刻 t 的实时系统时间戳, 该算法优化方案经使用检验可彻底消除因数据源延迟导致的时间窗口断层, 确保每条线路数据按实际产生时序完整采集, 同时仅在数据有效更新时执行存储操作, 减少I/O开销。

5.2 新增非数据库故障线路运行异常(连锁问题2)

(1)问题根源分析: 程序解决连锁问题1、2并投入实际应用后发现, 时间戳存储逻辑依赖键值对存在性校验机制(key-value pair existence verification mechanism, KPEVM)。原程序通过共享字典索引关联线路与时间戳时, 直接调用“line name”键值检查存储文件是否存在, 若键值不存在(如新增非数据库线路), 会触发“File Not Found Error”异常导致程序中断。该机制未考虑电力生产环境中可能出现的新线路故障数据(如临时新增线路、改造后未及时录入数据库的线路), 导致系统鲁棒性(system robustness, SR)不足, 无法自动识别和处理此类异常数据, 可能造成故障信息漏采或流程中断。

(2)技术优化方案: 在(式(5)、式(6))基础上为通过“异常驱动-动态适配机制”(anomaly-driven dynamic adaptation mechanism, ADDAM) (式(7)), 在读取时间戳文件前, 通过异常捕获机制拦截“File Not Found Error”, 识别是否为新增线路, 若判定为新增线路(即键值不存在), 则以本轮检索到的故障最早发生时间作为初始时间戳, 创建专属txt文件存储, 维持程序流程连续性。

程序运行方式判断: $G = \{Read_{f,d} | P_{line} = 1\}, G = \{Write_{f,d} | P_{line} = 0\}$

式中: G 为该程序的运行方式(read/write), $Read_{f,d}$ 为读取故障线路上轮检索截止时间存储文件, $Write_{f,d}$ 为写入新增非数据库故障线路时间存储文件, P_{line} 为文件存在性判断标识(1表示文件存在, 0表示文件不存在), 该优化方案彻底解决新增非数据库线路时导致的程序崩溃问题, 系统对未知线路的兼容性达100%, 实现“异常识别-动态建档-持续跟踪”的全流程自动化运行。

6 总结

文中从电力行业配网运维领域实际存在的数字化需求出发, 提出的DeepSeekR1与RPA双引擎智能决策方法, 通过哈希指纹比对、动态时间窗口等算法解决了多源数据过滤难题, 依托知识图谱与跨时段推演实现了故障智能研判, 结合RPA集群形成全流程自动化闭环。实验表明, 该方法在故障研判准确率、决策效率等指标上优于传统方法, 其核心算法可推广至“源网荷储”协同等场景。

未来系统优化研究将聚焦于以下方向: (1)融合数字孪生技术构建三维故障推演模型; (2)扩大知识图谱规模(目标1.2万份样本), 将研判准确率提升至80%; (3)设计RPA应急子程序, 提升系统在AI接口中断时的容错能力。

[参考文献]

[1]CHEN Xiaohong,FU Wenrun,LIU Zhaoming,et al.Discussion on applications of large-scale artificial intelligence models in power equipment operation and maintenance scenarios[J]. Chinese Engineering Science,2025,27(01):180-192.

[2]ZHANG Jun, XU Jian, XU Peidong,et al.Application of large-scale artificial intelligence models in power equipme

nt operation and maintenance scenarios[J].Engineering Journal of Wuhan University,2025,27(01):180-192.

[3]CHEN Xiaohong, FU Wenrun, LIU Zhaoming, et al. Review and prospect of large-scale artificial intelligence models in power system operation and control[J]. Chinese Engineering Science,2023,56(11):1368-1379.

[4]SUN Qiuye, YANG Lingxiao, ZHANG Huaguang. Smart energy: applications and prospects of artificial intelligence in power systems[J].Control and Decision,2018,33(05):938-949.

[5]TANG Wenhui, NIU Zhewen, ZHAO Baining, et al. Research and application of data-driven artificial intelligence in power equipment condition analysis[J]. High Voltage Engineering, 2020,46(09):2985-2999.

[6]PU Tianjiao, ZHAO Qi, WANG Xinying. Research framework, application status, and prospects of power artificial intelligence technology[J].Power System Technology,2024:1-21.

[7]LI Peng, HUANG Wenqi, WANG Xin, et al. Review of data-and-knowledge-driven artificial intelligence methods in power dispatch[J]. Automation of Electric Power Systems, 2024,

48(01):160-175.

[8]ZHAO Jinquan, XIA Xue, XU Chunlei, et al. Review of next-generation artificial intelligence technologies in power system dispatch[J].Automation of Electric Power Systems, 2020,44(24):1-10.

[9]LIU Haitao, DUAN Jing, WANG Yanhua, et al.. Application analysis and architecture design of RPA+AI-based digital employees in power industry[J].Electric Power Information and Communication Technology, 2022, 20(04):88-93.

[10]YANG Ting, GENG Yinan, GUO Jinghong, et al. Applications of artificial intelligence in smart sensing, communication, and data processing for new power systems[J]. High Voltage Engineering,2024,50(01):19-29.

作者简介:

刘振屹(1996--),男,壮族,广西玉林人,工学学士,广西电网有限责任公司贵港供电局,工程师,主要研究方向为人工智能、大数据分析、电力调度及生产智能化、流程自动化、数据可视化开发。