

# 基于海量数据的配电网运行分析系统架构与技术实现

张 稳<sup>1,2</sup>, 盛万兴<sup>2</sup>, 杜松怀<sup>1</sup>, 贾东梨<sup>2</sup>, 康田园<sup>2</sup>

(1. 中国农业大学信息与电气工程学院, 北京市 100083; 2. 中国电力科学研究院有限公司, 北京市 100192)

**摘要:** 灵活地分析和运用配电网多源异构数据, 进行配电网故障的准确预测和潜在风险扫描, 对实现配电网高效精准的运行分析决策很有意义。文中介绍了一种基于海量数据的配电网运行分析系统的整体架构和功能设计, 通过应用实例展示了系统各模块的功能和运行分析结果。该系统集成电网地理信息、营销业务应用和配电自动化等系统数据, 采用改进的机器学习算法和薄弱点辨识方法, 扩展了目标配电网数据相关性分析、故障风险等级预测和薄弱点辨识的功能, 有利于相关部门有针对性地提出相应技术和管理手段进行配电网运维检修工作, 提高了现有配电网分析系统的科学性和实用性, 为最终实现配电网运行分析的信息化、智能化、精益化奠定了基础。

**关键词:** 配电网运行分析; 辅助决策; 相关性分析; 故障风险预测; 薄弱点辨识

## 0 引言

随着智能配电网建设的不断推进, 配电网用电信息采集系统、营销业务应用系统、配电自动化系统和配变负荷监测系统等信息系统的不断完善, 如何灵活地将信息领域的大数据方法应用于配电网大数据分析, 实现对配电网故障的准确预测和潜在风险扫描, 找出薄弱环节加以改进, 已成为当前亟待解决的问题<sup>[1-3]</sup>。

国内的配电网信息化起步较晚, 数据质量、数据种类和数据完善度较差, 需对配电网信息系统进行调研和分析, 基于现有的、可靠的配电网数据信息进行预处理。考虑配电网多源异构海量数据蕴藏信息丰富, 通过特征提取和相关性分析方法, 从海量数据和众多故障特征中挖掘对配电网故障风险预测能力较强的特征<sup>[4]</sup>十分必要。配电网故障呈现弱特征和高频暂态特性, 使得配电网故障诊断及处理难度增大, 迫切需要研究适用于有源配电网的智能化的故障诊断与预测技术<sup>[5-6]</sup>。

配电网运维检修存在维护工作量大、运维检修人员不足等现象。为有效提高配电网运维分析决策水平, 增强配电网可控性, 缩短故障恢复时间, 计算机运行仿真辅助决策系统成为现代配电网规划建设不可或缺的工具<sup>[7-11]</sup>。目前, 国内外应用于配电网

运维管控的相关商业软件有 PSASP<sup>[12]</sup>, NETOMAC<sup>[13]</sup>, PSS/E<sup>[14]</sup>, ETAP<sup>[15]</sup>, PSAPAC<sup>[16]</sup>, BPA 和 DigSilent<sup>[17]</sup>等。这些软件虽然为配电网仿真计算分析提供了有力工具, 但并不直接针对配电网运维检修, 且缺乏配电网运行数据相关性分析、故障风险预测、薄弱点辨识等辅助决策功能。

本文设计了一种集数据相关性分析、故障风险等级预测和薄弱环节扫描辨识功能为一体的配电网运行分析系统。除了完成传统分析系统的数据汇集和预处理外, 采用改进特征选择算法实现了多源异构数据的相关性分析功能; 综合考虑了区域、时间和气象因素等配电网多维度、不确定性的故障影响因素, 扩展了基于改进决策树算法的故障风险等级预测功能和基于复杂网络理论的薄弱点综合辨识功能; 可以有效地降低配电网故障和风险发生的概率, 对于配电网高效精准化运行分析具有十分重要的意义。

## 1 系统架构设计

### 1.1 整体架构

配电网运行分析系统依托于电网现有业务系统进行高级应用的开发, 支撑配电网高效精准运维检修。系统整体框架设计如图1所示。

1) 数据层。基于国家电网有限公司公共数据模型(SG-CIM)标准对拟接入的设备资产运维管理系统(PMS2.0)、用电信息采集系统、地理信息系统(GIS)平台、配电自动化系统、气象信息系统、营销业务应用系统等系统的数据类型、格式、规范等进行

收稿日期: 2019-05-20; 修回日期: 2019-08-19。

上网日期: 2019-12-04。

国家重点研发计划资助项目(2017YFB0903000); 北京市自然科学基金资助项目(3172039)。

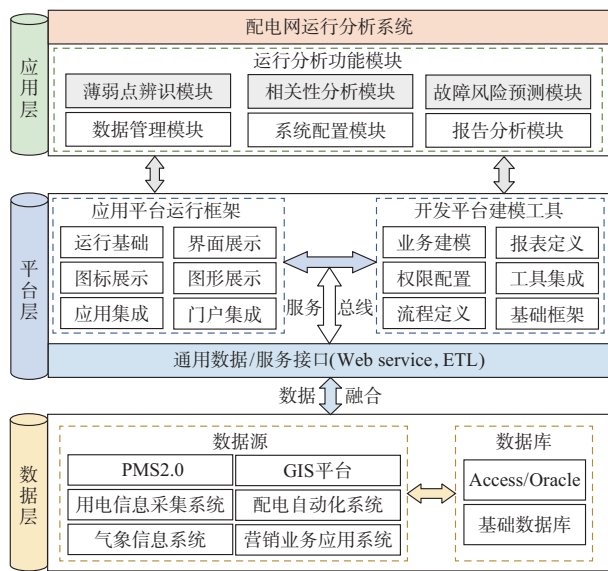


图1 配电网运行分析系统整体架构  
Fig.1 Overall structure of distribution network operation analysis system

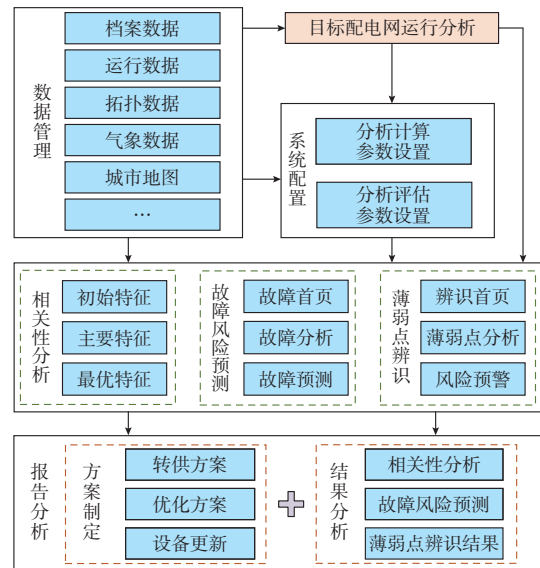


图2 系统功能实现流程图  
Fig.2 Flow chart of system function realization

深入分析,通过数据集成和清洗,根据配电网运行分析的实际要求,将故障数据、运行数据、气象信息和地理信息等数据进行深度融合,实现多业务数据系统的串行采集。该层由图形库和属性数据库组成,存储于 Access/Oracle 内,提供基础通用数据服务、数据库访问能力,供平台层调用。

2)平台层。平台层经过通用数据/服务接口,采用 Web service<sup>[18]</sup>、数据萃取-转置-加载(extract-transform-load, ETL)<sup>[19]</sup>等技术接入来自数据层的外部系统数据;该层由运行框架搭建和开发平台建模组成,包含工作流、事物处理、安全体系等平台的开源组件,应用平台运行框架由运行基础框架、界面展示框架、图标展示框架、应用集成框架等组成,开发平台建模工具由业务建模、报表定义、权限配置等组成。

3)应用层。底层使用 Java, C, C++ 等编程语言,基于 ECharts 图表库<sup>[20]</sup>进行前端数据可视化设计,分为数据管理、相关性分析、故障风险预测、薄弱点辨识、系统配置和报告分析模块,提供图表、报表报告、拓扑地图等展示,实现多级用户交互使用。

### 1.2 功能架构

配电网运行分析系统的核心功能通过应用层的6个模块实现。相较于传统分析系统,本系统对数据相关性分析、故障风险预测和薄弱点辨识功能进行扩展,并对分析和辨识结果进行报告生成和可视化展示。系统功能架构如图2所示。

1)数据管理模块主要实现数据集成、数据查询和数据处理。

2)相关性分析模块采用先进算法进行多源异构数据的关联分析,实现配电网故障最优特征的选择和薄弱点辨识指标的筛选功能。

3)故障风险预测模块根据气象、检修、负荷、范围和设备状态进行故障设置;进行停电原因分析、停电范围分析、停电等级分析;计算出区域、厂站和馈线各种故障的故障等级;对结果进行故障等级分析、故障位置分析、转供分析和优化分析;实现故障预测、抢修决策辅助决策等功能。

4)薄弱点辨识模块综合考虑网络拓扑结构和运行特性,实现配电网薄弱环节分析、扫描及风险预警。

5)系统配置模块主要完成风险自动分析计算设置和风险分析评估参数设置。

6)报告分析模块主要完成相关性分析、故障风险预测和薄弱点辨识结果的查询、生成和编辑,进而给出转供、优化和设备更新等方案。

## 2 关键技术

### 2.1 数据集成技术

配电网运行分析系统与部分配电网信息系统数据集成<sup>[21]</sup>实现流程如图3所示。

1)与营销业务应用数据集成设计。配电网运行分析系统与营销业务应用系统的数据集成主要包括标准代码,客户、计量点、表计、台区、配变等档案信息,以及配电网设备关系等。分析与营销系统数据集成需求的传输数据量、传输频率等要素,采用营销基础数据平台的技术路线。营销基础数据平台通过

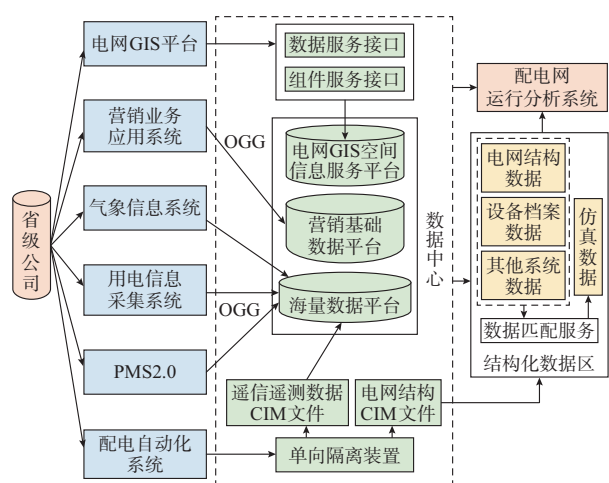


图3 数据集成结构  
Fig.3 Structure of data integration

数据同步工具 OGG(Oracle Golden Gate)<sup>[22]</sup>复制营销业务应用数据,配电网运行分析系统从数据中心营销基础数据平台获取数据,当用户、计量点、表计属性、计量箱等发生新增、更改和删除时,营销基础数据平台按需通过 OGG 的方式从营销业务应用系统进行数据处理,配电网运行分析系统通过 Java 数据库连接(Java database connectivity, JDBC)方式<sup>[23]</sup>调用实时更新数据。

2)与用电信息采集系统数据集成设计。采集系统通过海量数据标准接口方式将电量数据、电能表底示数等信息写入平台,部署配电网运行分析系统的计算服务模块,采用海量数据平台标准接口的方式提取电能表底示数。

3)与 PMS2.0 数据集成设计。PMS2.0 将输电部分的数据信息推送到数据中心,配电网运行分析系统通过 ETL 方式提取分析系统所需设备台账信息和网络拓扑信息。

4)与电网 GIS 平台数据集成设计。配电网运行分析系统与 GIS 平台的数据集成主要包括电网设备管理、基础地图和电网潮流图展示等信息。

5)与配电自动化系统数据集成设计。运行分析系统集成配电网设备及拓扑信息,由调控中心生成公共信息模型(CIM)格式文件<sup>[24]</sup>,数据中心负责数据解析。

6)与气象信息系统数据集成设计。气象信息系统通过海量数据平台接口将温度、湿度、风速等天气数据写入。

## 2.2 相关性分析技术

考虑配电网发生故障的影响因素众多,存在很多冗余、无关的故障特征,采用数据管理模块集成的多源数据:温度、湿度、风速等天气数据,设备位置数

据,配变容量、实时负荷数据、月最大负荷数据和停电时间、停电次数、缺供电量数据等,确定故障因素样本集,进行故障特征筛选。

传统的过滤式特征选择算法 Relief(relevant features),根据各个特征和类别的相关性赋予特征不同的权重,设置权重阈值选择特征。ReliefF 算法拓展了 Relief 算法处理多分类问题的能力<sup>[25]</sup>。

本文提出的相关性分析技术,在基于 ReliefF 算法完成特征选择后,进行了配电网故障特征之间的相关性分析,改进灰色关联分析方法实现故障特征的去冗余操作,最后输出最优故障特征集合(下文简称 G-ReliefF 算法)。故障风险特征相关性分析具体流程如图 4 所示。

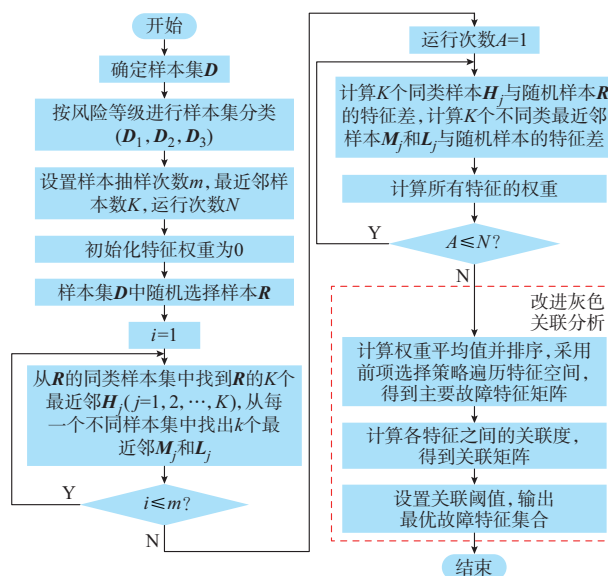


图4 故障风险特征相关性分析流程图  
Fig.4 Flow chart of correlation analysis on fault risk characteristics

改进灰色关联分析方法<sup>[26]</sup>主要步骤如下。

1)构建配电网主要故障特征矩阵,然后选择任意特征作为“参考序列”,剩余特征分别作为“比较序列”。

2)分别求解  $k$  时刻比较序列与参考序列之间的关联系数  $\xi(k)$ 。

3)求解故障特征之间的关联度  $\bar{\xi}(k)$ ,并建立关联度矩阵。

4)设置关联阈值,输出最优故障特征集合。

## 2.3 故障风险预测技术

机器学习理论中的典型数据预测算法有:神经网络算法、支持向量机(support vector machine, SVM)、决策树算法和提升算法等。

本文提到的 Adaboost 算法是一种典型的提升

算法,是通过针对同一训练集,将“弱分类算法”提升为“强分类算法”的过程<sup>[27]</sup>。采用了基于 Adaboost 改进决策树(简称 Ada-DT)算法进行风险等级预测。算法流程如图 5 中红色虚线框中所示,弱分类器采用决策树算法训练。通过弱分类器的偏差,更新样本的权重,然后重复迭代,得到强分类器,最终提高故障预测效果。在故障风险等级划分中,本文考虑不同供电区域配电网规模相差甚远,选取不同区域周故障停电频率、停电时长比例和缺供电量比例 3 类故障风险指标,将配电网故障停电风险程度分为一般、中度和严重 3 个等级。故障风险等级预测阶段的具体实现流程如图 5 所示。

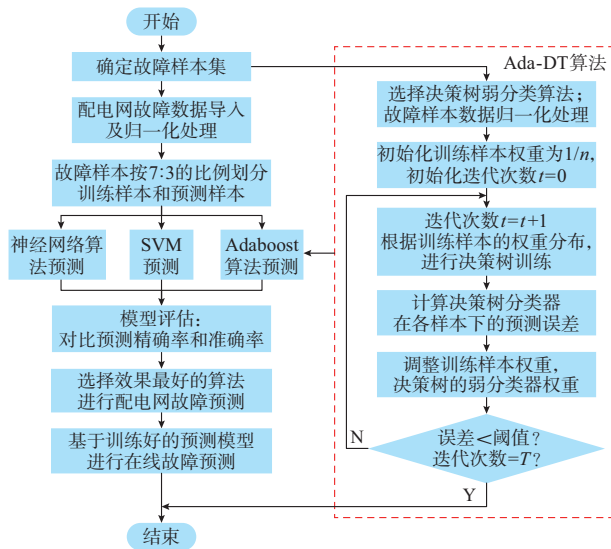


图 5 故障风险等级预测实现流程图

Fig.5 Realization flow chart of fault risk level prediction

通过 2.2 节中相关性分析技术筛选出用于故障风险等级预测的最优特征集合,并在此集合基础上有效确定预测阶段所需的输入样本集。首先,确定配电网故障样本集并导入,对样本数据经过“归一化”进行去量纲处理;对故障样本数据进行拆分,按 7:3 的比例随机拆分,70% 的数据用来训练模型,30% 的数据用来预测;选择机器学习中的反向传播 (BP) 神经网络、SVM 算法、Ada-DT 算法分别对配电网故障进行预测;对比预测的准确率,确定产生最佳效果的算法训练故障风险预测模型。

#### 2.4 薄弱点辨识技术

本文所提薄弱点辨识技术,综合考虑配电网安全性、经济性和网络拓扑结构 3 个方面,引入网络弹性性能变化、支路负载率变化风险、节点电压偏离风险和支路线损率变化风险作为关键节点辨识指标,结合主观经验和客观数据,采用综合评估方法分析指标权重,构建了多层次多角度的配电网关键节点

综合评估模型。

配电网薄弱点辨识方法流程如附录 A 图 A1 所示。主要步骤为:①读取配电网网络拓扑数据和运行状态数据,计算配电网系统初始稳定状态的潮流,进行节点  $N-1$  故障仿真试验;②计算配电网初始稳定状态的网络弹性变化值、节点电压偏离风险、支路负载变化风险和线损变化风险的指标值;③将所得指标值进行归一化处理,基于网络分析法<sup>[28]</sup>进行指标权重计算;④进行配电网薄弱点综合评分,根据综合评分和单一指标辨识薄弱点并进行风险预警和可视化呈现。

### 3 应用实例

#### 3.1 关键技术验证

系统在开发过程中分别以北方某地区和南方某地区实际配电网为实例进行了技术验证。采用北方某区县 52 个线段区域 2014 年 1 月至 2017 年 12 月的故障数据、运维数据和天气数据实现相关性分析和故障风险等级预测功能的测试;采用南方某配电网 2018 年夏大运行方式相关数据实现薄弱点辨识功能的测试。

##### 1) 故障特征相关性分析

统计区县不同线段区域不同月份的故障情况,经过数据集成和清洗,得到 18 个初始故障特征  $f_1 \sim f_{18}$  如附录 B 表 B1 所示。

采用本文改进 G-ReliefF 算法,初步选择主要特征集合  $\{f_2, f_{10}, f_{11}, f_5, f_3, f_4, f_9, f_{16}, f_{15}, f_{17}, f_{14}, f_{13}\}$ ,进一步得到关联性较高的 3 个特征  $f_5, f_3, f_4$ ,其中  $f_5$  和  $f_3$  之间关联度为 0.903,  $f_5$  和  $f_4$  之间关联度为 0.862,  $f_3$  和  $f_4$  之间关联度为 0.888。

经过相关性分析模块,筛选影响度较小的故障特征  $f_6, f_7, f_8, f_{11}, f_{12}, f_{18}$ ,剔除关联度很高的冗余特征  $f_3, f_4$ ,最终确定配电网故障风险最优特征集合为  $\{f_2, f_{10}, f_{11}, f_5, f_9, f_{16}, f_{15}, f_{17}, f_{14}, f_{13}\}$ 。

##### 2) 故障风险等级预测

故障风险预测模块,分别采用神经网络算法、SVM 算法、Ada-DT 算法进行风险等级预测,预测准确率对比如附录 B 表 B2 所示。

通过对比 3 种算法的准确率,最终选择指标值均较高的 Ada-DT 算法训练故障风险等级预测模型。将训练好的预测模型,以 2014 年 1 月至 2016 年 12 月的线段区域周故障数据作为训练样本,2017 年 1 月至 12 月的线段区域周故障数据作为测试样本预测校验,预测结果如附录 B 表 B3 所示。

附录 B 表 B3 中的“等级 1”代表配电网一般故障

风险;“等级2”代表配电网中度故障风险;“等级3”代表配电网严重故障风险;“综合”是3类样本总体预测准确率。由表B3可知,3类样本的故障风险等级的预测准确率达到88.3%。

### 3) 配电网薄弱点辨识

南方某地区配电网系统拓扑结构如附录B图B1所示。由于实际配电网系统中编号1和2的节点位于关键位置即电源节点处,仅进行除节点1,2外的薄弱点辨识校验,该地区配电网薄弱点辨识结果如附录B表B4所示。分析可知,经过薄弱点辨识模块得到各单一指标下的薄弱点排序,以及综合评分排名前十的薄弱点依次为:1,2,4,7,5,33,37,17,45,24。

## 3.2 系统应用成效

配电网运行分析系统在福建地市级供电公司进行了示范应用,通过分析配电网海量运行数据确定故障主要影响因素,对主城区各馈线和台区进行了故障风险等级预测,为科学制定抢修计划提供了技术支撑,提前辨识并治理配电网薄弱环节500多处,避免了直接经济损失。

系统薄弱点辨识分析环节在福建莆田进行了示例应用,通过选择莆田区馈线区域“工业I路633线”进行分析,将辨识得到的薄弱设备名称进行排序和可视化呈现,辅助工作人员制定检修方案。

## 4 结语

基于大数据、机器学习和复杂网络理论的配电网运行分析系统及方法,采用先进的计算机技术和配电网运行仿真技术,构建了集数据管理、数据相关性分析、故障风险预测、薄弱点辨识、报告分析等功能为一体的配电网运行分析系统,为配电网故障风险精准预测和薄弱点快速辨识奠定了基础,可以为城市配电网运行分析辅助决策方案的制定提供技术支撑。

目前,基于配电网海量数据的运行分析系统的研究与应用仍处于初步实现阶段,需要进一步进行大数据层面的配电网运行分析方法的仿真验证及现场实证。

附录见本刊网络版(<http://www.aeps-info.com/aeps/ch/index.aspx>),扫英文摘要后二维码可以阅读网络全文。

## 参 考 文 献

[1] 薛禹胜,赖业宁.大能源思维与大数据思维的融合:(一)大数据与电力大数据[J].电力系统自动化,2016,40(1):1-8.

XUE Yusheng, LAI Yening. Integration of macro energy thinking and big data thinking: Part one big data and power big data[J]. Automation of Electric Power Systems, 2016, 40(1): 1-8.

[2] 薛禹胜,赖业宁.大能源思维与大数据思维的融合:(二)应用及探索[J].电力系统自动化,2016,40(8):1-13.

XUE Yusheng, LAI Yening. Integration of macro energy thinking and big data thinking: Part two applications and explorations[J]. Automation of Electric Power Systems, 2016, 40(8): 1-13.

[3] 刘科研,盛万兴,张东霞,等.智能配电网大数据应用需求和场景分析研究[J].中国电机工程学报,2010,35(2):287-293.

LIU Keyan, SHENG Wanxing, ZHANG Dongxia, et al. Big data application requirements and scenario analysis in smart distribution network[J]. Proceedings of the CSEE, 2010, 35(2): 287-293.

[4] 牛东晓,谷志红,邢棉,等.基于数据挖掘的SVM短期负荷预测方法研究[J].中国电机工程学报,2006,31(18):6-12.

NIU Dongxiao, GU Zhihong, XING Mian, et al. Study on forecasting approach to short-term load of SVM based on data mining[J]. Proceedings of the CSEE, 2006, 31(18): 6-12.

[5] 毕天姝,倪以信,杨奇逊.人工智能技术在输电网络故障诊断中的应用述评[J].电力系统自动化,2000,24(2):11-16.

BI Tianshu, NI Yixin, YANG Qixun. An evaluation of artificial intelligent technologies for fault diagnosis in power network[J]. Automation of Electric Power Systems, 2000, 24(2): 11-16.

[6] 程乐峰,余涛,张孝顺,等.机器学习在能源与电力系统领域的应用和展望[J].电力系统自动化,2019,43(1):15-43.DOI: 10.7500/AEPS20180814007.

CHENG Lefeng, YU Tao, ZHANG Xiaoshun, et al. Machine learning for energy and electric power system: state of the art and prospects[J]. Automation of Electric Power Systems, 2019, 43(1): 15-43. DOI: 10.7500/AEPS20180814007.

[7] 徐特威,鲁宗相,乔颖,等.基于典型故障与环境场景关联识别的城市配电网运行风险预警方法[J].电网技术,2017,41(8):2577-2584.

XU Tewe, LU Zongxiang, QIAO Ying, et al. A risk warning method for urban distribution network based on associated recognition of typical fault and environment scenario[J]. Power System Technology, 2017, 41(8): 2577-2584.

[8] 余贻鑫,王成山,肖峻,等.城网规划计算机辅助决策系统[J].电力系统自动化,2000,24(15):59-62.

YU Yixin, WANG Chengshan, XIAO Jun, et al. Computer decision making support system for urban power distribution network planning[J]. Automation of Electric Power Systems, 2000, 24(15): 59-62.

[9] 罗凤章,王成山,肖峻,等.上海城市配电网规划辅助决策系统[J].电网技术,2009,33(3):79-83.

LUO Fengzhang, WANG Chengshan, XIAO Jun, et al. Decision-making support system for urban distribution network planning of Shanghai[J]. Power System Technology, 2009, 33(3): 79-83.

[10] 孟畅,顾洁,程浩忠,等.基于地理信息系统的配电网规划软件包的开发[J].电力自动化设备,2002,22(5):4-6.

- MENG Yang, GU Jie, CHENG Haozhong, et al. Development of distribution network planning software package based on GIS [J]. Electric Power Automation Equipment, 2002, 22(5): 4-6.
- [11] 龚博, 冯林桥, 卓佳, 等. 基于MapX组件技术的城市电网规划系统[J]. 湖南电力, 2007, 27(1): 1-4.
- GONG Bo, FENG Linqiao, ZHUO Jia, et al. Intelligent system development based on MapX component technology for substations planning in city distribution [J]. Hunan Electric Power, 2007, 27(1): 1-4.
- [12] 陆超, 唐义良, 谢小荣, 等. 仿真软件MATLAB PSB与PSASP模型及仿真分析[J]. 电力系统自动化, 2000, 24(9): 23-27.
- LU Chao, TANG Yiliang, XIE Xiaorong, et al. The analysis of power system models and simulation using MATLAB PSB and PSASP [J]. Automation of Electric Power Systems, 2000, 24(9): 23-27.
- [13] 郑三立, 韩英铎, 雷宪章, 等. NETOMAC在电力系统实时仿真中的应用[J]. 电网技术, 2003, 27(1): 18-21.
- ZHENG Sanli, HAN Yingduo, LEI Xianzhang, et al. Application of NETOMAC in real-time simulation of power systems [J]. Power System Technology, 2003, 27(1): 18-21.
- [14] 江全元, 彭旭东, 张宝星. PSS/E——一种先进的电力系统分析仿真软件[J]. 高电压技术, 2005, 31(1): 60-62.
- JIANG Quanyuan, PENG Xudong, ZHANG Baoxing. PSS/E—an advanced software package for power system analysis and simulation [J]. High Voltage Engineering, 2005, 31(1): 60-62.
- [15] 凌卫家, 吕东晓, 戴江江, 等. 面向用户的电力系统模拟计算高级应用软件[J]. 湖北电力, 2002, 26(2): 33-35.
- LING Weijia, LÜ Dongxiao, DAI Jiangjiang, et al. User-oriented software for the simulating calculation of power system [J]. Hubei Electric Power, 2002, 26(2): 33-35.
- [16] 韩祯祥, 张琦, 徐政. 电力系统分析软件的现状与发展[J]. 国际电力, 1999, 3(1): 51-54.
- HAN Zhenxiang, ZHANG Qi, XU Zheng. Present status and future of electric power system analysis software packages [J]. International Electric Power for China, 1999, 3(1): 51-54.
- [17] 董炜, 潘秋萍, 雷金勇, 等. BPA与DIgSILENT动态模型比较与数据转换[J]. 电力系统自动化, 2016, 40(8): 140-146.
- DONG Wei, PAN Qiuping, LEI Jinyong, et al. Comparison and data conversion between dynamic models of BPA and DIgSILENT [J]. Automation of Electric Power Systems, 2016, 40(8): 140-146.
- [18] 岳昆, 王晓玲, 周傲英. Web服务核心支撑技术: 研究综述[J]. 软件学报, 2004, 15(3): 428-442.
- YUE Kun, WANG Xiaoling, ZHOU Aoying. Underlying techniques for Web services: a survey [J]. Journal of Software, 2004, 15(3): 428-442.
- [19] 张宁, 贾自艳, 史忠植. 数据仓库中ETL技术的研究[J]. 计算机工程与应用, 2002, 38(24): 213-216.
- ZHANG Ning, JIA Ziyang, SHI Zhongzhi. Research on technology of ETL in data warehouse [J]. Computer Engineering and Applications, 2002, 38(24): 213-216.
- [20] 王子毅, 张春海. 基于ECharts的数据可视化分析组件设计实现[J]. 微型机与应用, 2016, 35(14): 46-48.
- WANG Ziyi, ZHANG Chunhai. Design and implementation of a data visualization analysis component based on ECharts [J]. Microcomputer & Its Applications, 2016, 35(14): 46-48.
- [21] 钱海龙. 一体化电量与线损管理系统的建设与应用[D]. 北京: 华北电力大学, 2018.
- QIAN Hailong. The construction and application of integrated power and line loss management system [D]. Beijing: North China Electric Power University, 2018.
- [22] 杨明珉, 陈勇. MySQL集群到Oracle数据库的数据同步方法[J]. 计算机系统应用, 2018, 27(6): 60-68.
- YANG Mingmin, CHEN Yong. Data synchronization method for MySQL cluster to Oracle database [J]. Computer Systems & Applications, 2018, 27(6): 60-68.
- [23] 罗荣, 唐学兵. 基于JDBC的数据库连接池的设计与实现[J]. 计算机工程, 2004, 30(9): 92-93.
- LUO Rong, TANG Xuebing. Design and realization for JDBC-based database connection-pool [J]. Computer Engineering, 2004, 30(9): 92-93.
- [24] 辛耀中, 米为民, 蒋国栋, 等. 基于CIM/E的电网调度中心应用模型信息共享方案[J]. 电力系统自动化, 2013, 37(8): 1-5.
- XIN Yaozhong, MI Weimin, JIANG Guodong, et al. Scheme of application model information sharing between control centers based on CIM/E [J]. Automation of Electric Power Systems, 2013, 37(8): 1-5.
- [25] 张稳, 盛万兴, 刘科研, 等. 计及天气因素相关性的配电网故障风险等级预测方法[J]. 电网技术, 2018, 42(8): 2391-2398.
- ZHANG Wen, SHENG Wanxing, LIU Keyan, et al. A prediction method of fault risk level for distribution network considering correlation of weather factors [J]. Power System Technology, 2018, 42(8): 2391-2398.
- [26] 王雁凌, 吴梦凯, 周子青, 等. 基于改进灰色关联度的电力负荷影响因素量化分析模型[J]. 电网技术, 2017, 41(6): 1772-1778.
- WANG Yanling, WU Mengkai, ZHOU Ziqing, et al. Quantitative analysis model of power load influencing factors based on improved grey relational degree [J]. Power System Technology, 2017, 41(6): 1772-1778.
- [27] 谭津, 邓长虹, 杨威, 等. 微电网光伏发电的Adaboost天气聚类超短期预测方法[J]. 电力系统自动化, 2017, 41(21): 33-39.
- TAN Jin, DENG Changhong, YANG Wei, et al. Ultra-short-term photovoltaic power forecasting in microgrid based on Adaboost clustering [J]. Automation of Electric Power Systems, 2017, 41(21): 33-39.
- [28] 陈柏森, 廖清芬, 刘涤尘, 等. 区域综合能源系统的综合评估指标与方法[J]. 电力系统自动化, 2018, 42(4): 174-182. DOI: 10.7500/AEPS20170605005.
- CHEN Baisen, LIAO Qingfen, LIU Dichen, et al. Comprehensive evaluation indices and methods for regional integrated energy system [J]. Automation of Electric Power Systems, 2018, 42(4): 174-182. DOI: 10.7500/AEPS20170605005.

张稳(1988—), 女, 博士研究生, 主要研究方向: 电力系统自动化、配电网风险评估与预警。E-mail: zhangwen\_098@163.com

盛万兴(1965—),男,博士,教授级高级工程师,主要研究方向:电力系统自动化、可再生能源发电。E-mail: wxsheng@epri.sgcc.com.cn

杜松怀(1963—),男,通信作者,教授,主要研究方向:电

力系统自动化、电力市场分析。E-mail: songhuaidu@cau.edu.cn

(编辑 章黎)

## Architecture and Technology Implementation of Massive Data Based Distribution Network Operation Analysis System

ZHANG Wen<sup>1,2</sup>, SHENG Wanxing<sup>2</sup>, DU Songhuai<sup>1</sup>, JIA Dongli<sup>2</sup>, KANG Tianyuan<sup>2</sup>

(1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China;

2. China Electric Power Research Institute, Beijing 100192, China)

**Abstract:** Accurate faults prediction and potential risk scanning of distribution network through flexible analysis and application of multi-source heterogeneous data are meaningful to realize efficient and accurate operation analysis decision of distribution network. The overall structure and function design of a distribution network operation analysis system based on mass data are introduced, and the functions and analysis results of each module are demonstrated by an application example. The system integrates data such as geographic information system (GIS), marketing business application and distribution automation. By using improved machine learning algorithm and weak point identification method, the functions of data correlation analysis, fault risk level prediction and weak point identification of target distribution network are extended. It is beneficial for relevant departments to propose corresponding technologies and management methods to carry out the operation and maintenance of distribution network, improve the scientific and practicality of the existing distribution network analysis system, and ultimately lay the foundation for the informatization, intellectualization and leanization of distribution network operation analysis.

This work is supported by National Key R&D Program of China (No. 2017YFB0903000) and Beijing Municipal Natural Science Foundation (No. 3172039).

**Key words:** operation analysis on distribution network; auxiliary decision-making; correlation analysis; fault risk prediction; weak point identification

