

大数据自动筛选技术在电力用户配电智能运维中的研究

熊威¹, 白天¹, 彭芬¹, 姜晓峰², 万元², 刘骏宇¹

(1. 五凌电力有限公司, 湖南 长沙 410004;

2. 湖南五凌电力科技有限公司, 湖南 长沙 410004)

摘要: 针对现行的配电智能运维的数据处理部分处理数据慢的问题, 分析大数据自动筛选技术在电力用户配电智能运维中的应用。首先, 清洗抽取电力用户端的海量数据, 并对数据中的缺失数据做聚类处理。之后, 使用小波变化特征提取技术提取筛选数据中的数据特征, 寻找数据中包含配电运维信息的数据, 读取信息。根据读取信息内容, 按照分类采取对应的运维措施。设计与现行配电智能运维的对比实验, 验证了应用大数据自动筛选技术的配电运维具有更佳的性能, 更适用于在大数据环境下对电力用户配电的运维。

关键词: 大数据; 自动筛选; 电力用户配电; 智能运维

中图分类号: TP311.13 文献标识码: A 文章编号: 1003-7241(2025)05-0169-04

Research on Big Data Automatic Screening Technology in Intelligent Operation and Maintenance of Power Distribution for Power Users

XIONG Wei¹, BAI Tian¹, PENG Fen¹, JIANG Xiao-feng², WAN Yuan², LIU Jun-yu¹

(1. Wuling Power Corporation Ltd., Changsha 410004 China;

2. Hunan Wuling Power Technology Co., Ltd., Changsha 410004 China)

Abstract: For the slow data processing part of the current intelligent distribution operation and maintenance, the use of big data automatic screening technology in the intelligent distribution operation and maintenance of power users is analyzed. First, the massive data of the power user end is cleaned and the missing data in the data is clustered. Later, the wavelet change feature extraction technology is used to extract the data characteristics in the screening data, find the data containing the power distribution operation and maintenance information, and read the information. According to the read information content, it takes the corresponding operation and maintenance measures according to the classification. The design and the comparative experiment with the current distribution intelligent distribution operation and maintenance verify that the distribution operation and maintenance applying big data automatic screening technology has better performance, which is more suitable for the distribution operation and maintenance of power users in the big data environment.

Keywords: big data; automatic screening; power user distribution; intelligent operation and maintenance

0 引言

在电能应用过程中, 对于配电运维的要求越来越高。电力用户配电和用电数据量不断增加, 从而构成大数据环境。利用大数据自动筛选技术能够通过海量数据, 对有用数据进行筛选, 使数据可用性得到提高, 并且方便后期数据工作。针对配电运维来说, 筛选海量数据中关于配电运维数据信息, 能够降低配电运维中的时间成本, 降低运维资源浪费, 还能够使电力用户使用感得到提高, 从而为用户提供优质电力服务^[1]。

将智能配电网设置在电力用户配电智能化运维过程中, 能够实现数据收集、传感和测量, 包括配电运维的数据汇总, 从而创建配电和用电大数据环境。目前, 还是使

用传统方法对电力用户智能化配电运维数据进行处理。在处理过程中要鉴别所收集数据的来源, 从而对数据中的信息进行鉴别分类。在对数据处理过程中, 运维数据分析时间会增加, 导致工作效率得到降低。为了使配电运维数据的处理效率得到提高, 可以使用大数据自动筛选技术处理配电运维数据, 对此方面的研究具有重要意义^[2]。

1 国内外研究现状

目前, 大数据被广泛应用到汽车、金融、电信、餐饮、娱乐等行业中, 随着现代信息化和智能电网技术的不断发展, 电力行业数据量也在不断增加, 在数据量、价值、多样性和速度方面都存在大数据的特性, 电力行业也进入到大数据时代中。通过音频通信设备、传感器、智能设备和

视频监控设备能够对电力大数据进行收集,使电力行业竞争力得到提高。国外IT公司都已经发布了大数据发展相关的白皮书与智能电网大数据分析系统,加拿大BC Hydro公司对于用户用电数据方面已经有相关研究成果。

法国电力公司大数据在客户行为特征方面被广泛应用,能够对客户与新产品的关联性、消费规律、客户群特征等进行分析,提高寻找目标客户的精准性,使市场份额得到扩大。美国BGE公司利用C3-Energy公司能源引擎平台集成内部数据源系统进行分析,并且研究了服务地区智能电表数据,一共有3500万条,10TB的云图像数据,每年的收益金额为200万美元。

在国家电网方面,十三五期间的公司一体化信息平台研究中对大数据平台基础体系架构进行了分析;内存计算、批量计算和掌握流计算方面的技术和数据实时收集、数据流复制等方面研究有一定进展。

通过电网科技发展与国家发展战略结合,国家电网提出了可以利用大数据技术和先进计算技术对计算机体系进行分析,并且研究了电力可视化展现技术、大数据分析挖掘算法等方面;基于智能电网的各行业大数据应用等。基于大数据建设实现国家电网的且有信息化,创新大数据的应用,从而加强数据挖掘,促进专业大数据应用的范围。

2 应用大数据自动筛选技术的电力用户配电智能运维

电力系统中存在各类数据与电力配电数据,影响实现配电智能运维的实际效果。需要利用相关技术将如图1所示的配电大数据筛选出来,以便于后续的配电运维管理。

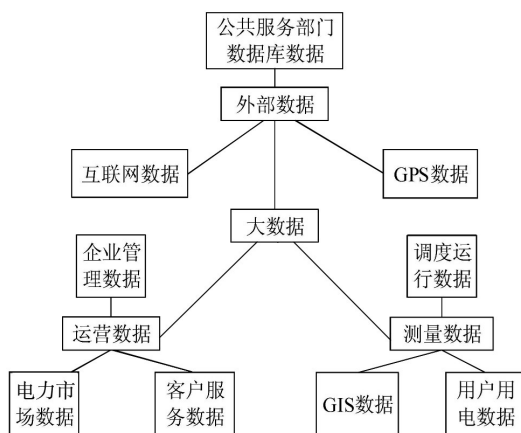


图1 智能配电数据类型

现行的配电智能运维在数据处理时,需要多个步骤,从海量数据中筛选有关信息。为优化现行的数据处理部分,使用大数据自动筛选技术,先对大数据做清洗等预处理后,筛选出与配电运维相关的大数据,之后从筛选后的

大数据中获取信息,按照获取的信息采取对应措施,实现配电智能运维。

2.1 电力用户配电大数据预处理

抽取所有电力用户端发出的电力大数据,组成数据集,使用离群数据清洗算法对数据集清洗。按式(1)计算数据集中各数据点间的距离 d_{ij} 。

$$d_{ij} = \left[\sum (g_i - g_j)^q \right]^{\frac{1}{q}} \quad (1)$$

式(1)中, g_i, g_j 为电力用户数据集中的数据点, $i=1, 2, \dots, n, j=1, 2, \dots, n, i \neq j; q$ 为距离计算公式的参数^[3]。数据集中,当部分数据与某一数据点间距离值大于 d_{ij} ,则该数据点为离群数据。即将距离 d_{ij} 作为数据清洗的阈值。为确保数据清洗的准确率,定义一个最小边界矩阵,来近似描述电力用户数据集的特征。

求出最小边界矩阵中各点与第 m 个最近距离的最大值和最小值,根据最小边界矩阵中的最大、最小距离,来判断是否存在离群数据,以此来简化计算。定义 $D(s, k)$,即距离的平方值为最小边界矩阵中各点与第 m 个最近距离值的度量, $D(s, k)$ 的计算式如公式(2)所示^[4]。

$$D(s, k) = \frac{d_{ij}^2}{d_{im} d_{jm}} \quad (2)$$

式(2)中, s, k 为最小边界矩阵中的两数据点,其中 s 是 n 维数据点, $s=[s_1, s_2, \dots, s_n]$ 。选取区域两对角端点 $q=[q_1, q_2, \dots, q_n]$ 和 $q'=[q_1', q_2', \dots, q_n']$,并且 $\forall i \in [1, n], q \leq q'$,组成矩阵 Q 。所以点 s 与矩阵 Q 之间的最短距离为:

$$\text{dist}_{\min}(s, Q) = \sum_{i=1}^n x_i^2 \quad (3)$$

式中, x_i 为点 s 与矩阵 Q 之间的距离,其取值如下:

$$x_i = \begin{cases} q_i - s_i, & \text{当 } p_i < r_i \\ s_i - q_i', & \text{当 } q_i' < s_i \\ 0, & \text{其它} \end{cases} \quad (4)$$

点 k 与矩阵 Q 之间的最大距离如下:

$$\text{dist}_{\max}(k, Q) = \sum_{i=1}^n x_i^2 \quad (5)$$

式中, x_i 的取值如下:

$$x_i = \begin{cases} q_i' - s_i, & \text{当 } s_i < \frac{q_i + q_i'}{2} \\ s_i - q_i, & \text{其它} \end{cases} \quad (6)$$

按照上述步骤完成对电力用户配电数据的清洗后,使用聚类算法对数据中的缺失数据集做聚类处理^[5-6]。如图2所示,将清洗后的数据集划分为缺失数据集 Q_1 和完整数据集 Q_2 。其中, $Q_1=(B, P), B$ 表示缺失数据集中的对象, P 为对象所具有的属性, $b_n \in B$,令 b_n 的个数为 $N; Q_2=(T, P), T$ 表示完整数据集中的对象, P 为对象所具有的属性, $t_m \in T$,令 t_m 的个数为 m 个^[7-8]。

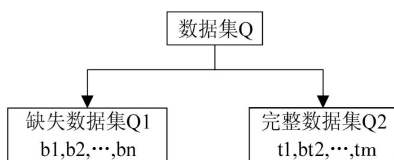


图2 数据集Q划分

以缺失数据集 Q_1 中的数据为聚类中心,将完整数据集 Q_2 中的数据按照欧氏距离与缺失数据聚类。计算完整数据集 Q_2 中的各数据对象与聚类中心的距离,将其放至距离最小的缺失数据类中。重复上述步骤,直至完整数据集 Q_2 中所有数据对象都聚类完毕。对用户端数据处理后,使用大数据自动筛选技术筛选出与配电运维有关的数据。

2.2 筛选提取配电运维信息

将预处理后的数据按照图3所示的映射方式,映射到数据筛选空间,在筛选空间中应用大数据筛选技术筛选出与配电运维有关的配电网网络故障、配电网网络负载、配电设备运行状态、配电网风险预警等数据。从筛选出的数据中,提取配电运维信息。

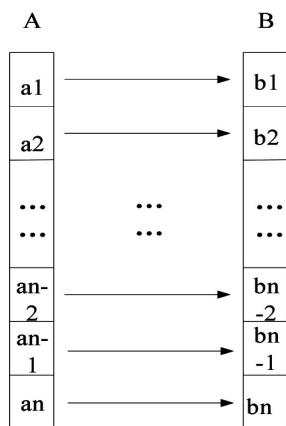


图3 映射关系

使用小波变化特征提取技术提取筛选数据中的数据特征,按照提取的数据特征将数据分类成设备故障数据、配电运行状态数据、电力用户负载数据、配电网风险预警数据等。将这些数据按照分类与样本数据拟合,并将拟合度90%的数据按照数据内部分段,再次与样本数据拟合,找出影响拟合度的数据段,并对数据段标记。读取被标记数据段内信息,按照信息结合配电运维中的处理办法,对数据对应的配电运行维护问题进行处理,实现电力用户的配电智能运维。

2.3 实现用户配电运行维护

根据标记数据段读取信息,对不同数据采取对应的处理方法。被标记数据代表配电设备出现故障,在保证电力用户用电基础上,对故障设备所在线路进行运行检修,排除线路影响导致设备不正常运行。对故障设备按

照互感器、变压器和电容器三种不同的维修方式,完成对故障设备的维护,使配电设备恢复运行。

当从配电运行状态数据中读取到状态异常的信息时,与电力调度沟通后,重置配电网络。通过短时重置,能够解决大部分的配电运行状态异常问题。

分析读取的电力用户负载信息,对电力负载过高的电力用户区,提高送电电压;对负载较低的电力用户区,降低送电电压。若电力负载过高区域过多,还可以降低负载正常地区的电压,以此来保证配电网络中的负载均衡,使配电正常运行。

读取配电网络风险预警信息,通过电力调度,降低预警等级,保障电力用户配电的正常运行。

至此,完成了应用大数据自动筛选技术的电力用户配电智能运维的研究。

3 性能测试

本文利用大数据自动筛选技术设计了电力用户配电智能运维,利用大数据自动筛选技术提高了电力用户智能运维的效率。为了测试性能,设计对比实验。通过与现行的电力用户配电智能运维的对比,测试本文设计的配电运维性能。

3.1 实验内容

为确保实验的真实可靠性,按照图4所示的拓扑结构,从实际电力用户配电网络中划分两个相同的实验区域。

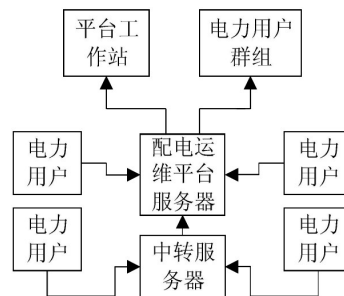


图4 拓扑结构

以本文设计的应用大数据自动筛选技术的电力用户配电智能运维为实验组,传统的配电智能运维作为对照组,两组均在如表1所示的测试环境中运行。

表1 测试环境

设备	参数
操作系统	CentOS release 6.5 (Final)
系统硬盘	2 TB
处理器	Intel Core Processor (Haswell) @2.09 GHz
处理器内存	2 GB
程序开发工具	My Eclipse8.0
Java 执行环境	jdk1.6

在实验区域对电力用户数量进行增减处理,并改变

电力用户内部群组的电力连接方式,使实验区域的电力用户的配电运维需求发生变化,共进行6次调整。实验组和对照组两个配电智能运维对实验区域进行管理,记录在实验区域配电需求变化后,两组对异常情况的监测反应时间和运营维护调整时间,计算平均时间,通过对比平均时间长短,验证性能优劣。

3.2 实验结果

与现行电力用户配电智能运维的对比实验结果如图5所示。

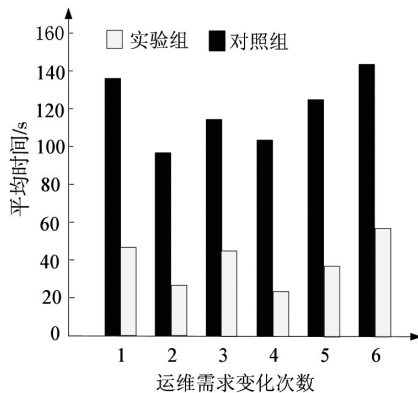


图5 实验结果

对图5进行分析,得出实验结论。对照组运维平均时间均高于90 s,说明对照组的监测反应和运营维护调整的总时间高于3 min。实验组运维平均时间均低于60 s,说明实验组的监测反应和运营维护调整总时间低于2 min。对每一次电力用户需求的变化,实验组的平均时间均远远低于对照组一半平均时间。在调整电力用户的电力需求,从而改变配电运维关系时,相比对照组,实验组能够更快反应并在最短的时间内对用户配电进行运维调整。综上所述,本文设计应用大数据自动筛选技术的电力用户智能配电运维能够提高配电运维效率,相比现行配电运维更具有优越性。

4 结束语

电力系统运行中,对电力用户的配电运维要求较高。

(上接第69页)

显著性检测[J].计算机辅助设计与图形学学报,2020,32(2):203-212.

[8] 沈洋,戴月明.基于加权模糊隶属度的二叉树多分类算法[J].计算机应用研究,2020,37(11):3281-3286.

[9] 张俊华,薛辉.基于无线网络的心理健康教育资源和分布式方法[J].自动化技术与应用,2023,42(11):114-117.

[10] 蒋科.基于数据挖掘算法的医疗网络安全风险评估[J].自动化技术与应用,2023,42(5):99-102,127.

[11] 袁江辉.智慧社区配电网供电可靠性风险动态综合评估方法[J].自动化技术与应用,2023,42(4):171-175.

[12] 杨力,秦红梅,苏华文.基于云模型和组合权重的SCADA

随着社会的发展,对电力用户数量越来越多,电力系统形成了海量数据,原有的运维对数据处理较慢,影响了电力系统的效率。本文利用大数据自动筛选技术对原有电力运维处理数据部分进行性能提升。通过与传统配电智能运维的对比测试,验证了本文设计的应用大数据自动筛选技术的电力用户配电智能运维能够提升整体运行效率,实现智能配电的高效率运维,更加具有优越性能。

参考文献:

[1] 李学军,张世元.大数据背景下基于改进RNN的低压配电网线损智能分析方法[J].沈阳工业大学学报,2025,47(1):130-136.

[2] 萧展辉,张世良,邓丽娟,等.数据缺失情况下配电网时间序列数据分类算法[J].沈阳工业大学学报,2025,47(1):29-36.

[3] 张明泽,栾文鹏,艾欣,等.基于边缘计算的台区短期负荷预测方法[J].电测与仪表,2024,61(4):93-99.

[4] 黄瑞,肖宇,曾伟杰,等.低压配电网HPLC通信信道特征分类研究[J].电网与清洁能源,2022,38(11):1-7.

[5] 董晨霓,狄婷.基于大数据技术的电力运维故障诊断及自动告警系统设计研究[J].电子产品世界,2025,32(2):53-56.

[6] Hou Weiwei.Design of power operation and maintenance fault diagnosis and automatic alarm system based on big data[J].Wireless Interconnection Technology,2022,19(8):60-61.

[7] 梁燮凡,王焯杰,谭喆.基于PLC技术的配电网雷击信号检测系统设计[J].自动化与仪表,2023,38(5):76-80.

[8] Shan Zehua.Application of voice prompt technology in distribution network line loss analysis[J].Electric Sound Technology,2025,49(2):89-91.

作者简介:熊威(1989-),男,工程师,研究:电力市场政策及增值服务研究。

系统安全风险评估研究[J].智能系统学报,2022,17(5):969-979.

[13] 贺德强,柳国强,陈彦君,等.基于正态云模型和模糊层次分析法的列车通信网络性能评估方法[J].交通运输工程学报,2022,22(2):310-320.

作者简介:程华福(1977-),男,硕士,中级工程师,研究方向:信息技术、网络安全,数据安全。