

# 基于大数据技术的重要用户供电安全分析

丁坚勇<sup>1</sup>, 周凯<sup>1</sup>, 田世明<sup>2</sup>, 潘明明<sup>2</sup>

(1. 武汉大学 电气工程学院, 湖北省 武汉市 430072;

2. 中国电力科学研究院, 北京市 海淀区 100192)

## Big Data Technology Based Important User Power Supply Security Analysis

DING Jianyong<sup>1</sup>, ZHOU Kai<sup>1</sup>, TIAN Shiming<sup>2</sup>, PAN Mingming<sup>2</sup>

(1. School of Electrical Engineering, Wuhan University, Wuhan 430072, Hubei Province, China;

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

**ABSTRACT:** In this paper, power supply security of important users is analyzed based on big data technology. Firstly, data about important user supply safety is discussed on its contents and sources both from distribution network and inside user, and unique characteristics of these data are analyzed. Then this paper probes storage processing analysis and big data diagnosis techniques about important user security. Finally, it takes an example to analyze potential risks in important user supply security based on cable temperature information using cluster analysis method. Result indicates that the device is deteriorating with serious potential risk, the example illustrates feasibility in user security diagnosis based on big data of important user.

**KEY WORDS:** important users; power supply safety; big data; potential risks mining

**摘要:** 结合大数据技术对重要用户供电安全问题进行分析, 首先从配网和用户内部 2 方面对重要用户供电安全相关的数据内容与数据来源进行梳理, 并分析了这些数据的独有特征, 在此基础上, 探讨了重要用户安全大数据存储处理技术和分析诊断技术。最后以某重要用户为例, 基于电缆温度信息, 采用聚类分析的方法对其供电安全潜在风险进行分析, 结果表明该设备处于劣化过程中, 存在较大风险隐患, 该应用场景验证了重要用户大数据在用户安全诊断中的可行性。

**关键词:** 重要用户; 供电安全; 大数据; 潜在风险挖掘

**DOI:** 10.13335/j.1000-3673.pst.2016.08.035

## 0 引言

重要电力用户是指在国家或者一个地区(城市)的社会、政治、经济生活中占有重要地位, 对供电可靠性有特殊要求的用电场所<sup>[1]</sup>。重要用户供电安全尤为重要, 国内外学者对重要用户用电安全已经

开始研究: 文献[2]介绍了重要用户供电现状, 并提出了保障用户供电的措施; 文献[3]对重要用户供电风险评价方法进行了研究; 文献[4]分析了不同电源路数对重要用户供电可靠性的影响。总的来说, 重要用户供电安全研究还处于起步阶段, 一方面在评价重要用户供电安全时, 大多只考虑了用户供电侧的运行情况, 而缺乏对用户内部用电安全的探讨; 另一方面, 当前的研究多采用传统建模方法, 只针对其安全的某一个或几个方面的问题进行分析。随着电力公司对用户管理的不断深入, 重要用户供电安全相关信息迅猛增长, 传统的建模方式已经无法全面实时评价其供电安全。

中国电机工程学会信息化专委会于 2013 年发布了《中国电力大数据发展白皮书》<sup>[5]</sup>, 阐述了电力大数据“3V3E”的特征, 以及将会给电力行业乃至社会带来的价值, 最后对其发展前景进行展望。随后国内学者掀起了电力大数据研究的热潮。文献[6]分析了智能电网大数据应用现状, 对发、输、配、变各领域大数据的来源和特点进行了分析; 文献[7-8]阐述大数据技术在配用电系统中的应用; 文献[9]提出了大数据技术的应用场景, 并验证了大数据技术的优越性, 电力大数据技术开始覆盖到电力行业的每个角落。但专门针对电力用户, 特别是重要用户的大数据分析应用场景还未见报道。

鉴于此, 本文以重要用户供电安全作为电力大数据的另一应用场景, 就包含配网侧和用户侧 2 方面的重要用户大数据内容及来源进行了全面梳理, 并结合重要用户安全数据独有的特征, 分析了大数据技术在其存储处理及其诊断分析方面的适用性, 最后以重要用户潜在风险挖掘为例, 论述了重要用户大数据在用户安全诊断中的可行性。

基金项目: 国家电网公司科技项目(521820140017)。

Project Supported by Science and Technology Foundation of SGCC(521820140017).

## 1 重要用户供电安全大数据来源和特征

随着电力部门对用户管理差异化、精细化需求的日益增加,以及电力信息化的推进,与重要用户供电安全相关的数据种类和规模快速增长,逐步形成了重要用户供电安全大数据。

### 1.1 用户供电安全数据来源

重要用户供电安全不仅与配电网供电可靠性有关,还与用户自身运行状态相关,因此,重要用

户供电安全大数据包括 2 个部分:配电网供电可靠性与用户内部的电力运行状态。

1) 配电网供电可靠性。大多数地市已具有多个配电管理系统,包括配电自动化系统、调度自动化系统、生产管理系统、电能质量监测管理系统、用电信息采集系统等,这些系统为用户供电安全提供了丰富的数据源,配网供电可靠性相关数据如表 1 所示。

表 1 重要用户供电安全影响因素数据分析

Tab. 1 Important user power supply security data analysis

数据种类	影响因素类型	信息内容	数据来源	数据结构
配网供电 可靠性数据	供电网络结构	用户供电电源回路/供电方式	地理信息/生产管理	结构化/半结构化/非结构化
	供电故障记录	设备故障记录/停电故障记录/检修记录	生产管理	结构化
	供电电压质量	电压合格率/频率偏差	用电信息采集	结构化
	负荷转移能力	容载比/设备利用率	负荷控制/负荷监测	结构化/半结构化
用户内部 电力运行 状态数据	负载率	电压/电流/功率	用电信息采集	结构化
	用户运行方式	变压器投运情况/开关状态/备用电源的运行情况	用户监测	结构化
	用户电压质量	电压暂升/暂降/电压三相不平衡/电压谐波畸变	用户监测	结构化
	用户运行状态	役龄/电压/电流/负荷率/运行温度	用户监测	结构化
	用户历史运行	故障记录/维修记录/巡检记录 设备说明书/用户电力拓扑图	人工记录	结构化/非结构化
用户运行环境	室温/湿度/强磁环境/雷雨 极端天气/用电管理水平/政治经济事件	用户监测/互联网	结构化/非结构化	

2) 用户内部电力运行状态。重要用户内部的安全数据几乎为空白。一方面用户内部监测装置不具有通信功能,只能依靠人工记录数据,而且监测信息中仅有功率、电压等,监测内容过于单薄;另一方面各用户用电管理模式参差不齐,管理水平差别很大。因此,要想实时、全面地了解重要用户运行状态,建立一套信息量全、分辨率高、实时性好的用户监测系统是电力公司对重要用户安全管理的迫切需求。本文结合国内外学者对重要用户安全运行的研究,对用户的用电安全信息进行梳理,见表 1。

①运行方式。文献[10]通过分析用户在不同运行方式下用户电力可靠性的情况,说明用户的运行方式对用户的运行有一定影响,而用户内部的运行方式主要通过切换母联开关和变压器的运行状态来改变,此外备用电源的投切也会影响到用户的运行。因此,对重要电力用户的变压器运行状态、母联开关开断情况和备用电源的投切进行监测,即可实现对用户运行方式的监测。

②电压质量。重要电力用户供电连续性至关重要,重要电力用户供电连续性不仅体现在供电是否中断,电压质量的优劣也会影响到重要电力用户的敏感设备能否正常工作,给用户带来巨大的生命财产损失,例如大型纺织企业,短时的电压跌落会导致整批产品的质量问题的,造成重大的经济损失。文献[11]对电力用户电压质量要求进行规定,本文也

针对反映电压质量的 3 个方面进行监测,保证重要电力用户正常、稳定运行。

③设备运行状态。电力设备的运行状态直接影响到用户运行安全稳定。国内外对影响设备状态的相关因素进行了研究,包括设备负荷率<sup>[12]</sup>、设备运行温度<sup>[13]</sup>、设备役龄<sup>[14]</sup>等。另外,用户内部大功率电器设备的启停也会对用户安全运行产生影响,如:大型酒店水泵和空调的启停会影响用户电压质量。

④历史运行记录。历史运行记录表征用户过去的运行状况,一定程度上反映了用户设备的老化情况和用户用电管理上的缺陷。重要电力用户内部历史运行记录主要包括设备故障记录、设备维修记录、设备巡检记录、用户档案资料。用户档案资料指保证用户安全运行的基础资料,例如电力设备说明书,用户用电结构图等。

⑤环境因素。重要用户安全可靠运行不仅与其实时运行状态和历史记录有关,还与用户运行环境因素有关。其运行环境不仅包括天气状况、用户设备所处室内环境等客观因素,还包括用电管理条件和政治经济事件等人为主观因素。环境因素信息的搜集需根据信息类型采用不同的渠道获取。

### 1.2 供电安全数据特征分析

重要用户供电安全数据除具有大数据的共性,还有一些独有特点。

1) 安全等级要求高。重要用户的信息安全关

系到一个地区乃至国家的安全,其供用电信息作为用户信息的重要组成部分,其中包含了大量保密数据,如地理位置,用户规模、作息时等。而且不同重要用户之间的信息管理水平参差不齐,许多用户的信息管理水平还相当落后,这些信息一旦落入不法分子之手,将会造成不可估量的损失。

2)数据混杂。重要电力用户供电数据内容混杂,一方面因为重要用户用电安全受到多方面因素的影响,其不仅与用户自身用电情况有关,还受到包括配电网运行情况、国家政治和经济政策调整等其他方面的影响;另一方面,重要用户供电安全数据结构多样,不仅有实时监测和历史记录等结构化数据,还包括文本数据、时间序列数据等各类半结构化及非结构化数据。除此以外,其供电安全数据分布广泛,一部分数据来源于配电网各级生产、管理系统,还有一些来自用户自身监测与记录等。

3)实时性强。重要用户用电要求比起一般用户要高,对电能监测的内容更多、精度更高、采集的频率也更高,例如,对重要用户电压暂升或暂降监测,按照文献[11]要求,监测频率为微秒级。另外,由于用户用电的随机性、以及政治经济事件等因素的影响,在用户内部会出现冲击性负荷的情况,重要电力用户负荷曲线在峰谷用电规律的基础上,会有短时间的波动,而这些波动往往信息含量很高,是数据挖掘的关键,因此,重要用户供电安全信息采集实时性要求高。

此外,重要电力用户大数据还具有数据交互性差的特点。重要电力用户分布在不同城市中,所属地域比较分散,且重要电力用户涉及行业众多,包括政府、教育、医院、企业等,各用户之间相对独立,自身条件和周边环境各不相同,各用户之间的监测数据难以进行交互。

## 2 重要用户供电安全大数据关键技术

重要用户大数据的形成必将给电力用户和电网公司带来双赢局面:电力公司能够实时了解重要用户的运行状态,真正实现对电网的全面管理;重要用户可依据自身运行状态和电网公司的评价分析,对自身用电行为进行调整,以达到安全、智能、经济的用电模式。

### 2.1 数据采集与集成

根据 1.1 节对重要用户大数据来源的分析,重要用户大数据一方面来自于用户内部运行状态监测,这些数据可分为 2 类,即实时变化的数据(如电缆温度、开关状态等)以及历史运行数据(如各设备

的故障记录、图纸说明书等)。其中,实时变化的数据采用安装传感器的方式采集,历史运行数据采用人工获取的方式记录。另一方面的数据来源于配电网的多个管理系统,但这些数据以各种数据形式存在于不同的数据库中,运行管理已形成竖井模式,不同单位、不同部门之间独立工作,交互性较差,数据查询面临困难,数据共享还存在很多问题。

因此,需要把不同来源、不同格式、不同特征的数据有机地集中,形成一系列面向同一主题的数据集合,实现系统全面的数据共享,即大数据的数据集成管理。它包含关系型和非关系型数据库技术、数据融合和集成技术、数据抽取技术、过滤技术和数据清洗等<sup>[5]</sup>。图 1 为重要用户安全数据交互平台,该平台收集用户内部运行状态信息的同时,也与相关配网管理系统接口,获取配电网相关数据。此外,重要用户安全数据服务器还能从互联网获取天气、政治经济事件等影响因素。

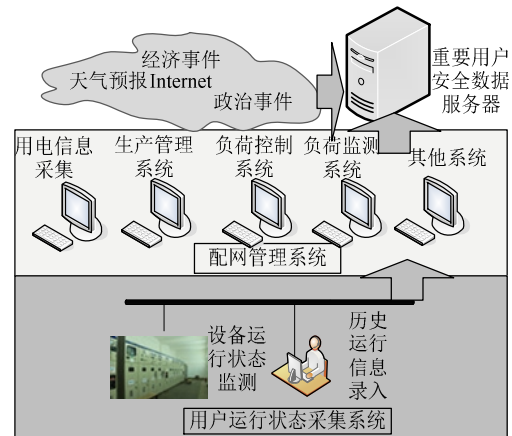


图 1 重要用户供电安全数据交互示意图

Fig. 1 Schematic diagram of important user power supply security data exchanging

### 2.2 数据存储与处理

传统关系性数据库不能满足上送数据量过大和相关用户行为分析数据的处理要求<sup>[6]</sup>,重要用户安全信息中大量非结构化数据涌现,这些数据无法用传统数据库处理<sup>[15-16]</sup>,因此急需能够应用于大数据的存储与处理技术。大数据存储与处理模式主要包含 2 种类型,流处理和批处理。流处理技术适用于数据量大、持续性好且速度快的信息,它将数据看作流,对新数据立即处理并返回结果,这种处理方式可满足重要用户安全大数据中实时性要求较高的在线监测数据,如用户电压暂升、暂降事件获取。批处理技术适用于先存储后计算、实时性要求不高,而准确性和全面性要求高的信息,例如,在分析用户设备故障概率时,不仅要分析设备自身故障事件,还应分析同型号设备的家族缺陷等信息。

重要用户供电安全各数据处理方式如表 2 所示。

**表 2 重要用户供电安全数据处理方式**  
**Tab. 2 Important user power supply security data processing**

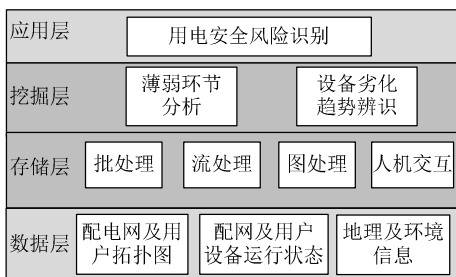
信息类别	获取方式	存储与处理模式	分析诊断
运行状态	传感器采集	流处理	实时分析/海量
运维管理	人工录入	人机交互	离线分析/TB
运行方式	传感器/人工	批处理/图处理	实时分析/TB
运行环境	传感器/互联网	流处理	实时分析/TB
供电网络结构	配网管理系统获取	流处理/图处理	实时分析/海量
供电故障记录	配网管理系统获取	批处理	离线分析/TB
负荷转移能力	配网管理系统获取	流处理	实时/海量
负载率	配网管理系统获取	流处理	实时/TB

另外，还有交互式处理和图数据处理模式<sup>[17]</sup>，交互式处理模式以人机对话的方式采用问答模式逐步实现，按这种模式，存储在系统中的数据文件能够被及时处理修改，处理方式直观、灵活。图数据处理模式根据图中点和边的强关联性，对图模式进行挖掘以及图数据的分类、聚类等。

**2.3 大数据安全分析诊断**

重要用户供电安全大数据分析诊断技术是传统数据挖掘技术在海量数据挖掘下新的应用与发展，但由于数据量大、增长速度快、种类繁多，且结构化、半结构化和非结构化数据共存，传统的数据表达模型和方法通常是简单的浅层模型学习，无法对底层数据特征进行抽象，学习精度和效率不尽人意。

基于重要用户大数据特点，可以从 3 方面对用户数据分析方法进行分类：1) 数据实时性，根据数据的实时性可分为实时分析和离线分析。例如，EMC 的 Greenplum<sup>[18]</sup>和淘宝开源的 Timetunnel<sup>[19]</sup>分别是实时分析和离线分析的工具；2) 数据规模，根据数据的规模可分为内存级、TB 级和海量级。3) 业务的算法复杂程度，例如，针对用户电压质量和用户可靠性分析 2 个不同业务时，涉及的数据量以及算法复杂程度都有很大差距，电压质量分析采用传统的数据分析方法就能实现，而可靠性分析需要综合设备运行状态，供电拓扑结构，运行环境等多方面的因素，最终得出用户可靠性值。图 2 为基于监

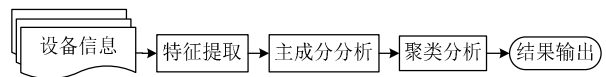


**图 2 基于监测数据挖掘的潜在风险识别平台技术框架**  
**Fig. 2 Technology framework of potential risk identification platform based on monitoring data mining**

测数据挖掘的潜在风险识别平台技术框架，该构架基于多维度的状态信息和数据分析手段，对基础数据进行挖掘，最终实现用户用电安全风险识别。

**3 基于监测数据挖掘的潜在风险案例分析**

用户设备状态监测数据中往往存在一些显著不同于其他数据的异常点，在很多数据挖掘过程中，异常点被当成噪声，直接滤除。但也有一部分异常点预示着故障、突变等罕见情况，比其他常见数据具有更大的信息量与价值，需要对这些数据进行挖掘。图 3 为设备状态信息挖掘流程，其关键技术包括特征提取、主成分分析、基于距离计算的聚类分析。特征提取和主成分分析是对数据进行分类和筛选，实现数据挖掘计算的简化；距离计算是聚类分析中较为关键的部分，基本思想为，利用某历史样本的特征距分布中心的距离来表征该样本的异常程度，距离越大，则出现的概率越低，异常程度越高。



**图 3 设备状态信息挖掘流程**  
**Fig. 3 Device status information mining process**

在江西某重要用户内部建立用电安全预警系统，对重要用户内部各设备进行状态监测，取用户某条电缆在一段时间内的监测数据，根据该设备不同阶段的运行温度变化情况挖掘故障隐患，选定距离最后一次采集数据 10 天内较新的数据，和距离最后一次测量 10 天前较老的数据，采用上述方法进行数据挖掘，对这 2 批数据分别进行计算，从而得到局部聚类中心。最后将 2 个局部聚类结果置于总体聚类之上，进行对比分析，评估在 2 个时间段内温度变化及稳定程度。

如图 4 所示为温度数据的分析结果，区域 3 为所有样本点进行聚类分析和距离判断，95%置信度判定正常温度的样本点。区域 2 表示最近 10 天数据的聚类中心，区域 1 表示 10 天前数据的聚类中心，虚线框内表示正常数据的范围。横纵坐标分别为温度的异常程度和稳定程度，异常程度通过设备阈值来判定和计算，稳定程度通过温度变化率来描述。对数据分布进行马氏距离计算，得到图 4 结果。可以看出 10 天前的数据聚类中心距离正常状态中心较近，且比较稳定，而 10 天内的数据距离正常数据中心较远，也不稳定，说明设备正处于劣化状态，设备存在较大安全风险，需对该设备进行全面检查。

另外，重要用户供电安全事故大多数始于系统薄弱环节某个设备元件的故障，从而引发用户一系

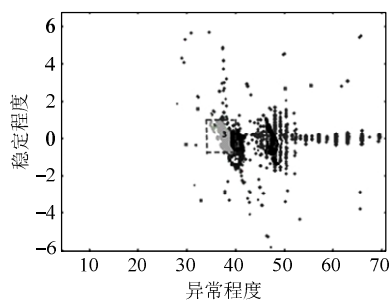


图4 用户电缆运行温度数据聚类分析

Fig. 4 Data clustering analysis of user cable operating temperature

列的故障，最终导致供电安全事故。因此，在设备故障事件挖掘的同时，还需要对用户供电系统中的薄弱环节进行分析。

#### 4 结论

未来的智能电网将是依托大数据处理分析技术的全景实时电网。重要用户作为智能电网的重要组成部分，其供电安全信息系统的建立为电网安全分析提供了各种异构且多样化的数据。电力公司能够实时了解重要电力用户的运行状态，真正实现对电网的全面管理；电力用户可依据自身运行状态和电网公司的评价分析，对自身用电行为进行调整，以达到安全、智能、经济的用电模式。

重要用户供电安全大数据在给电力公司和电力用户带来双赢的同时，又存在数据交互安全性难以保证、分析能力有限导致决策错误的挑战，如何面对这些挑战，需要后续研究工作中进一步探讨。

#### 参考文献

- [1] GB/Z 29328—2012 重要电力用户供电电源及自备应急电源配置技术规范[S]. 2012.
- [2] 徐阿元. 重要用户供电现状分析及其建议[J]. 供用电, 2014(6): 18-21.
- [3] 田世明, 潘明明, 苑嘉航, 等. 有源重要用户供电风险评估方法研究[J]. 电网技术, 2015, 39(1): 35-41.  
Tian Shiming, Pan Mingming, Yuan Jiahang, et al. Power supply risk element assessment for important consumers equipped with standby power supplies[J]. Power System Technology, 2015, 39(1): 35-41(in Chinese).
- [4] 徐浩, 陈健, 马帅, 等. 不同电源数对重要用户供电可靠性的影响研究[J]. 供用电, 2014(6): 36-39.
- [5] 中国电机工程学会. 中国电力大数据发展白皮书[R]. 北京: 中国电机工程学会, 2013.
- [6] 宋亚奇, 周国亮, 朱永利. 智能电网大数据处理技术现状与挑战[J]. 电网技术, 2013, 37(4): 927-935.  
Song Yaqi, Zhou Guoliang, Zhu Yongli. Present status and challenges of big data processing in smart grid[J]. Power System Technology, 2013, 37(4): 927-935(in Chinese).
- [7] 赵腾, 张焰, 张东霞. 智能配电网大数据应用技术与前景分析[J]. 电网技术, 2014, 38(12): 3305-3312.  
Zhao Teng, Zhang Yan, Zhang Dongxia. Application technology of big data in smart distribution grid and its prospect analysis[J]. Power System Technology, 2014, 38(12): 3305-3312(in Chinese).
- [8] 苗新, 张东霞, 孙德栋, 等. 在配电网中应用大数据的机遇与挑战[J]. 电网技术, 2015, 39(11): 3122-3127.  
Miao Xin, Zhang Dongxia, Sun Dedong. The opportunity and challenge of big data's application in power distribution networks[J]. Power System Technology, 2015, 39(11): 3122-3127(in Chinese).
- [9] 周国亮, 朱永利, 王桂兰, 等. 实时大数据处理技术在状态监测领域中的应用[J]. 电工技术学报, 2014, 29(增刊 1): 432-437.  
Zhou Guoliang, Zhu Yongli, Wang Guilian, et al. Real-time big data processing technology application in the field of state monitoring[J]. Transactions of China Electrotechnical Society, 2014, 29(Sup.1): 432-437(in Chinese).
- [10] 潘明明, 丁坚勇, 周凯, 等. 基于分层搜索算法的超高层建筑供用电可靠性评估[J]. 电网技术, 2015, 39(6): 1631-1636.  
Pan Mingming, Ding Jianyong, Zhou Kai, et al. A hierarchical searching algorithm based assessment on reliability of power supply and consumption of super high-rise building[J]. Power System Technology, 2015, 39(6): 1631-1636(in Chinese).
- [11] GB/Z 30137—2013 电能质量电压暂降与短时中断[S]. 2013.
- [12] 段盼. 电力系统负荷及负荷率的可靠性影响模型[D]. 重庆: 重庆大学, 2012.
- [13] 廖瑞金, 郑升讯, 杨丽君, 等. 不同温度下 XLPE 电缆中电树枝的生长特性[J]. 高电压技术, 2010, 36(10): 2398-2404.  
Liao Ruijin, Zheng Shengxun, Yang Lijun, et al. Growth characteristics of electrical tree in XLPE cable under different temperatures[J]. High Voltage Engineering, 2010, 36(10): 2398-2404(in Chinese).
- [14] 暴英凯, 王逸飞, 文云峰, 等. 考虑人为因素的设备可靠性评估及定检周期决策[J]. 电网技术, 2015, 39(9): 2546-2552.  
Bao Yingkai, Wang Yifei, Wen Yunfeng, et al. Equipment reliability evaluation and maintenance period decision considering the impact of human factors[J]. Power System Technology, 2015, 39(9): 2546-2552(in Chinese).
- [15] 王德文, 孙志伟. 电力用户侧大数据分析与并行负荷预测[J]. 中国电机工程学报, 2015, 35(3): 527-537.  
Wang Dewen, Sun Zhiwei. Big data analysis and parallel load forecasting of electric power user side[J]. Proceedings of the CSEE, 2015, 35(3): 527-537(in Chinese).
- [16] 田世明, 王蓓蓓, 张晶. 智能电网条件下的需求响应关键技术[J]. 中国电机工程学报, 2014, 34(22): 3576-3589.  
Tian Shiming, Wang Beibei, Zhang Jing. Key technologies for demand response in smart grid[J]. Proceedings of the CSEE, 2014, 34(22): 3576-3589(in Chinese).
- [17] 程学旗, 靳小龙, 王元卓, 等. 大数据系统和分析技术综述[J]. 软件学报, 2014, 25(9): 1889-1908.  
Cheng X Q, Jin X L, Wang Y Z, et al. Survey on big data system and analytic technology[J]. Journal of Software, 2014, 25(9): 1889-1908(in Chinese).
- [18] 孟小峰, 慈祥. 大数据管理: 概念、技术与挑战[J]. 计算机研究与发展, 2013, 50(1): 146-169.  
Meng Xiaofeng, Ci Xiang. Big data management: concept, techniques and challenges[J]. Journal of Computer Research and Development, 2013, 50(1): 146-169(in Chinese).
- [19] Ren Z, Xu X, Wan J, et al. Workload characterization on a production Hadoop cluster: a case study on Taobao[C]//2012 IEEE Int. Symp. on Workload Characterization(IISWC). Piscataway, NJ: IEEE, 2012: 3-13.



丁坚勇

收稿日期: 2015-09-09.

作者简介:

丁坚勇(1957), 男, 教授, 博士生导师, 研究方向为电力系统运行与控制、用户可靠性分析;

周凯(1991), 男, 博士研究生, 研究方向为配电网可靠性技术。

(责任编辑 李兰欣)