

基于大数据技术的电力用户行为分析及应用现状

沈玉玲, 吕燕, 陈瑞峰

(上海电气集团股份有限公司中央研究院, 上海 200070)

摘要: 电力行业是大数据技术应用的重要领域之一, 电力大数据产生于电力生产的各个环节, 蕴藏着丰富的商业和社会价值。介绍了大数据技术在电力行业的应用现状, 重点对基于大数据技术的电力用户行为分析方法进行了归纳总结, 最后给出了几个面向电力用户行为分析的大数据应用案例。

关键词: 电力行业; 大数据技术; 数据挖掘; 电力大数据; 用户行为分析

DOI: 10.3969/j.issn.1000-3886.2016.03.017

[中图分类号] TM71 [文献标志码] A [文章编号] 1000-3886(2016)03-0050-03

Power User Behavior Analysis and Application
Status Based on Big Data Technology

Shen Yuling, Lyu Yan, Chen Ruifeng

(Central Research Academy, Shanghai Electric Group Co., Ltd., Shanghai 200070, China)

Abstract: Electric power industry is one of important application fields of big data technology. Big power data, generated in each aspect of power production, bears a great deal of commercial and social value. This paper introduces the status of application of big data technology in the electric power industry, summarizes methods for analyzing power user behavior based on big data technology, and finally gives several big data application cases for power user behavior analysis.

Keywords: electric power industry; big data technology; data mining; big power data; user behavior analysis

0 引言

目前, 大数据技术已经得到广泛关注, 据权威部门预测, 2016 年中国大数据市场规模将达 6.17 亿美元, 而全球规模将达 238 亿美元。未来“互联网+”思维将推动新一波生产率的增长, 而大数据就是“互联网+”的发动机, 通过大数据来创造需求和预测未来, 进行资源合理分配, 从而达到社会效率的整体提升。

电力大数据是大数据理念、技术和方法在电力行业的实践, 它涉及到发电、输电、变电、配电、用电、调度等各个环节, 是跨单位、跨专业、跨业务的数据集合。电力大数据按来源不同可划分为三类: 一类是来自于电力生产企业的发电量数据, 二类是国家电网的运营和管理数据, 包括交易电价、售电量以及 ERP、一体化平台等方面的数据。三类是电力用户侧数据, 包括用户侧管理平台和智能电表采集的数据。电力大数据是大数据理念、技术和方法在电力行业的实践, 是大数据应用的重点领域之一。电力大数据具备普遍 5V (Volume; Variety; Velocity; Veracity; Value) 特征^[1-2], 即数量大、类型多、处理速度快、准确度要求高和价值高。2013 年《中国电力大数据白皮书》^[3]发表, 提出电力大数据还具备的独有 3E (Energy; Exchange; Empathy) 特点, 即能量、交互和共情。

1 大数据技术

大数据处理是个复杂的过程, 从各种类型的数据中“去冗分类”、“去粗存精”并快速获得有价值信息, 需要经过多个处理阶

段。目前, 比较经典的大数据处理模型是 Usama Fayyad 等设计的多阶段处理模型^[4], 如图 1 所示, 主要包括数据采集、数据预处理、数据存储与管理、数据的挖掘与分析及知识成果的展示五个处理步骤。

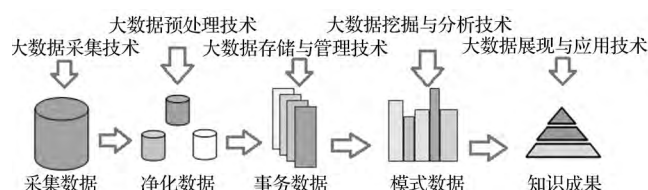


图 1 大数据多阶段处理模型

大数据采集一般分为大数据智能感知层和基础支撑层。大数据采集技术包括大数据的智能识别与感知技术、大数据传输与接入技术、分布式虚拟存储技术、大数据可视化接口技术、大数据的网络传输与压缩技术和大数据隐私保护技术等。

大数据预处理技术指对大数据进行辨析、抽取、清洗等操作。由于获取的数据可能具有多种结构和类型, 数据抽取过程就是将复杂的数据转化为单一的或者便于处理的构型, 便于快速分析处理, 对于大数据中无利用价值或者不关心的内容, 要通过过滤“去噪”剔除掉。

大数据存储与管理要用存储器把采集到的数据存储起来, 建立相应的数据库, 并进行管理和调用, 解决大数据的可存储、可表示、可处理、可靠性及有效传输等几个关键问题。大数据存储与管理技术囊括了大数据存储技术、分布式非关系型大数据管理与处理技术、数据融合技术、数据组织技术、大数据建模技术、大数据索引技术、大数据可视化技术、数据库技术、大数据移动-备份

定稿日期: 2015-11-16

基金项目: 上海市优秀技术带头人项目(14XD1420900)

-复制等众多关键技术。目前,比较常用的大数据存储与管理方案有 HDFS(Hadoop Distributed File System)、Tachyon、QFS(Quantcast File System) 和 ClusterFS 等^[9],其中 HDFS 是支持 Hadoop 计算框架的分布式大数据存储系统,具有很高的容错性、可扩展性、高并发性,并且可基于廉价的存储服务器设备,是目前最为流行的大数据存储系统。

数据挖掘就是“数据库中的知识发现(KDD)”,前面的几个步骤都是为数据挖掘分析做准备。数据挖掘的过程就是提取隐含在其中的、人们事先不知道的、但又是潜在有用的信息和知识的过程。传统的数据挖掘分析方法可分为:机器学习方法、统计方法、神经网络方法和数据库方法。传统的分析方法虽然仍可应用于大数据领域,但在处理规模较大的数据集时有一定局限性,这时就需要借助云计算技术,将大数据集的挖掘任务分解到多台计算机上并行处理,从而提升数据挖掘效率。

知识成果的展示与应用是大数据处理的最后一步,也是关键一步。如果数据分析得结果无法正确显示,反而会误导决策者。数据可视化技术是大数据展示的最有力方式,常见的可视化技术有基于集合的可视化技术、基于图标的技术、基于图像的技术、面向像素的技术和分布式技术等。目前,面向智慧城市的物联网大数据应用、面向在线社交网络的大数据应用、面向智慧医疗的健康大数据应用,以及面向智能电网的大数据应用都在如火如荼的进行中。

2 基于大数据技术的电力用户行为特征分析

由于我国电力行业的特殊性,电力用户行为分析仅限于电力消费的终端用户。目前,比较热门的电力用户行为特征分析研究方向,包括电力用户的分类及细分建模研究^[5-6]、电力用户用电负荷预测研究^[7]、电价与激励下的电力用户响应行为研究^[8]等。

2.1 电力用户分类研究

我国电力市场逐渐进入以市场需求为导向的新阶段,电力用户分类是实现电力资源的优化配置的必要手段之一,电力企业通过了解用户需求、用电行为、用电习惯等,可以选择用户,并有针对性的采取价格措施和激励政策,促进电力系统的平稳高效运行。

目前常用的电力用户分类方法主要有三类:一种是基于电力用户的综合价值评价体系的分类方法,评价体系由市场价值、潜在市场价值、区域贡献价值三部分组成。采用基于熵权法的用户评价分类方法,首先要采集相关数据,计算各参与评价指标,利用熵权法确定各个指标的权重,计算综合评价价值,并根据综合评价价值进行用户分类,利用聚类分析方法,将用户分为指定类。第二类用户分类方法是基于负荷曲线特征的用户分类方法,从电力数据中提取用户一段时间内的日负荷曲线,并根据季节和节假日进行分类,对多条日负荷曲线求平均负荷曲线,作为用户的代表曲线,采用模糊 C 均值、K-means 等聚类方法,最终确定用户的分类。第三类是依照经验规则的分类方法,包括基于业务逻辑规则和基于决策树的用户分类方法。无论哪种分类方法,都是为了确定电力用户的行为特征属性,提供给电力企业,以便提供有针对性的服务。

2.2 电力用户负荷预测研究

基于电力数据的用户负荷预测研究分为短期负荷预测和中长期负荷预测。短期电力负荷预测指未来几小时、几天的电力负

荷预测,短期负荷预测已作为能量管理系统的重要功能模块,为安排电力调度计划、供电计划等提供依据。中期符合预测指未来一年内的负荷预测,预测结果作为机组维修计划、电网规划等的重要依据。长期符合预测指未来 1~10 年的用电负荷预测,为电网的规划、增容和扩建等工作开展提供参考依据。

目前负荷短期预测理论研究逐渐成熟,包括回归分析法、时间序列法、小波分析、支持向量机、人工神经网络、模糊预测、综合模型预测等多种方法。相比短期符合预测,中长期负荷预测更容易受到不确定因素的影响,比如天气情况、自然环境和人类活动等,因此,中长期负荷预测需要的数据更多,难度更大。中长期负荷预测方法可以分为两大类:基于参数模型的方法和基于非参数模型的方法。基于参数模型的中长期预测方法包括电力弹性系数法、时间序列法、相关分析法等。基于非参数模型的方法包括灰色预测技术等。无论哪种预测方法,核心问题就是要基于电力用户历史负荷数据,建立预测模型,模型的精准度决定了预测水平的高低。

随着海量电力数据的出现,传统的负荷预测算法已无法满足预测速度和精度的要求,将传统的负荷预测方法与并行化计算模型 MapReduce、内存并行化计算框架 Spark 等技术相结合展开短期电力负荷预测,是大数据时代进行短期负荷预测的有效方法。

2.3 电力用户响应行为研究

从电力用户侧管理(DSM)到电力用户侧响应(DR),再到未来互动性更强的电力用户侧需求调度(DD)阶段,是电力企业发展模式面临的必然转变。电力需求侧响应的动机通常有两类:基于价格的需求侧响应和基于激励的需求侧响应。基于时变电价的需求侧响应又可分为分时电价、尖峰电价和实时电价。基于激励的需求侧响应可分为直接负荷控制、可中断负荷、需求侧竞价、紧急需求响应和容量、辅助服务计划等。

对电力用户响应规律的探寻离不开对响应行为的建模,目前电力用户响应行为的建模方法,一种是基于用户响应激励的分析模拟来建立机理模型,比如基于加权最小二乘法的用户响应曲线参数辨识模型、基于多智能体和模糊逻辑的响应规则以及基于市场均衡原理下的用户响应行为分析等等。随着电力大数据的形成,一些非机理性的建模方法也得到了广泛应用,比如基于知识学习规则的电力用户电价响应建模、基于模糊决策的用电态度建模、基于支持向量机回归算法的用户响应模型等。

3 电力用户行为特征分析应用实例

在电力大数据的理论研究及工程应用方面,欧美国家始终走在前列。在智能电网建设政策的支持下,我国近几年开始着眼电力需求侧领域,开展大数据关键技术的研究,并初步实现了一些应用^[9]。

3.1 法国

法国电力公司(EDF)自 2009 起对个人家庭用户安装智能电能表,用于采集个体家庭的用电负荷数据,并结合气象数据、用电合同及电网数据等信息,开发了基于大数据的用电采集应用系统。目前,法国电力公司的运行分析中心大数据项目组承担了客户数据的分析工作,以用户用电负荷曲线的海量存储和处理为突破口,利用大数据技术,开发能够在规定延迟时间内完成的复杂

并行处理计算能力,并搭建了大数据存储架构,以此为基础构建了分布式数据发生器 CourboGen 系统,用于生成用户用电负荷曲线及其关联数据。采用神经网络、聚类等方法对数据进行分析,预测电力需求侧的负荷变化,进行客户群分类,为法国电力公司的销售营销指出改进方向。

3.2 美国

比较著名的电力大数据应用就是“洛杉矶电力地图”,它由美国加州大学洛杉矶分校和当地政府机构共同研发,将每个街区信息、用户个人信息、用电实时信息以及地理、气象等信息全部整合,能准确地反映当地的经济状况及各用户群体的用电习惯,为城市和电网规划提供直观有效的负荷预测数据。利用先进的可视化分析工具,“洛杉矶电力地图”甚至可以辅助进行能源投资和能效决策。美国 C3 energy 公司的能源分析引擎平台,则是专门利用电力大数据提供能源投入冗余分析、节能计划和电力用户空间视图等服务的平台。

在用户侧响应机制上,美国也同样走在前列。纽约州的电力市场由 NYISO 管理,采用了四种方案的激励型需求响应机制,做到必要时削减负荷,从而保证纽约州电力系统的安全、可靠、经济的运行。激励型需求响应在美国已运行了 40 年,对其激励方案和运作形式的研究,将为我国电力市场改革提供借鉴。

3.3 中国

自 2013 年起,我国电力企业逐渐开始重视电力用户侧领域,开展电力数据中心的建设以及大数据应用技术的研发。截止目前国家电网公司已在 27 个省完成智能用电采集系统的部署,覆盖用户数达 2.42 亿户。

2013 年江苏省电力公司率先开展了大数据智能分析系统建设,采用 Hadoop 分布式批处理技术,建立包括电量、电力负荷、地理信息、气象数据等在内的多维度数据分析模型,开发了对数据分析结果的可视化展示界面,初步实现了电力看经济、电力看民生、用户用电行为分析三个方面的应用。

2015 年上海电力开展了首个电力需求响应试点实验工作,

(上接第 6 页)

5 结束语

目前电流源型变换器的研究很多集中在三相 CSR 整流器上,研究背靠背电流源型风电变流器的文献比较少,本文比较全面的介绍了背靠背电流源型风电变流器。首先建立并分析了机侧和网侧的数学模型;其次根据网侧和机侧的控制目标,结合数学模型提出了机侧和网侧控制算法;针对 LC 谐振问题,提出引入电感电压和电流的有源阻尼策略;最后介绍了系统参数设计方法并以 10 kW 的 MATLAB 仿真验证了上文提出的各种控制策略。

参考文献:

- [1] 赖纪东. 基于 CSC 永磁直驱风力发电系统协调控制方法与策略研究[D]. 合肥: 合肥工业大学, 2012.
- [2] 李玉玲. 电流型 PWM 整流器及其控制策略的研究[D]. 杭州: 浙江大学, 2006.
- [3] 吴奎华. 三相电流型 PWM 并网逆变器的研究[D]. 杭州: 浙江大学, 2008.
- [4] 茆美琴, 徐斌, 张邵波, 等. 基于小型风电系统电流源型逆变器的控制研究[J]. 电力电子技术, 2011, 45(8): 78-80.

为配合该项目的实施和为政府决策提供数据支撑, 2015 年年底上海将建成一个深度分析挖掘区域性、行业性用电特征的电力需求响应中心, 该中心对 2.8 万个用电大户的负荷数据进行采集分析, 立体展示企业的实时和历史负荷数据, 并通过专业分析, 提供给用电企业电力能源的分类使用情况, 以便进行用电行为分析, 更合理的利用电能, 减少能耗, 实现用户成本效益的最大化。

4 结束语

电力行业数据量大且价值高, 对数据利用率的提高, 可以加快电力企业盈利与控制水平的提升, 无论是行业内还是行业外应用, 都会带来新的发展机遇。同时, 这也对大数据的认知水平提出了挑战, 如何通过足够的数据存储和分析处理能力来有效地应用这些数据, 从中提取出有价值的信息, 并最终转化成可以优化管理模式和提升服务水平的决策知识, 这是大数据价值能否被充分展现和应用的关键所在。

参考文献:

- [1] 王继业. 大数据与电力企业[J]. 电力信息化, 2012, 10(8): 7.
- [2] 闫龙川, 李雅西, 李斌臣, 等. 电力数据面临的机遇与挑战[J]. 电力信息化, 2013, 10(4): 1-4.
- [3] 中国电力大数据发展白皮书(2013) [R]. 北京: 中国电机工程学会, 2013.
- [4] FAYYAD U, PIATETSKY-SHAPIO G, SMYTH P. From data mining to knowledge discovery [M]. An overview. AAAI/MIT Press, Cambridge, Mass, 1996.
- [5] 王臻, 冯勤超. 基于价值评价的电力用户分类研究[J]. 价值工程, 2009, 28(5): 64-67.
- [6] 王雷. 基于数据挖掘的电力行业客户细分模型研究[D]. 上海: 上海交通大学安泰经管学院, 2007.
- [7] 廖旋煊, 胡智宏, 冯莹莹, 等. 电力系统短期负荷预测方法综述[J]. 电力系统保护与控制, 2011, 39(1): 147-152.
- [8] 伍伟华, 庞建军, 陈广开, 等. 电力需求侧响应发展研究综述[J]. 电子测试, 2014, 21(3): 86-94.
- [9] 史梦洁, 韩笑, 程志艳, 等. 面向电力需求侧的大数据应用研究分析[J]. 供用电, 2014, 31(12): 20-23.

【作者简介】沈玉玲(1980-), 女, 山东人, 博士, 专业: 自动化。

- [5] DAI J, D. DAVID XU, B. WU. A novel control scheme for current-source-converter-based PMSG wind energy conversion systems[J]. IEEE Transactions on Power Electronics, 2009, 24(4): 963-972.
- [6] 刘福炎. 基于三相电流源型逆变器的永磁直驱风力发电系统的综合控制研究[D]. 合肥: 合肥工业大学, 2012.
- [7] WISEMAN J. C, W. BIN. Active damping control of a high-power PWM current-source rectifier for line-current THD reduction [J]. IEEE Transactions on Industrial Electronics, 2005, 52(3): 758-764.
- [8] 文教普. 三相电流型 PWM 整流器的研究与实现[D]. 南京: 南京航空航天大学, 2014.
- [9] Power Electronics and Motion Control Conference. A novel control method for IGBT current source rectifier, 2008 [C]. Poznan: IEEE Conference, 2008.
- [10] ZHIHONG B, HAO M, DEWEI X et al. Resonance damping and harmonic suppression for grid connected current-source converter [J]. IEEE Transactions on Industrial Electronics, 2014, 61(7): 3146-3154.

【作者简介】张亮亮(1991-), 男, 河南人, 硕士生, 研究方向为风电变流器。