

基于大数据的电力客户行为分析体系研究及实践

凌德祥, 黄 拓, 关晓林, 赵 严

(国网抚顺供电公司, 辽宁 抚顺 113008)

摘 要:随着我国电力市场的发展,供电企业紧跟数字化发展步伐加快信息系统建设,各业务系统在运行过程中积累了海量信息数据资源,利用这些数据资源进行数字化分析已经渐渐成为供电企业的重点工作。在供电企业不断地拓展分析范围,优化分析方法的过程中,大量的数据资产被盘活,巨大企业价值与社会价值逐渐显现。大数据技术的使用可体现在电力系统运行的各个环节,为电力信息的商业化利用提供有力支撑。基于大数据技术的电力客户用电行为分析能提高电力系统的运行效率,为供电企业的实际运行提供决策依据。文章分析了国内、外电力行业大数据应用的进程及现状,根据不同类型客户的用电行为建立了基于大数据技术电力客户行为分析体系,对大数据技术的实践情况进行探究,实现电力客户用电行为的深度分析,为客户提供更优质的服务体验。

关键词:大数据;电力客户;行为分析;聚类分析

文章编号:2096-4633(2018)10-0013-05 **中图分类号:**C39 **文献标志码:**B

随着我国电力市场改革的不断推进,售电侧市场引入竞争已经是大势所趋。如何在竞争中脱颖而出,是供电企业必需面临的挑战。电力客户是供电企业的直接消费者,抓住客户的需求,给客户以方便快捷的服务才是供电企业生存的根本所在,提高客户满意度是我们需要研究的重要课题。通过大数据挖掘技术来不断探究沉睡的数据资产潜在的价值,有利于制定个性化的营销策略,从而提高客户的满意度,留住现有客户,开发增量客户。

1 国内、外大数据应用进程及现状

法国从 2009 年开始,就注重使用大数据技术对用户的行为进行探索,到目前法国电力公司专门构建了大数据系统。该系统采取分布数据计算方式,能处理海量的信息,自动生成用户用电负荷曲线,统计与其有关的所有关联数据,在有效时间内完成复杂的数据处理工作。该系统还具有预测功能,将未来的负荷变化情况用曲线的方式表现出来,以便电力公司随时改变自己的运营对策。美国也相当重视大数据应用技术,2012 年在洛杉矶设立了电力大数据应用系统,该系统由当地政府与加州大学组织建设,采用大数据技术,系统按照街区地图统计每个与用户相关的数据,将用电情况与个人结合在一起,在这个过程中,加上地理信息和气象信息,将用电情况与更多的因素综合考虑到一起,能够全面地整合当

地用电情况。我国电力行业从 2013 年开始逐步运用大数据技术,对用户需求侧进行研究^[1-3]。目前已经全面铺开智能用电研究工作,目前已经可以通过用电信息采集系统直接将客户的用电数据进行采集,通过建立多维度数据分析模型及多类型的应用场景,为服务于企业和客户提供了有利的条件。电力客户的行为分析在整个行业都是概念和 POC 阶段,需要依靠智能电网有效支撑^[4],建立分析体系对客户行为进行分析。

2 电力客户行为分析体系建立的问题与对策

电力客户行为分析体系的建立是大数据技术思维方式很好的体现^[5-6],它的建立前提是在数据资产高质量的基础上,有效的对数据资产治理才能使供电企业数据资产更加准确、完整、一致、安全、安全,降低总体拥有成本,才能数据资产应用到各类监测、分析中去,提升企业管理决策水平,实现公司数据资产的真正价值。

2.1 各业务系统助互独立,存在专业壁垒

企业数据横跨营销业务应用系统、电能量信息采集系统、95598 工单管控系统等多个系统。数据字段烦多,匹配情况繁杂,跨业务应用间的数据融合成为开展电力客户行为分析的基础条件。解决数据开放与融合共享问题需要从两个方面入手:一是针对跨业务系统的大数据量分业务协同处理。通过各业务

明细数据的分布式计算,再将计算结果再进行集群存储再重新计算的方法,方便、快捷地实现各业务系统间的数据共享与整合;二是通过可视化的分析工作,建立友好的可视化展示界面,降低大数据开发及数据共享的门槛,应各业务管理人员能够快速的接受与使用,有效提升数据分析能力和工作效率。

2.2 数据结构繁杂冗复杂,数据质量有待提高

业务应用系统在运行过程中会产生许多垃圾数据,降低了数据资产挖掘地价值。为了保证数据资产的准确性、完整性与及时性,管理部门要在不断的应用过程发现数据问题、改进业务系统应用,促进数据资产质量的不断提高。提高数据资产质量依然从两方面入手:一是供电企业从企业战略的层面入手,将各业务应用系统进行数据资产专项治理工作,有序拓展数据资产管理范围,逐步实现数据全业务覆盖。二是建立公司数据资产管理体系,从资产架构上调整,将元数据统一管理,形成元数据全生命周期管理,并将元数据应用到各各业系统,以实现在数据资产的质量管理、调度运维、数据开放等管理的每一个环节。

2.3 数据资产管理体系不健全,保障机制需建立

数据资产的概念近几年刚刚提出,许多企业还没有意识到数据资产的价值,业务应用的数据疏于管理,许多数据都沉睡在数据库当中,不能充分发挥数据资产的价值。为了唤醒沉睡的数据资产,保证数据资产的有效利用,需要建立数据资产管理体系,健全数据资产保障机制与信息数据脱敏机制:一是以现有的业务应用系统数据资产为基础,设计数据质量管理方案,从数据准备、数据治理、数据评估、数据质量改进等方面建立数据资产管理体系。使数据资产实现资产全生命周期管理,使工作过程标准化,快速找到各业务应用系统中的数据质量问题,及时修正数据差错,改进数据流转与应用的问题,对今后对数据资产进行大数据挖掘提供有效地支撑;二是建立数据安全保障机制。从数据资产的安全、数据资产的保密与数据资产的共享等方面进行研究,使数据从采集、传输、使用与开放各环节实在加密与备份,而成数据资产开放策略,实行不同权限、不同级别不同的开放标准与审查机制,以保障数据资产的安全。

3 电力客户行为分析体系

基于大数据的电力客户用电行为分析能提高供电企业的服务质量,提高电力客户的体验舒适度,电

力客户行为分析体系主要由客户贡献程度分析、客户诉求行为分析、用电及窃电分析和诉缴费及欠费行为分析等内容构成^[7],构成体系如图 1 所示。

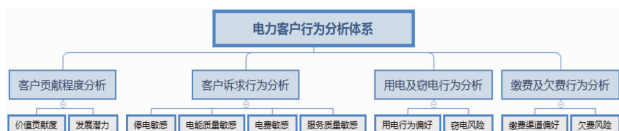


图 1 电力客户行为分析体系

Fig. 1 Behavioral analysis of power clients

3.1 客户贡献程度分析

供电企业对客户贡献程度的评价依据是客户用电量与缴纳电费的多少。因为供电企业主要利润来源用售电量,在不产生欠费的情况下,电费为主营业务收入。供电企业对客户发行的电量及缴费情况有详细的记录,根据客户用电量来计划客户应该交纳的电费,并按缴费情况分析客户的经济潜力,分析供电企业未来经济增长状况。供电企业能通过对客户贡献程度进行分析,规划供电企业未来的发展道路,保障和增长充足的客户量,尽快实现供电企业的高效发展。

3.2 客户诉求行为分析

供电企业对客户服务主要包括客户的咨询、业扩报装、投诉、故障报修等内容。使用大数据技术记录客户的诉求内容,这样供电企业可以针对客户诉求的内容对服务业务进行整改,适当改良公司颁布的政策制度,以提高公司的服务质量和效率。供电企业诉求行为分析的对象是不同种类的人群,比如大客户希望提高电网的可靠性,事业型人群想要提高业务办理速度,供电企业可以通过开设多种咨询渠道了解顾客的需求,此时供电企业需为其制定较好的营销策略,满足客户的需求,避免出现客户投诉等事件的产生,提高客户的满意程度和业务体验舒适度^[8-9]。

3.3 用电及窃电行为分析

供电企业通过电能量采集系统收集客户用电信息后,提取客户时点、电压、电流、负荷曲线等信息,结合季节、气温、地理信息等重要因素完成对客户用电行为的整理,增强供电企业对客户用电行为的了解程度,有利于供电企业应对客户的实际用电需求,并根据整理好的数据适当采取节能、无功补偿,扩展容量等技术维持电力系统的正常运行。供电企业还可以通过开展客户用电行为分析工作,及时检测出盗电客户,避免对供电企业造成损失。所以供电企

业对客户的用电行为进行分析,能提高公司的工作效率,保障居民的正常用电,提高公司的经济效益。

3.4 缴费及欠费行为分析

供电企业通过营销业务应用系统可以查询客户的缴费记录,对客户的缴费时间、缴费金额、欠费次数、欠费金额缴费方式等数据进行统计。根据收集的数据对客户进行分类,分为高风险、低风险和中风险客户。供电企业针对不同类别的客户,可采取不同的服务政策,比如针对高风险客户,供电企业需对其进行风险评估,结合客户的缴费情况为客户提供较好的降低风险的措施,提高客户的风险防范意识。同时,供电企业需收集客户对缴费方式的评价,并根据收集的信息对缴费方式进行改良,提高供电企业的服务效率。电力客户行为分析涉及众多的企业数据资源。

4 电力用户行为分析应用实践

电力用户行为分析要从大量数据中挖掘出隐含

的、有潜在价值的规律和过程。数据挖掘过程包括数据准备、数据存储、数据运算、数据分析和数据可视化五个过程。面对多个业务应用系统中的增量数据,大数据挖掘技术可以在庞大复杂的数据海洋中,剔除无效的和错误的数 据,得出隐含在数据资产中某种的影响因素,确定数据资产的内在联系,找到真正变化规律,从而提高分析与预测的准确度^[10-12]。

4.1 数据整合与清洗

电力客户用电行为分析需要从相关业务系统的数据库中提取相关数据。包括营销业务应用系统中电量电费模块中的客户发行电量数据、发行电费数据、缴费模块中的缴费数据、缴费方式、欠费数据、95598 业务模块的业报装数据、故障报修数据、投诉举报数据;采集电力用户用电信息采集系统中的实时数据,电流、电压负荷等数据。然后按照客户编号进行匹配,去除无效数据、空白数据及无法匹配的数据。形成主题数据库^[13]。具体业务应用系统提取字段见表 1 电力客户用电行为分析数据需求表:

表 1 电力客户用电行为分析数据需求表

Tab. 1 Data demand table for behavioral analysis of power clients

数据分类		所 属 字 段						
用电客户信息	客户编号	所属地区	行业类别	电压等级	电价类别	电价码	电表号码	客户地址
客户服务数据	客户投诉数量	有责任投诉数量	回访满意率	不属实及无责任投诉数量	属实投诉按投诉原因类型分布	属实投诉按地区分布情况	投诉原因	投诉类型
客户用电数据	上一年度用电总量	上一年度用电增长变化率	本月有功总电量	电压	用电时间	电流	发行月份	本月无功总电量
客户缴费信息	欠费次数	发行电费	当月连续欠费次数	欠费金额	欠费时间	缴费方式	电费月份	当年总欠费次数
电能质量数据	用户平均停电时间	城网用户供电可靠率	电压合格率	农网用户供电可靠率	电压等级	一个月停电两次及以上台区名称	停电时间	停电用户数
生产数据	输电线路长度	变电容量	配电线路长度	线路损失率	设备故障率	设备型号	设备类型	投运日期
调控数据	抢修工单数量	故障报修到达现场时间兑现率	城区用户抢修到达现场平均时长	农村用户抢修到达现场平均时长	7×8 小时外紧急非抢修类工单	用户地址	台区名称	故障原因
外部数据	区域地图	天气情况	自然灾害	气温情况	日期	时间	全民生产总值	经济增加值

4.2 数据存储与计算

任何强大的单一服务器都满足不了大型系统持续增长的业务需求,数据库读写分离随着业务的发展最终也将无法满足需求,需要使用分布式数据库及分布式文件系统来支撑。建立基于 Hadoop 的数据仓库 Hive。现在 Hadoop 已经发展成为了一个包

括分布式文件系统(hadoop distributed file system, HDFS)、分布式数据库(HBase Cassandra)以及数据分析处理 MapReduce 等功能模块在内的完整生态系统(ecosystem),是目前最流行的大数据处理平台^[14-15]。根据选择业务场景需求的不同,选择不同的数据分析模型,包括聚类分析、关联分析、回归分

析、时间分类分析和分类预测。按照模型评估器评估,选择最适合模型。电力客户行为分析可对用户群进行分类,所以本次选择聚类分析模型,在对客户进行分类中采用最常用的 K-means 算法,标出用户的分类特征^[16-18]。

4.3 数据分析与可视化

可视化是展示大数据挖掘结果最直接而有效的手段之一,它是利用科学地运算将抽象的数据分析结果展示出来,在许多行业得到广泛使用。数据可视化(data visualization)技术是运用现代计算机图形学和图像处理技术,将数据转换为图形或图像在显示器上展现出来,并可以实现交互处理的技术、方法或者理论。图形化的交互方式比传统的文字更直接与对象进行交流,更容易被人脑的视觉思维能力,将抽象的数据表现成为可见的图形或图像,帮助人们发现数据中隐藏的内在规律^[16]。我们在主题库的基础上,利用 Hive 做为数据源,选择 R 做为数据挖掘和界面展示工具,经过开发,将计算结果展现,如图 2。



图 2 用电行为及用电量聚类曲线

Fig. 2 Clustering curve graph of using electricity behavior and electricity consumption

根据计算展示结果我们可以发现行为异常客户,从而判断优质潜力客户与潜在风险客户最终实现用户用电行为的深度分析,针对不同用户群体制定相对应的营销策略,为电力信息的商业化利用提供基础数据。

5 结束语

目前我国电力行业数据量巨大,加强大数据技术的应用,也是为了提高数据资产的利用率,从中把握住高价值的数据,以增加电力企业的盈利水平,使风险能够得到可控、能控、再控。对外,可以使电力市场更加活跃,对内,为电力企业提供了更多的发展机会。与此同时,我国需要改变对大数据技术的看法,能从数据存储和分析的角度来实施大数据技术,从而提高数据资产的利用率,从而使我国电力企业能从数据中获取更多有用的资讯,提高我国的电力行业整体服务水平。

参考文献:

- [1] 林弘宇,张晶,徐鲲鹏,等. 智能用电互动服务平台的设计[J]. 电网技术,2012,36(07):255 - 259.
LIN Hongyu,ZHANG Jing,XU Kunpeng,et al. Design of interactive service platform for smart power consumption[J]. Power System Technology,2012,36(07):255 - 259.
- [2] 张东霞,姚良忠,马文媛. 中外智能电网发展战略[J]. 中国电机工程学报,2013,33(31):1 - 14.
ZHANG Dongxia, YAO Liangzhong, MA Wenyuan. Development strategies of smart grid in China and abroad[J]. Proceedings of the CSEE,2013,33(31):1 - 14 (in Chinese).
- [3] 赵鸿图,朱治中,于尔铿. 电力市场中需求响应市场与需求响应项目研究[J]. 电网技术,2010,34(05):146 - 153.
ZHAO Hongtu, ZHU Zhizhong, YU Erkeng. Study on demand response markets and programs in electricity markets[J]. Power System Technology,2010,34(05):146 - 153 (in Chinese).
- [4] 何恒靖,赵伟,黄松岭,等. 云计算在电力用户用电信息采集系统中的应用研究[J]. 电测与仪表,2016,53(01):1 - 7.
HE Hengjing, ZHAO Wei, HUANG Songling, et al. Research on the application of cloud computing in power user electric energy data acquisition system[J]. Electrical Measurement & Instrumentation, 2016,53(01):1 - 7 (in Chinese).
- [5] 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(01):1 - 8.
XUE Yushen, 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(01):1 - 8 (in Chinese).
- [6] 王堃,杨飞,李斌. 用电采集大数据的用户用电行为分析研究[J]. 电力大数据,2017,20(10):1 - 6.
WANG Kun, YANG Fei, LI Bin. Analysis and research of user electricity consumption behavior based on big data collected by electricity[J]. Electric power data, 2017, 20 (10):1 - 6 (in Chinese).
- [7] 彭小圣,邓迪元,程时杰,等. 面向智能电网应用的电力大数据关键技术[J]. 中国电机工程学报,2015,35(3):503 - 511 (in Chinese).
PENG Xiaosheng, DENG Diyu, CHENG Shijie, et al. Key technologies of electric power big data and its application prospects in smart grid[J]. Proceedings of the CSEE,2015,35(3):503 - 511 (in Chinese).
- [8] 孙景,李莉,胡宏力. 商业银行个人客户价值评价模型研究[J]. 金融论坛,2013,18(01):29 - 35.
LI Jing, LI Li, HU Hongli. A study of the model to evaluate the personal customer's value of commercial bank[J]. Finance Forum,2013,18(01):29 - 35 (in Chinese).
- [9] 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(03):86 - 94.
WU Weihua, Pang Jianjun, CHEN Guangkai, et al. Research on

- the development of electric power demand side response [J]. Electronic testing, 2014, 21 (03): 86 - 94.
- [10] 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.
- [11] 廖旎煊,胡智宏,马莹莹,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.
LIAO Nihuan, HU Zhihong, MA Yingying, et al. Review of the short-term load forecasting methods of electric power system [J]. Power system protection and control, 2011, 39 (01): 147 - 152 (in Chinese).
- [12] 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(04): 927 - 935.
SONG Yaqi, ZHOU Guoliang, ZHU Yongli. Pres present status and challenges of big data processing in smart grid [J]. Power System Technology, 2013, 37(04): 927 - 935 (in Chinese).
- [13] 陈新贺,宋伟,陆莹. 互动化营业厅信息集成需求分析与数据建模[J]. 电力系统自动化, 2013, 37(07): 77 - 81.
CHEN Xinhe, SONG Wei, LU Ying. Information integration demand analysis and data modeling for interactive electric power business hall [J]. Automation of Electric Power Systems, 2013, 37 (07): 77 - 81 (in Chinese).
- [14] 刘智慧,张泉灵. 大数据技术研究综述[J]. 浙江大学学报(工学版), 2014, 48(06): 957 - 972.
LIU Zhihui ZHANG Quanling. Research overview of big data technology [J]. Journal of Zhejiang University (Engineering Science), 2014, 48(06): 957 - 972 (in Chinese).
- [15] 周江,王伟平,孟丹,等. 面向大数据分析的分布式文件系统关键技术[J]. 计算机研究与发展, 2014, 51(02): 382 - 394.
ZHOU Jiang, WANG Weiping, MENG Dan, et al. Key technology in distributed file system towards big data analysis [J]. Journal of Computer Research and Development, 2014, 48 (06): 957 - 972 (in Chinese).
- [16] 王媛媛,丁毅,孙媛媛,等. 数据可视化技术的实现方法研究[J]. 现代电子技术, 2007, 30(4): 71 - 74.
WANG Yuanyuan, DING Yi, SUN Yuanyuan, et al. Research on data visualization implementation methods [J]. Modern Electronics Technique, 2007, 30(4): 71 - 74 (in Chinese).
- [17] 赵锡艺,曾茜. 电力大数据在大客户信用评级服务中的应用研究[J]. 电力大数据, 2017, 20(10): 65 - 68.
ZHAO Xiyi, ZENG Qian. The application exploration of big power data in key account credit rating service [J]. Power systems and big data. 2017, 20(10): 65 - 68.
- [18] 崔立卿,贺伟军,田晶,虞伟,张叶. 基于 K 均值聚类算法的大客户用电行为分析[J]. 浙江电力, 2017, 36(12): 47 - 52.
CUI Liqing, HE Weijun, TIAN Jing, YU Wei, ZHANG Ye. Analysis on power consumption behavior of large customers based on K-Means clustering algorithm [J]. Zhejiang Electric Power, 2017, 36(12): 47 - 52.
- [19] 梁浩波. 基于文本挖掘的用电客户诉求智能聚类研究[J]. 广东电力, 2016, 29(8): 45 - 50 + 66.
LIANG Haobo. Research on Intelligent clustering for electricity customers' demands based on text mining [J]. Guangdong Electric Power, 2016, 29(08): 45 - 50 + 66.

收稿日期: 2018 - 08 - 12

作者简介:



凌德祥(1981),男,本科,高级经济师。主要从事运营监测及大数据应用相关工作。

(本文责任编辑:范斌)

Research and practice of power client behavior analysis system based on large data

LING Dexiang, HUANG Tuo, GUAN Xiaolin, ZHAO Yan

(State Grid Fushun Power Supply Company, Fushun, Liaoning, 113008)

Abstract: With the development of our country electric power market, power supply enterprises with digital development pace to speed up the information system construction, various business systems in operation process accumulated the massive information data resources, using the data resources of digital analysis has gradually become the focus of the power supply enterprise work. In the process of continuously expanding the scope of analysis and optimizing the analysis methods, a large number of data assets are revitalized, and great enterprise value and social value gradually emerge. The application of big data technology can be used in all aspects of the current power system operation, providing strong support for the commercialization of power information. The analysis of power customer behavior based on big data technology can improve the operation efficiency of power system and provide decision basis for the actual operation of power supply enterprises. This paper analyzes the electric power industry in domestic and foreign large data application process and the present situation, according to the power of different types of customer behavior based on big data technology power customer behavior analysis system, the practice of big data technology, realize the depth of the electric power customer electricity behavior analysis, to provide customers with more quality service experience.

Key words: big data; power customers; behavioral analysis; cluster analysis