

# 电力需求侧大数据应用模型的建立

王 健<sup>1</sup>, 宋述停<sup>1</sup>, 兰俊美<sup>2</sup>, 刘冰漪<sup>3</sup>

(1. 国网山西省电力公司 信通分公司, 太原 030001; 2. 对外经济贸易大学 国际商学院, 北京 100029;  
3. 华大天元(北京)电力科技有限公司, 北京 102206)

**摘 要:**通过分析大数据的发展形势和最新研究成果,构建了大数据在电力需求侧的应用模型,以提升电力企业对电力客户的认知水平,以及电力企业的综合管理水平和经营效益。该模型以不低于 6 次/h 的频率进行数据采集,通过电网现有光纤信道和无线宽带进行数据传输,利用云存储、数据消冗技术、数据压缩技术进行数据存储,利用云计算处理技术与传统数据处理技术结合进行数据挖掘,从而得出电力需求侧分析结果,用于指导电力企业的生产、销售、服务,并通过反馈机制,对国家政策、外部环境等产生积极的影响,从而使该模型不断自我适应。

**关键词:**大数据;需求侧管理;应用模型;数据挖掘

**基金项目:**山西省电力公司重点科技项目(KJ2013-130)

**中图分类号:**TM73;F426.61 **文献标志码:**A **文章编号:**2095-1256(2014)03-0283-04

## Establishment of Application Model of Big Data in Power Demand Side Management

Wang Jian<sup>1</sup>, Song Shuting<sup>1</sup>, Lan Junmei<sup>2</sup>, Liu Bingyi<sup>3</sup>

(1. Information & Telecommunication Branch, State Grid Shanxi Electric  
Power Company, Taiyuan 030001, China;

2. Business School, University of International Business and Economics, Beijing 100029, China;

3. Huada Tianyuan (Beijing) Electric Power Technology Co., Ltd., Beijing 102206, China)

**Abstract:** This paper analyzes the development situation of and the latest research achievements on big data, and proposes an application model of big data in power demand side to help electric power enterprises improve in respects of the cognitive level about electricity customers, the comprehensive management level and operational performance. This model will collect data with a frequency of no less than 6 times/h, transfer data through the optical fiber channel and wireless broadband existing in power grid, store data using cloud technology, data redundancy technology and data compression technology, and mine data by combining cloud computing technology with traditional data processing technologies. The power demand side analysis results could be used to guide electric power generation, sales and service. The feedback in this model would positively impact on national industrial policy and other outside circumstances to make this model self-adaptive continuously.

**Key words:** Big data; Demand side management; Application model; Data mining

## 0 引言

近年来,大数据的相关研究得到世界各地及相关领域的高度重视。随着电力行业信息化不断深入,电力行业采集的数据种类和数量越来越多,除了实时动态监控数据及用电信息的采集数据外,还有大量地理信息数据、气象监测数据等,同时电力行业对数据的安全性要求也越来越高。因此,如何在电力现有数据中挖掘有效信息并进行

存储分析、如何利用分析结果指导电力企业更好的运营都成为目前亟待解决的问题,电力行业成为一个典型的大数据行业。电力大数据会在电力行业的生产、运输、销售、运营、服务等方面带来突破性的推进作用,可提升电力企业的精益化管理水平,促进电力行业的可持续发展。

## 1 电力大数据概况

在“十一五”期间,国家电网的 SG186 工程构

筑了公司一体化信息集成平台,“十二五”期间构筑了一体化企业资源计划系统(SG-ERP)<sup>[1]</sup>。随着国家智能电网建设的开展及“三集五大”管理体系的实施,电力企业对数据的管理、共享、交互、采集都提出了更高要求。电力大数据主要在电力的生产、使用、服务的过程中产生,伴随着电网投资规模的不断扩大、新技术的不断产生、设备的更新更替,电力信息数据的范围日益扩大。有必要运用高效、快捷的数据挖掘技术,探索适用于电力行业数据特征的科学方法,对现有数据进行分析处理存储。

目前,关于大数据的研究已经涉及到各行各业。在数据处理方面,数据挖掘技术、存储技术也不断地更新完善,空间数据挖掘技术、时空数据挖掘技术、统一存储系统(HUS,也称多协议存储系统)等方面都成为研究热点<sup>[2-3]</sup>。在大数据的应用领域方面,地理信息系统(GIS)领域的相关软件研究具有较多的研究成果,如面向交通网络的TransCAD、面向电力网络的SmallWorld<sup>[4]</sup>等。但是关于电力大数据的研究成果较少,尤其在关于电力需求侧方面。

## 2 电力大数据的相关研究

### 2.1 电力大数据的特征

全新的数据源和与众不同的数据分析类型构成了大数据。大数据的主要特征为:数据规模庞大而且快速增长、数据结构丰富多样、客户对数据分析的要求较高。EMC(美国易安信公司)对大数据规模的界定条件为:规模为10 Tb左右、源自不同数据源的多个大型数据集组合在一起形成Pb级的数据量。

电力大数据通常指在电力生产和使用过程中产生,数据主要在电力行业的发、输、配、变、用、售电、调度的各个环节产生。电力大数据有大数据的所有特征<sup>[5]</sup>,概括为5V(数量庞大 Volume、数据类型众多 Variety、数据处理速度快 Velocity、数据准确 Veracity、数据有价值 Value)<sup>[6]</sup>。

1) 数量庞大 Volume 随着智能电网的普及、服务质量的升级,电力行业中采集的数据量越来越庞大并且呈现快速增长的模式。

2) 数据类型众多 Variety 电力行业的数据除了传统的负荷数据、变电站数据等结构化数据外,近年来地理信息数据、资产管理信息、需求侧

用电设备类型等非结构化信息也成为电力大数据的重要组成部分,针对这部分数据的研究成为未来的发展趋势。

3) 数据处理速度快 Velocity 在电力生产、使用和销售各个环节,电力都需要时刻保持供需平衡,因此电力数据的处理速度都是实时的。

4) 数据准确 Veracity 电力行业需保证运行的安全性和可靠性,为了使后期的数据统计和决策能够与社会发展水平相适应,电力的大数据应具有很高准确性,能真实反应社会的经济和生产水平。

5) 数据有价值 Value 电力大数据应能准确反映行业的运营规律,能够起到指导电力行业生产、帮助电力企业进行经营管理的作用,达到更好地安排生产、节能减排的效果。

### 2.2 电力大数据需求侧研究的关键技术

#### 2.2.1 数据采集

数据采集是进行数据分析研究的基础性工作,而数据采集的正确性、全面性、有效性与否会直接影响数据挖掘分析的可靠性。通过数据采集,可以将杂乱的数据收集起来,再转变表达形式,为数据更好地共享、分析和可视提供必要前提。

电力需求侧的数据特点决定了数据采集技术必须具有实时性、灵活性、全面性的特点。数据采集的对象应为所有的需求侧客户;数据采集的范围应包括客户类型、用电设备类型、用电量、交费信息等;数据采集的频率应该提高至6次/h或者更高水平。

#### 2.2.2 数据传输

Pb级数据量的传输不仅要求数据传输过程的快捷,还应保证数据的完整性和安全性。目前,坚强智能电网的建设恰恰为数据的传输提供了可靠的平台。

智能电网中主要发展的通信建设设施可以成为大数据传输的主要通道。电网的光纤信道和无线宽带不仅能成为各种业务接入网的重要方式,也会成为大数据传输的有力保障。

#### 2.2.3 数据存储

越来越多的数据量和数据格式增大了企业数据管理的难度。要满足企业利用数据和信息创造价值的需求,就必须要求企业能够有合理的数据存储方式。

传统的存储方式已经满足不了大数据容量的存储要求<sup>[7-9]</sup>,大量的数据需要在存储的过程中选择合适的处理方式,减少数据量的同时达到数据保真的效果,可采用的技术有云存储、数据消冗技术、数据压缩技术等。

#### 2.2.4 数据挖掘分析

大数据挖掘分析同现有的数据分析工作最大的区别就是在于海量的数据当中分析所有数据的特征,不再用样本代表总体,而是直接对总体展开分析,提取其中隐含的、潜在的可用信息和知识。有可能在通过大量的分析后得到的可用知识量和信息量有限,但是这些少量的信息带来的价值却极高。

因此,开展数据挖掘工作首先要明确数据挖掘的目的,对开展的业务有深入的了解;其次要对存储的原始数据进行组织、清理、整合、变换等过程,然后依据现有的数据挖掘研究成果,选用合适的数据挖掘算法、工具和模型,得出具有指导意义的挖掘分析结果。不论是在对结构化还是非结构化数据的处理过程中,都要用到最基本的统计分析、关联分析、聚类、机器学习、模型仿真等传统分析技术。大数据时代的到来,使得需要处理的样本数量和类型迅速增加,催生了诸如 hadoop 这一类的云计算处理技术<sup>[10-11]</sup>。

### 3 电力需求侧的大数据应用模型研究

#### 3.1 电力需求侧大数据应用模型的建立及说明

针对电力需求侧的特征,建立大数据应用模型如图 1 所示。

1) 电力需求侧具有数据量大、数据种类繁多、更新速度快的特征。确定合适的数据收集对象,锁定客户购电量、新装、新增及变更用电数据、用电设备及类型、95598 客户服务数据、客户档案

资料这几项作为数据采集的数据源,开展对电力需求侧的用电分析工作,能够全面地反应需求侧的用电类型及用电习惯,使得分析的结果更加科学合理有效。

2) 通过传统数据采集方法与新型数据采集方法相结合的形式,主要采集手段有智能电表、无线采集系统、抄表计费手段、问卷调查、档案查询、视频手机等。通过这些采集手段能采集到不同需求侧的用电数据,为后面的数据分析提供合理的数据来源。

3) 将采集到的数据和一些国家政策法规、经济发展形势等外部环境影响因素结合在一起,并通过高效的传输手段上传至服务端,通过数据压缩、消除冗余、云存储等方式得到规范化的数据,并进行储存。

4) 一般情况下系统存储的数据较多,首先要进行数据预处理,剔除错误的、对研究目的无用的数据。同时,需要把握好目前的信息系统,在统一的数据收集、处理、整合框架下采取传统分析技术和云计算相结合的方式,得到形式多样、内容丰富的分析结果,如报表、视频、图形、报告等。

5) 得出的分析结果在存入存储系统的同时还能作为模型前端的有效反馈,在电价指导方面、需求侧用电控制、电力生产调度、用电服务质量提升方面都具有很好的指导作用,为以后的模型应用提供参考意见。

#### 3.2 应用模型建立的价值及意义

电力需求侧大数据应用模型的建立在数据采集过程中规范并且拓宽了数据采集的手段,在数据传输中充分运用智能电网的通信网络,在数据挖掘中利用先进的挖掘手段最终得出科学合理的数据分析结果。

1) 提升电网企业的数据处理能力,有效应对

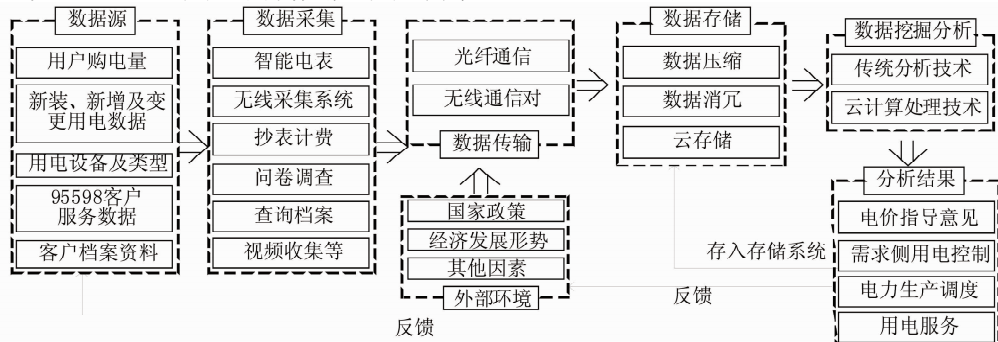


图 1 电力需求侧大数据应用模型

大数据的发展趋势 在电力需求侧大数据应用模型中,提出了数量大、规范化的数据规范,对以后的数据收集整理工作提出了更好的规范保证。同时,结合最前沿的数据挖掘分析技术,提升电网企业的数据处理水平。

2) 顺应数据发展趋势,提升电网信息价值 随着电力行业信息化的普及,越来越多的数据在电力的生产、输送、配送、使用中产生,在模型中充分应用智能电网的发展成果,提升数据通信水平以及电网的整体价值。

3) 从用电需求侧数据出发,构建电力企业的整体运维数据系统 应用模型从需求侧的数据采集入手,经过数据处理后进行客户细分,探寻客户的用电行为、发现客户的主要耗电设备。根据这些分析结果可以制定年度、月度、周度、日度的电力生产调度计划;制定具有区分性、引导性的电价,鼓励客户更加科学、节能的用电习惯;还可以通过分析提供高附加值的增值服务。通过这些服务加强电网安全检测与控制,实现电力企业的精细化运营管理和科学的需求侧管理。另外,此模型还可以通过改进应用到电力企业的电力生产、物联网构建、协同办公等方面,最终构建统一的、具有大数据特征的电力企业信息化系统。

## 4 结语

电力需求侧管理作为电力企业现代化管理中非常重要的一方面,加强需求侧管理不仅对于建设资源节约型社会具有重要意义,同时也对提升电力企业的管理水平、保证设备的正常运行具有不可缺少的作用。

(上接第 282 页)

延伸到客户侧,构建家庭户内的通信网络。通过智能交互终端、智能插座、智能家电等,实现对家用电器信息自动采集、分析、管理,实现家电经济运行和节能控制。在此基础上,通过电力大数据技术提供互动的用电服务,包括有供电信息服务、家电互动控制服务、家庭用电管理、自助缴费服务等方向。

### 参考文献:

- [1] 洪天妍,程德山. 浅谈智能电网建设与线损管理[J]. 安徽电气工程职业技术学院学报,2011(S1):138-122.
- [2] 孙德强,郑军卫,贾俊松,等. 低碳经济视角下智能电网与新

虽然大数据在电网的发展中仍处于发展的初级阶段,但是大数据在各行各业的应用中都已经展示出了巨大价值。本文提出了适应电网需求侧大数据发展趋势的应用模型,有助于提高电力企业的需求侧管理水平。

### 参考文献:

- [1] 李皎. 大数据时代到来对电力行业发展提出新要求[J]. 华北电业,2012(4):82-83.
- [2] 赵彬彬,李光强,邓敏. 时空数据挖掘综述[J]. 测绘科学,2010,35(2):62-65.
- [3] 陈捷,唐世渭,杨冬青. 面向移动环境的时空数据挖掘研究现状与展望[J]. 计算机工程与应用,2002,38(16):1-3.
- [4] 张林广. 大数据量 GIS 网络分析算法的实现和优化研究[D]. 中国科学院研究生院,2006.
- [5] 王继业. 大数据与电力企业[J]. 电力信息化,2012,10(8):7.
- [6] 闫龙川,李雅西,李斌臣,等. 电力大数据面临的机遇与挑战[J]. 电力信息化,2013,11(4):1-3.
- [7] 邢晶,熊劲,孙凝晖,马捷. 一种支持 EB 级存储的可扩展存储空间管理方法[J]. 计算机研究与发展,2013,50(8):1573-1582.
- [8] 彭成,王树鹏,贾志凯. 基于纠删码的数据消冗存储系统可靠性增强研究[J]. 计算机研究与发展,2011,48(z1):1-6.
- [9] 吴非. “云存储”助力构建海关报关单证电子档案库——基于 OpenStack 对象存储技术的集群存储系统的设计和实现[J]. 上海海关学院学报,2013,34(1):109-115.
- [10] 邓自立. 云计算中的网络拓扑设计和 Hadoop 平台研究[D]. 中国科学技术大学,2009.
- [11] 陈康,郑纬民. 云计算:系统实例与研究现状[J]. 软件学报,2009,20(5):1337-1348.

收稿日期:2014-02-21

作者简介:王 健(1973-),男,高级工程师,硕士研究生,从事通信管理工作。

(责任编辑:赵艳粉)

能源综合利用[A]. 第十二届中国管理科学学术年会论文集,2010.

- [3] 孙乔,焦艳斌. 实时数据库系统研发及其在智能电网中的应用[J]. 电力信息化,2012,11:87-92.
- [4] 贾东梨,孟晓丽. 实时数据库在用电信息采集系统中的应用研究[J]. 电力建设 2012,01:17-21.
- [5] 吴晶. 实时数据库在电能量集成系统中的应用[J]. 电力信息化,2010,07:93-97.
- [6] 中国电机工程学会信息化专委会. 中国电力大数据发展白皮书[R]. 北京:中国电力出版社,2013.

收稿日期:2014-01-22

作者简介:张 弦(1986-),男,工程师,从事实时数据库、海量数据处理、大数据分析与管理、数据交换等方面的研发及管理工作。

(责任编辑:赵艳粉)