



DIANLI DASHUJU YINGYONG ANLI JIEXI

电力大数据 应用案例解析

张素香 徐家慧 主 编



中国电力出版社
CHINA ELECTRIC POWER PRESS

DIANLI DASHUJU YINGYONG ANLI JIEXI

电力大数据 应用案例解析

主 编 张素香 徐家慧

副主编 王东升 高德荃

参 编 袁彩霞 张 东 陈 芳

阎 博 曹津平 曹 宇

王靖然 陈方正



中国电力出版社
CHINA ELECTRIC POWER PRESS

内 容 提 要

本书共7章,内容包括电力数据交互技术与应用、电力调控系统大数据应用、智能配用电大数据应用、电力信息通信运维大数据应用、人机对话及电力智能客服系统及新能源大数据应用,将视角放到大数据在电力企业应用的不同环节,通过介绍大数据实际案例来阐述在大数据研究过程中的技术路线以及贡献,不仅为读者提供了该领域的基础性知识,还提供了应用实践指南。

本书可供电力行业从事大数据统计挖掘的学习者和实践者阅读和学习,也可与其他行业广大数据分析从业人员提供参考。

图书在版编目 (CIP) 数据

电力大数据应用案例解析/张素香,徐家慧主编. —北京:中国电力出版社,2019.5

ISBN 978-7-5198-1061-0

I. ①电… II. ①张… ②徐… III. ①数据处理—应用—电力工程
IV. ①TM7-39

中国版本图书馆 CIP 数据核字 (2019) 第 041437 号

出版发行:中国电力出版社

地 址:北京市东城区北京站西街19号(邮政编码100005)

网 址:<http://www.cepp.sgcc.com.cn>

责任编辑:王杏芸(010-63412394)

责任校对:王小鹏

装帧设计:赵姗姗

责任印制:杨晓东

印 刷:三河市航远印刷有限公司

版 次:2019年5月第一版

印 次:2019年5月北京第一次印刷

开 本:710毫米×1000毫米 16开本

印 张:13

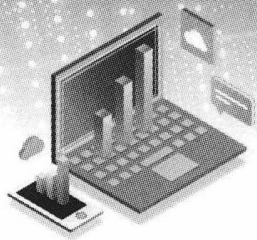
字 数:243千字

印 数:0001-2000册

定 价:50.00元

版 权 专 有 侵 权 必 究

本书如有印装质量问题,我社营销中心负责退换



前言

大数据时代，电力工业从发电、输电、配电、用电、调度以及信息通信支撑各环节带来了巨量的生产控制、经营管理的历史数据和实时数据。

本书定位于电力行业从事大数据统计挖掘的学习者和实践者，将视角放到大数据在电力企业应用的不同环节，通过介绍大数据实际案例来阐述在大数据研究过程中的技术路线以及贡献，不仅为读者提供了该领域的基础性知识，还提供了应用实践指南。

全书共7章。第1章介绍了电力大数据以及价值发现、电力大数据技术面临的挑战、电力大数据应用趋势以及大数据与云计算的关系；第2章以智能配用电信息系统为基础，依托 IEC 61850、IEC 61970、IEC 61968 标准，阐述了信息交互总线、公共信息模型、配用电信息模型形式化建模方法以及相应案例；第3章介绍了电力调度领域大数据应用的关键技术及案例，包括监控大数据分析、智能搜索、检修计划编排及调控机器人等；第4章介绍了智能配用电数据采集与处理及配用电大数据应用场景，包括负荷预测、用户行为分析、有序用电等案例；第5章介绍了电力信息通信运维大数据应用，包括系统运行状态检测、通信告警分析、通信网状态评估、线路覆冰分析等；第6章介绍了人机对话与电力智能客服系统关键技术与应用案例；第7章为新能源大数据应用，着重介绍了新能源大数据处理特点以及在风电设备异常、风电功率预测等应用场景。

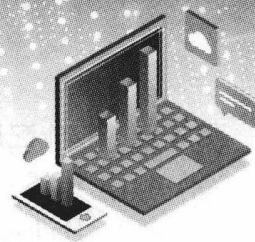
本书第1章由张素香、高德荃编写；第2章由张素香、曹津平编写；第3章由徐家慧、阎博、曹宇、王东升编写；第4章由张素香、张东编写；第5章由陈芳、高德荃、陈方正编写；第6章由袁彩霞、高德荃、张素香编写；第7章由王东升、王靖然、徐家慧编写，本书由张素香、高德荃统稿。同时感谢张翼英教授、郑蓉蓉处长在本书编制过程中提出了大量宝贵意见和指导。

希望本书能够对从事数据分析的高校师生、科研工作者能有所帮助，为我国大数据处理与人工智能发展贡献绵薄之力。

由于笔者水平有限，书中难免会有局限性和诸多不足之处，欢迎各位专家和读者指正。

作 者

2019年5月



目 录

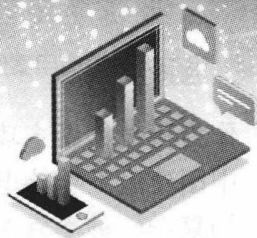
前言

第 1 章 概述	1
1.1 电力大数据	1
1.2 电力大数据的价值发现	3
1.3 电力大数据处理与分析的挑战	4
1.4 电力大数据分析应用领域	10
1.5 大数据与云计算的关系	13
1.6 小结	14
参考文献	14
第 2 章 电力数据交互技术与应用	16
2.1 数据模型	16
2.2 电力企业信息系统发展带来的问题	17
2.3 国内外电力行业数据交互规范	19
2.4 信息交互关键技术	24
2.5 一体化数据平台	45
2.6 小结	47
参考文献	48
第 3 章 电力调控系统大数据应用	50
3.1 电力调控数据概述	50
3.2 电力调控数据应用面临的挑战及大数据发展需求	52
3.3 电力调控大数据技术体系	54
3.4 典型应用案例	57
3.5 小结	71
参考文献	72

第4章 智能配用电大数据应用	74
4.1 概述	74
4.2 Hadoop 以及并行计算框架	76
4.3 系统拓扑结构	81
4.4 数据采集与处理	84
4.5 系统功能设计	91
4.6 配用电大数据应用场景	97
4.7 小结	116
参考文献	116
第5章 电力信息通信运维大数据应用	118
5.1 概述	118
5.2 电力信息通信运行数据采集	119
5.3 电力信息通信运行大数据分析	125
5.4 电力信息通信运行大数据应用场景与案例	131
5.5 小结	151
参考文献	151
第6章 人机对话及电力智能客服系统	153
6.1 概述	153
6.2 人机口语对话系统	154
6.3 自然语言理解	157
6.4 对话管理	158
6.5 自然语言生成	162
6.6 电力智能客服应用场景	165
6.7 电力智能客服系统构建	169
6.8 小结	177
参考文献	177
第7章 新能源大数据应用	180
7.1 概述	180
7.2 新能源基本知识	181
7.3 新能源大数据	183
7.4 新能源数据价值挖掘	185
7.5 应用案例解析	188

7.6 小结 196

参考文献 196



第 1 章 概 述

本章导读

智能电网的最终目标是建设成为覆盖电力系统整个生产过程，包括发电、输电、变电、配电、用电及调度等多个环节的新一代电力系统。而支撑智能电网安全、自愈、绿色、坚强及可靠运行的基础是电网实时数据采集、传输和存储，以及累积的海量多源数据快速分析。大数据的意义并不在于大容量、多样性等特征，而在于我们如何管理和分析电网，以及挖掘其价值。实现大数据驱动的智能电网，需要具备对电力大数据更强的决策力、洞察力和流程优化能力。如果在分析处理上缺少相应的技术支撑，大数据应用无从谈起。

● 本章将学习以下内容：

电力大数据及其价值发现。

电力大数据技术面临的挑战。

电力大数据应用趋势。

1.1 电力大数据

近年来，互联网、物联网、社交网络等技术的突飞猛进，引发了数据规模的爆炸式增长，大数据已经普遍存在，能源、交通运输业等领域都积累了 TB 级、PB 级乃至 EB 级的大数据。数据的迅猛增长预示着以数据驱动的“数据科学”时代到来，国内外都制定和启动了大数据研究计划，投入大量资金支持大数据研究。2012 年，美国政府宣布启动“大数据研究与开发计划”。2013 年美国电力科学研究院（EPRI）启动了两项大数据研究项目：输电网现代化示范项目和配电网现代化示范项目。德国联邦经济和技术部启动了未来能源系统技术促进计划，在 6 个示范项目中，普遍利用了大数据技术，分别从促进可再生能源发展、开发商业模式、能源服务、能源交易及传统的化石能源如何融入能源互联网等方面推出了能源互联网初步解决方案。加拿大实施了 Power Shift Atlantic 项目，其目的是通过控制热水器、空调等跟踪风力发电的变动，从而保持电力需求实时平衡，是对信息与能源交互的进一步探索。2012 年，中国在国家层面提出把“大数据”

作为科技创新主攻方向之一，打造以大数据驱动的智能电网。《“十三五”数据中国建设下智能电网产业投资分析及前景预测报告》分析认为智能电网大数据结构复杂、种类繁多，具有分散性、多样性和复杂性等特征，这些特征给大数据处理带来极大的挑战。智能电网是以物理电网为基础，将现代先进的传感测量技术、通信技术、信息技术、计算机技术和控制技术与物理电网高度集成而形成的智能电网，如图 1-1 所示。

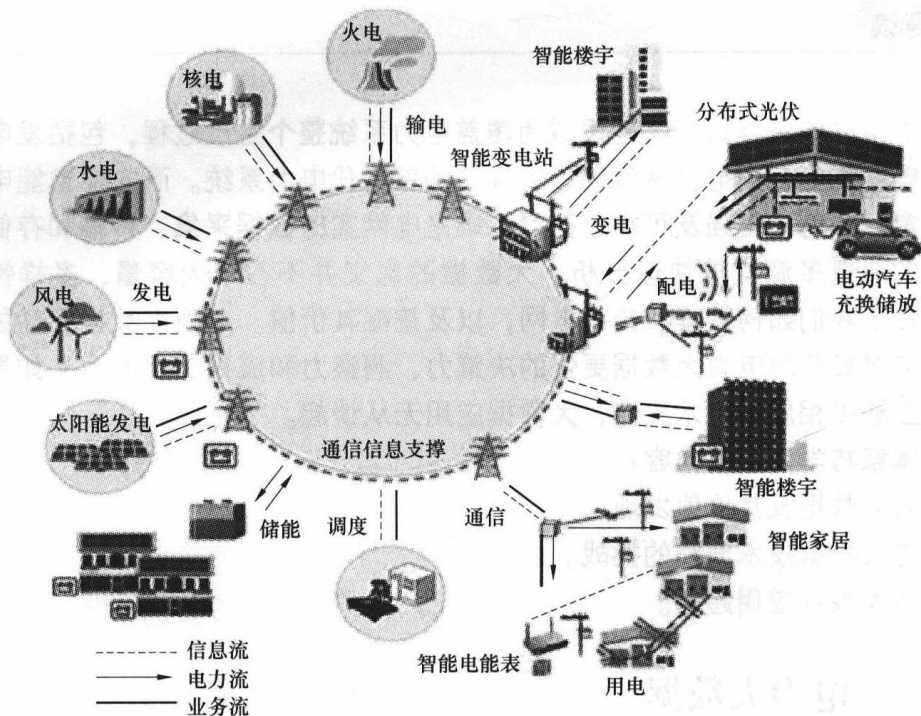


图 1-1 智能电网

智能电网建设的最终目标是成为覆盖发、输、变、配、用及调度等过程的灵活新一代电力系统，而支撑系统得以准确、安全、实时及可靠运行的基础是电力系统多源异构大数据的快速采集、响应和分析。未来智能电网既要支持个人终端用户与电网系统的交互，也要满足控制系统对电网安全稳定性的需求，智能电网中的多数应用需要海量数据处理技术的支撑。随着智能电网建设在广度和深度上的不断推进，在智能电网运行过程中会收集到系统内外的海量全景数据，形成电力大数据。从 2013 年开始，国家电网公司在输变电运行管理、智能配电网、用电与能效、电力信息与通信、决策支持等专业领域开始开展大数据应用关键技术研究，启动了多项智能电网大数据应用研究项目。2014 年国家电网公司启动了大数据应用试点研究，在电网设备状态监测、营配贯通、用电信息、客户服务信息等 7 个方面开展大数据研究和示范应用，并启动了营销大数据应用场景分析和



应用规划研究。

在电力领域，其大数据具有以下特征：

(1) 数据体量巨大。智能电网部署了大量的智能电能表和其他监测设备，产生了大量的历史数据且在迅速增长，从 TB 级别，跃升到 PB 级别。常规数据采集与监视控制（Supervisory Control and Data Acquisition, SCADA）系统 10 000 个遥测点，按采样间隔 3~4s 计算，每年产生 1.03TB 数据。

(2) 数据结构复杂，种类繁多。电网数据具有显著的分布式和异构特性，数据广域分布、种类众多，包括实时数据、历史数据、文本数据、多媒体数据、时间序列数据等各类结构化、半结构化数据，以及非结构化数据，各类数据查询与处理的频度和性能要求也不尽相同。比如，电力设备状态监测数据中的油色谱数据 0.5h 采样一次，而绝缘放电数据的采样速率高达几百 kHz，甚至 GHz。

(3) 实时性要求（速度）高且增长快。能源生产、转换和消费要求瞬间完成，电网大数据中包含着很多实时性数据，数据的分析结果也往往具有实时性要求。

(4) 数据价值高。大数据应用贯穿电力发、输、配、用、调度等各个环节，通过大数据技术应用可对各个阶段进行预测，及时发现潜在风险，保证电网的安全性和电价的经济性；大数据应用催生了更多的互联网商业模式，甚至将影响不同区域一次能源的价格体系。

1.2 电力大数据的价值发现

电力大数据大致可分为 3 类：电力系统运行和设备检测、实时状态数据，电力企业营销数据和电力企业管理数据。其中，电力企业营销数据又包括交易电价、售电量、用电客户等方面的数据。随着我国智能电网以及能源互联网的建设越来越深入，大数据技术已成为支撑智能电网安全运行最重要的方法，因此要充分挖掘电力大数据的价值。

(1) 辅助电源系统协同运行决策。电力多源系统协同优化决策的前提是对大量、翔实、可靠的信息进行及时处理，缺乏全面的信息资源将会造成决策的偏差、失误，以及管理效率的低下。具体而言，在电源端，包含了大量分布式电源、微网、储能装置，还需实现电力与供热（冷）、供气和交通系统的互动，只有建立起电力与其他能源的营、配、调一体化数据融合系统，利用大数据技术进行分析，才能保证多种能源的智能生产与配送。

(2) 支持电力安全稳定经济运行。电力系统本身是一个动态实时变化的系统，因此必须实时监测系统的运行状态，快速处理各种情况，保证系统的安全稳定运行。统计发现，大多数电网故障主要是设备故障问题引发的，通过收集设备



的全生命周期数据（实验数据、运行数据及气候环境数据等），建立设备运行模型，有利于实时评估设备状态，从而避免由于设备故障造成的电网事故。

（3）催化新商业模式的形成。在构建能源互联网过程中，数以百亿计的设备需要与网络互联互通，不断累积的海量数据有待于挖掘和运用的同时，在各国电网广泛互联状态下，电力价格属性将会催生更多的互联网商业模式，甚至将影响不同区域一次能源的价格体系。同时将涌现出基于开放的电源端、售电端活跃市场，以及在节能增效的大背景出现的新型能源公司。这一概念下的“能源互联网”，呈现的是前所未有的互联网激情和各行各业的主动融合。“互联网+”和用电市场的大力结合，将全力带动相关产业健康发展。

（4）提高能源管理水平。利用数理统计、模式识别、神经网络、机器学习、人工智能等技术，可从海量数据中挖掘出能源生产和消费中能量损失的原因，为能源生产和消费效率提升找到方向。基于天气数据、环境数据、能源互联网设备监控数据，可实现动态定容、提高设备利用率，并提高设备运检效率与运维管理水平；通过搜集、整理、分析、检索各能源消费终端、生产链等能源信息，最快地传输能源需求和能源供给，在整个能源互联网中实现能源调配，满足用户、企业、生产商、运营商等各方的需求，可最大限度地避免能源浪费与低效利用。

（5）提高调度精益化水平。根据电力生产发、供、用必须同时完成的瞬时平衡规律及电能不能大规模有效存储的特点，需要科学调度，保持电网正常运行。电力系统调度当前主要应用 SCADA 系统，以实现数据采集、设备控制、测量、参数调节，以及各类信号报警等各项功能。但随着大规模间歇性能源的接入，电力系统的结构更加复杂多变，海量、分散、异构的实时信息大量涌入数据中心，面对大量实时信息时，传统调度系统的实时性和合理性很难满足要求，无法实现电网的精益化调度。因此，基于大数据的调度技术关键问题是设计基于大规模多源细节数据的电力系统调度模型和实时流数据分析处理技术，实现电力系统调度从粗放型向集约型的转变。

电力大数据的价值发现需要处理好数据采集、数据聚合和集成、数据建模、大规模数据分析和高效计算、数据决策、确保数据安全等过程。

1.3 电力大数据处理与分析的挑战

海量数据的使用正成为电力企业在业绩方面超越其同行的一种重要方式，而持续、爆炸式增长的数据量对多源、异构、高维、分布、非确定性的数据及流数据的采集、存储、处理及知识提取提出了挑战。电力大数据的管理和分析中存在的这些挑战，需要新的方法和技术来支撑。



1.3.1 数据集可用性方面

一个正确的大数据集至少应满足以下5个性质：

- (1) 一致性。数据集中每个信息都不包含语义错误或相互矛盾的数据。
- (2) 精确性。数据集中每个数据都能准确表述现实世界中的实体。
- (3) 完整性。数据集中包含足够的数据来回答各种查询和支持各种计算。
- (4) 时效性。信息集中每个信息都与时俱进，不陈旧过时。
- (5) 实体同一性。同一实体在各种数据源中的描述统一。

一个数据集满足上述5个性质的程度是该数据集的可用性。

确保数据可用性是一项十分困难的任务，考虑到大数据的数据量大、数据产生速度快、数据类型复杂、价值大密度低4个特点，确保大数据可用性将变得难上加难，我们需要针对大数据的4个特点，解决如下大数据可用性的5个挑战性研究问题：

(1) 数据采集与传输问题。能源的生产、传输、消耗过程的数据呈爆炸性增长趋势，势必给数据传输、存储和分析带来挑战，同时，冗余的数据在一定程度上影响了系统的性能，因此，需要确保对有效数据的可靠采集。能源互联网中，电网在传输效率等方面具有无法比拟的优势，将来仍然是能源互联网中的“主干网”，能源互联网的数据传输需要依托电力通信网络作为主要工具。

(2) 数据融合问题。各业务部门的信息化系统由不同研发团队、围绕不同应用需求设计开发，采集的各类数据存在种类交叉、数据冗余、数据不一致、采集频率和存储频率差异性大、数据格式不统一等问题；同时，由于电力行业缺乏行业层面的数据模型定义与主数据管理，存在较为严重的数据壁垒，业务链条间尚未实现充分的数据共享，这为数据融合带来了很大的技术挑战。

(3) 数据质量问题。数据在体量上越来越大，但信息缺乏、数据质量低、防御脆弱、基础不牢、共享不畅等瓶颈依然存在。目前，电力领域数据可获取的颗粒程度，以及数据获取的及时性、完整性、一致性和数据源的唯一性、及时性、准确性有待提升，且缺乏完整的数据管控策略和组织管理流程。

(4) 数据识别和挖掘问题。相对于其他行业而言，电力行业的大数据资源更为丰富，对于海量数据处理难度更大。电力大数据应用是一项跨学科、跨专业的复杂性技术，现在的研究方式主要是基于团队的试验性研究方法，没有系统方法论的指导。

(5) 数据隐私和信息安全问题。大数据由于涉及众多用户的隐私，对信息安全提出了更高的要求。电力行业各单位防护体系建设不平衡，信息安全水平不一致，安全性有待提高。需引入新的防护措施，提升安全传输、安全存储的防护水平，安全防护手段和关键防护措施需要进一步加强。

1.3.2 技术方面

大数据平台作为大数据技术的一个综合载体，集成数据采集、存储、处理、分析等功能，为大数据应用提供支撑。

电力大数据平台总体上包括数据采集、数据存储、数据计算、数据分析、平台服务五个层次以及安全管理、数据管理、平台管理、广域分级协作四个保障功能，提供数据存储、计算、分析、展现能力，支撑业务应用建设，如图 1-2 所示。

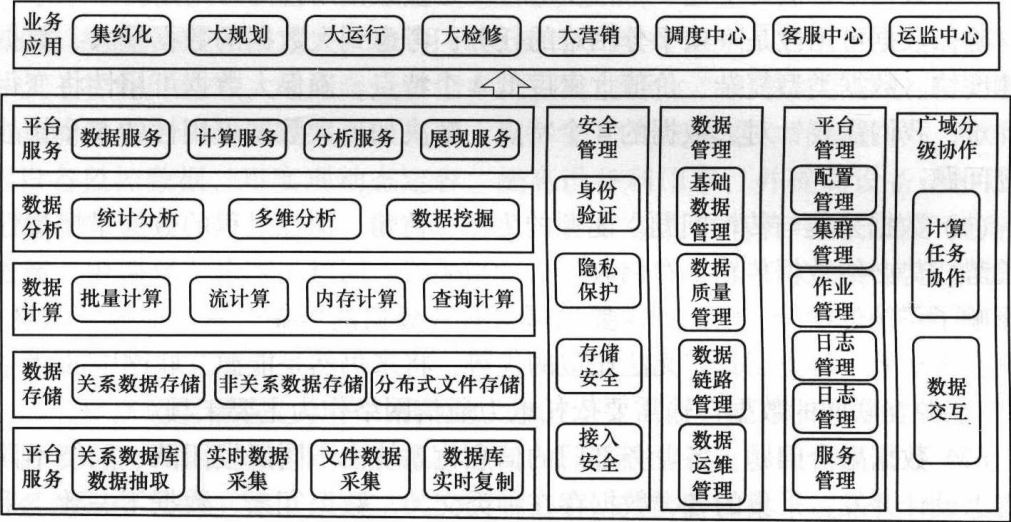


图 1-2 大数据平台功能架构

(1) 数据采集。提供关系数据库数据抽取、实时数据采集、文件数据采集、数据库实时复制等整合方式，从外部数据源导入结构化数据（关系库记录）、半结构化数据（日志、邮件等）、非结构化数据（文件、视频、音频、网络数据流等）等不同数据类型、不同时效的数据。

(2) 数据存储。负责进行大数据的存储，针对全数据类型和多样计算需求，以海量规模存储、快速查询读取为特征，提供关系数据存储、非关系数据存储、分布式文件存储三种存储方式，存储来自外部数据源的各类数据，支撑数据处理层的高级应用。通常情况下，非结构化数据存放在分布式文件存储中；半结构化数据存放在非关系数据存储中，采用列式数据库或键值数据库进行存储；结构化数据存放在关系数据存储中；实时性高、计算性能要求高的数据存储在内数据库或实时数据库中。

(3) 数据计算。对多样化的大数据提供流计算、批量计算、内存计算、查询计算等计算功能，允许对分布式存储的数据文件或内存数据进行查询和计算，对

实时数据流进行实时分析处理,实现实时决策、预警等。

(4) 数据分析。对多样化的大数据进行加工、处理、分析、挖掘,产生新的业务价值,发现业务发展方向,提供业务决策依据。

(5) 平台服务。提供数据服务、计算服务、分析服务、展现服务,将底层存储、计算、算法等能力以服务的形式,为业务系统大数据应用提供技术支撑。

(6) 数据管理。提供基础数据管理、数据质量管理、数据链路监测、数据运维管理等功能,从接入、存储、使用等方面对大数据平台中的数据进行全生命周期管理。

(7) 平台管理。必备功能模块,实时监测大数据处理全过程中的整体运行状态、资源使用情况和接口调用情况等性能指标并对关键系统险情进行告警,支持大数据组件安装、配置和状态管理,可快速扩展应用功能和能力,可实时性监控和调度任务计划。

(8) 安全管理。解决从大数据环境下的数据采集、存储、分析、应用等过程中产生的诸如身份验证、授权过程和输入验证等大量安全与隐私问题。

(9) 广域分级协作。可选功能模块,在总部与省(市)大数据平台集群间建立计算任务协作、数据交互机制,将地理上分散的、集群节点闲置的平台计算资源通过广域网络连接起来,集成为一个统一的计算环境,用来解决一些用传统的计算方法与计算环境无法解决或者解决起来较困难的计算问题。

面向电力大数据的处理与分析还存在以下挑战:

(1) 性能操作方面。目前主流的 Hadoop 批处理方式可以适应电力大数据的历史数据分析,但 Hadoop 采用先存储后计算的模式,且需要频繁的磁盘操作,难以满足对生产系统的快速响应需求。因此,电力大数据分析平台面临着严重的性能挑战,具体而言主要包括如下两个方面:

1) 利用大规模复杂细节多源数据的电力系统分析算法。在智能电网环境下,PMU (Phasor Measurement Unit, 相量测量单元)、AMI (Advanced Metering Infrastructure, 高级电表架构) 及 IED (Intelligent Electronic Device, 智能电子设备) 提供了大规模可利用的复杂细节数据,这些数据为提高电力系统的操作水平、建模技术和计划调度提供了机遇。通过现代计算机高并行计算平台,设计快速、高精度的安全稳定分析算法,能够提高数据利用率,并保证系统的实时性要求。

2) 大规模实时流数据调度、分析技术为了保证监测、调度等关键业务的时效性,大规模实时流数据的接入和分析性能至关重要。将常用数据存储于分布式内存中,并充分利用多核或众核处理器的并行技术和大内存的缓存技术以及高效的调度算法,以保证流数据的实时处理效率。

3) 缺乏对非结构化数据的处理能力。传统的关系型数据库对数据的处理只局限于某些数据类型,比如数字、字符、字符串等,对非结构化数据的支持较



差。然而随着用户应用需求的提高、硬件技术的发展和互联网上多媒体交流方式的推广,用户对多媒体处理的要求从简单的存储上升为识别、检索和深入加工,面对日益增长的处理庞大的声音、图像、视频、E-mail 等复杂数据类型的需求,传统数据库已显得力不从心。

(2) 内存计算。在电力系统中多项业务处理任务需要高性能计算技术的支持。内存计算主要用于数据密集型计算的处理,面向数据量大且需要实时分析处理的情况。针对电力大数据价值密度低的特点,依据数据的使用频率,将电力大数据分为热数据和冷数据,热数据存储在内存中,冷数据存储在磁盘上。

目前比较成熟的内存计算平台包括 HANA 和 Spark 等。HANA 是由 SAP 提出的基于内存计算技术的高性能实时数据计算平台,采用的主要技术包括内存数据读取和处理、行列混合存储、并行计算、数据压缩等。

在流数据分析平台包括 Storm、S4 方面,Storm 是具有分布式和高容错的实时计算系统,以简单、高效、可靠的方式处理流式数据,并保证消息不丢失,处理严格有序。其主要特征包括编程模型简单、支持多种编程语言、作业级容错、水平扩展、快速消息计算等。但其资源分配策略并没有考虑系统拓扑结构,任务调度机制过于简单,因此很难直接应用于电力大数据处理。电力系统中各种监测、采集装置的流数据和极端天气情况下的报警数据构成了电力流式大数据。电力流式大数据具有实时性、突发性、无序性等特征,并要求在有限时间内处理完成,不能遗漏数据,因此,流式大数据的实时处理对系统提出了更高的要求。

1.3.3 管理层面

(1) 数据管理模式有待创新。电力数据存储于不同系统中,而这些系统由不同企业和部门开发和运维管理,未考虑数据跨系统、跨部门共享和交互的需求,这给数据在跨部门、各业务环节的顺畅流通带来了困难。而大数据分析需要开展跨部门、跨业务的分析,为更好地发挥数据资产的价值,需要建立数据交互与共享机制,创新数据管理模式。

(2) 新型跨专业合作模式需要探索。大数据是交叉学科,需要电力领域专家、数据分析专家、信息通信技术专家、社会学专家等协作开展研究,但目前各专业人员还存在交流上的知识壁垒,跨专业的复合型人才较少,需要探索新型联合攻关机制。

(3) 大数据价值评估方法和体系缺乏。大数据需要大量人员和资金的投入,而大数据价值评估方法的缺失,将导致大数据价值不清晰,影响能源行业领域对大数据分析的决策。

(4) 数据集成管理机制缺乏。在电力互联网框架下,能源类型既有传统的煤

电、大型水电，又有各种可再生分布式能源，在能源传输方面既有电网与输配电线路，又有智能化的气、热管道，同时考虑信息传输通道、气象、政策等因素，数据来源广泛且结构繁杂。因此，采用集成管理方法实现对能源互联网数据的妥善保管非常必要。

1.3.4 安全方面

电力系统涉及国家能源安全，电力大数据安全是一项包括技术层面、管理层面、法律层面的社会系统工程，其保障体系的框架由组织管理、技术保障、基础设施、产业支撑、人才培养、环境建设组成。安全方面应该是研究数据源和传输的可靠性，研究信息系统故障或受到攻击时的行为，以及信息的阻塞、淹没、丢失和出错对大能源可靠性的影响。

1.3.5 应用方面

电力大数据应用必须结合具体的业务场景和目标问题，大数据研究应该遵循问题导向、需求牵引及数据共享的原则。电力大数据的应用场合涵盖发、输、变、配、用、调等电力行业的各个环节，在风电场选址、降低网损、风电并网、电网安全监测、大灾难预警、电力企业精细化运营、电力设备状态监测等领域有非常强的可实现性。随着智能电网建设的进一步推进，大数据技术在智能电网中将发挥越来越大的作用。

目前，电力大数据典型业务应用见表 1-1。

表 1-1 电力大数据典型业务

序号	业务	业务模型
1	输电	故障量预测
		负荷预测
		中期/短期配变重过载预测
		变压器油温预测
		欠费风险预测
		违约用电风险预测
2	配电	防窃电异常分析
		配网投入产出分析
		设备运行年限与故障关系分析
		电能表状态监测
		停电优化模型
		抢修效率分析



续表

序号	业务	业务模型
3	用电	用电行为分析
		用户感知度
		负荷特性分析
		物资合理供货周期分析
		设备健康状况分析

电力大数据分析的商业价值，根据服务对象的不同，可以分为以下 4 类：

- (1) 提高电网智能，提高电网运行可靠性，改善电网规划精度。
- (2) 提高资产智能，提高资产利用率。
- (3) 提高用户智能，了解用户用电行为，创新服务模式，为用户提供定制化服务。
- (4) 提高社会智能，为全社会节能减排提供定量化服务。

电网各信息系统大多是基于本业务或本部门的需求，存在不同的平台、应用系统和数据格式，导致信息与资源分散，异构性严重，横向不能共享，上、下级间纵向贯通困难，例如，电力系统中存在监控、能量管理、配电管理、市场运营等各类信息系统，大多相互独立，数据信息不能共享。使用云平台实现各独立系统的集成，可实现这些分散孤立系统之间的信息互通。另外，智能电网的基础设施规模庞大，数量众多且分布在不同地点。如何有效管理这些基础设施、减少数据中心的运营成本是一个巨大的挑战。

信息交互方面，在电网异构多源信息融合和管理中，建立类似 IEC 61968 或 IEC 61970 的信息互操作模型是必要的，研究异构数据融合与挖掘的集成方案，以及实时挖掘算法是非常急需的，但是由于系统及数据质量、挖掘算法等各方面因素，实用化程度不高。

可视化方面，可视化方法已被证明是解决大规模数据分析的有效方法，并在实践中得到广泛应用。智能电网各类应用产生的大规模数据集，如何从这些庞大复杂的数据中快速而有效地提取有用的信息，成为智能电网应用中的一个关键技术难点。

1.4 电力大数据分析应用领域

1.4.1 电力系统监控运行

电网监控运行数据具有多源、高维、先验、异构的特点，传统依靠人工经验