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5 V (Volume ;Variety ; Velocity ; Veracity ; Value)特征，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》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。

**应用研究**

**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、电力用户用电负荷预测研究、电价与激励下的电力用户响应行为研究等。

**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，计算综合评价值，并根据综合评价值进行用户分类，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MapReduc;e、内存并行化计算框架印ark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。

**2电力大数据应用分析研究综述**

**2.1电力大数据应用分析综述**

公司为深入推进“两个转变”，实现“一强三优”现代公司的战略目标，提出了建设三集五大体系的重大举措。三集五大体系的建设不仅是传统电网设施的升级和改造，还是更全面、更深入的电网运行模式和业务模式的革新。首先，三集五大体系使信息通信技术正以前所未有的广度、深度与电网生产、企业管理快速融合，三集五大等业务协同性更强，业务流、信息流趋于一致;电网业务数据将从时效性层面进一步丰富和拓展。其次，充分利用这些基于电网实际的数据，对其进行深入分析，挖掘更深层次的数据价值，便可以提供大量的高附加值服务。这些增值服务将有利于电力企业精细化运营管理，提高公司管理水平和运营效率，并可以产生很多创新性业务。

**2.2大数据在大规划应用分析**

随着配网规划业务的覆盖面进一步扩大，将逐渐积累TB级数据，对于数据的收集、存储和分析处理将提出更高的要求。届时可以利用大数据技术进一步提升如下业务应用的价值:

(1)用电量预测。基于海量历史电量数据，规划区域面积、历史人口、历史国民经济数据、三产比例等变化情况，对区域用电量进行预测，作为进一步规划设计依据。

(2)空间负荷预测。基于全网中各小区的占地面积、用地类型、容积率，行业的建筑面积负荷密度、占地面积负荷密度，小区目标年占地面积、小区目标年建筑面积，总负荷值、行业负荷值等数值，对远景年负荷进行预测。

(3)多指标关联分析。从多个外部系统(如GIS}PMS} OMS等)抓取所需数据的时间一致性切片，进行综合分析利用，从而支持规划设计。

**2.3大数据在大建设应用分析**

在大建设应用方面，系统目前每年管理的项目可以达到2 000~3 000，但由于自身数据存储和处理能力有限，在管理项目时只能保留少量的建设现场信息(仅图片，无视频)，这导致管理人员无法全面地了解项目工作细节，且由于缺乏自动化的分析与决策应用支持，这些数据也无法得到有效的利用。利用大数据技术在“大建设”基建系统中的应用主要可在如下几个方面提升:提高系统非结构化数据存储能力，扩大现场信息收集范围和频率，提升建设项目管理全面性和准确性;通过大数据技术，逐步形成并提高对大量非结构化数据的分析处理能力，基于此能力之上，采用自动对比，模式识别等挖掘技术，对建设管理的整个过程提供良好的、自动化的智能分析和监控功能。

**2.4大数据在大检修应用分析**

设备状态信息的获取手段有很多种。目前最常用的是通过各种传感器，实时或定期获取设备状态信息。视频作为一种特殊的传感器，可以作为设备状态信息获取的一种重要而有效的手段。随着视频智能分析技术的发展，目前已经能够实时准确识别变电站多种表计、刀闸、开关与隔离开关的位置、状态或读数。通过对视频数据的智能分析，还可以起到其他类型传感器无法实现的功能。利用大数据技术，主要可在如下几个方面提升:

(1)状态评价。综合设备各种基础及运行监测数据，加上视频数据智能识别技术识别出的设备状态量等，建立综合评价模型，得出设备总体健康状态。

(2)趋势分析。研究基于历史数据的趋势分析算法，建立反映设备健康状态的数学模型，对设备故障进行趋势分析，掌控设备风险。

(3)实时报警。基于视频监控的设备自动巡检技术和安防技术，发现外观变化、表计变化、发热缺陷、非法入侵、物体靠近、现场烟火等设备健康危害因素，进行实时报警，保证设备正常运行。

(4)检修、技改、大修决策支持。根据设备评价结果，优化检修策略，为技改、大修计划制定、筛选评审提供决策依据。

**2.5大数据在大运行应用分析**

设备状态管理是公司“大运行”重要业务模块，设备状态管理可对计划类应用进行校验并提供完善的建议，为电网调度提供辅助决策，提升电网安全性与可靠性。设备状态管理目前只根据其他调度业务进行当前设备状态断面的记录，没有对历史设备断面的查询及未来状态的预测分析。随着调度业务规范的进一步完善，对设备状态模块的提升需求日趋迫切。设备状态管理基于电网设备台帐信息、设备拓扑信息、设备遥信遥测信息等相关信息。可以进行设备状态的以下三类大数据应用:

(1)历史时刻查询。实现设备数据的海量存储，直接快速查询到某时刻的设备状态及设备的操作历程。

(2)设备数据质量分析。综合设备信息得出当前设备状态断面，进行系统间设备状态的互检比对，提醒其他系统或功能进行检查与核对。

(3)未来时刻断面预测。基于海量存储历史状态数据，自动总结电网调度规则，在任意时刻断面的基础上，自动收集涉及设备状态变化的计划类信息，进行智能编排，演算出未来一段时间内的调度操作历程，最终预测电网在某一时刻的设备断面。

**2.6大数据在大营销应用分析**

大营销重点是优化现有营销组织模式，拓展面向智能化、互动化的服务能力，加快用电信息采集系统建设，科学配置计量、收费和服务资源，实现计量检定配送、95598服务等业务向省级集中，构建营销稽查监控体系，推行统一的业务模式、服务标准和工作流程。目前公司大营销业务相关的支撑系统主要有SG186工程核心之一的营销业务系统，以及负责用户侧电能量自动采集的用电信息采集系统、支撑营销数据日常统计工作的辅助决策系统等。其中营销业务系统负责业扩报装、电费核算收缴等日常业务的流程管控，用电信息采集系统负责用户电量信息采集为营销业务系统电费核算提供数据，辅助决策系统用于完成统计报表制作等功能。另外，为了满足营销业务新的需求，相关网省公司建设了诸如智能客户档案管理系统，用于将供电合同、用户身份证明等纸质资料扫描存档，方便检索和查找。

大营销相关系统数据量的急剧增加，给数据存储、分析处理、统计计算带来极大挑战。根据大营销业务、信息系统及其数据的现状和需求，大数据可以在以下方面促进大营销的进一步提升:营销海量数据的高效分布式存取和并行计算。深度挖掘营销数据价值，从管理、客户服务、业务拓展等不同的方向促进营销业务提升。其中管理方向主要有线损分析、用电异常分析、计量和采集设备的智能化运维等;客户服务方向主要为用电客户提供有针对性的用电优化建议，促进用户节约用电;业务拓展方向积极策划数据类产品，引导和促进公司由“只卖电”向“也卖电”发展。

**2.7大数据在运监中心应用分析**

国家电网两级运营监控中心以服务公司战略为目标，以核心资源和主要经营活动管控为重点，以信息化手段为支撑，打造全天候、全方位、全流程综合管理运营监测平台。利用大数据技术在运监中心中的应用主要可在如下几个方面提升：

(1)构建大数据背景下的在线监测、在线分析和在线计算工作台，满足日常监测、协调控制与综合管理需要。建设总部和省公司两级运营监测(控)中心，实现对公司经营管理24小时即时在线监测分析，及时发现公司运营过程中的异动和问题并自动预警。

(2)跨业务多专业关联分析。为了满足运监跨业务跨专业分析和挖掘需要，在大数据环境下，迫切需要构建多种业务模型和分析挖掘算法，形成大数据模型库和算法库，利用聚类和模式识别技术，实现运营监测多业务关联分析，如对营销和财务收支情况建模进行关联分析。

(3)运监数据及监测对象的质量诊断体系。通过在“大数据”环境下，对海量异构数据进行实时监控分析，在数据处理全过程中不断提升数据质量，持续改进数据可用率，为拓展传统的决策分析提供保证。

**2.8大数据在客服中心应用分析**

客户服务中心是公司优化整合服务资源，打造“全业务、全天候，服务专业化、管理精益化、发展多元化”的供电服务平台。客服中心核心业务系统分为95598智能互动网站、95598业务支持系统、基础支撑平台三个独立系统。95598业务支持系统采用全国一级部署模式。按数据的种类来分，主要有两类，一类是档案、交易和GIS等结构化数据，全国约有3 TB，日增约1 GB;还有是语音音频类非结构化数据，以文件形式保存，全国

约有近10 TB，年增7 TB，并有加速增长趋势。利用大数据技术在客户服务中心应用可在如下几个方面提升：

(1)服务质量实时监控。采用大数据技术，实现对服务通话的准实时/实时监控和分析，通过对客服过程的实时质检，最大程度保障通话服务质量。

(2)热点问题集中处理。采用大数据技术挖掘一段时间内的用户关注热点，并及时答复和处理，提高客户满意率。

(3)座席答案智能推送。人工搜索答案及时性很难满足要求，利用大数据技术实现答案的智能匹配并自动推送给座席。