



电力数据挖掘在电网内部及各领域间的应用

耿亮, 吴燕, 孟宪楠

(北京国电通网络技术有限公司 北京 100070)

摘要:信息技术得到了空前的发展,电力数据呈指数规模增长,数据挖掘扮演着越来越重要的角色。主要介绍了数据挖掘的交叉产业标准及数据挖掘的主要方法。结合电力行业可获得的数据,分析了数据挖掘在电力系统内部在电网设施建设规划、动态安全评估、电力负荷预测、电力系统故障诊断及优调度选择等方面的应用,同时分析了数据挖掘在支撑智慧城市发展过程中各行业的应用,为城市发展提供助力。

关键词:电力系统;数据挖掘;应用

doi: 10.3969/j.issn.1000-0801.2013.11.022

Application of Power Data Mining Within and Between Grid Fields

Geng Liang, Wu Yan, Meng Xiannan

(Beijing Guodiantong Network Technology Co., Ltd., Beijing 100070, China)

Abstract:With an unprecedented development of info-technology, the power data has an exponential increasing, and the data mining (DM) plays a very important role in this process. The CRISP-DM and its primary approaches were presented. Combined with the power industry data, the applications of grid facilities construction planning, dynamic security assessment, power load forecasting, power system fault diagnosis and optimal scheduling options with the power system were analyzed, and then the applications in some other industries were also analyzed, supporting the development of the city.

Key words: power system, data mining, application

1 引言

随着信息技术的不断发展,各行业大量的数据随之产生,为从大量的数据中挖掘出潜在价值,为生产生活提供支撑,数据挖掘技术逐步发展,并在多个领域得到初步应用。如在金融市场,通过在业务开展和资金流动频率、额度等方面,对用户群体进行识别和目标市场分析;在零售业,通过对顾客购买记录、货物进出等的分析,更精确地挑选潜在客户。

对于电力系统这样一个庞大的系统,源源不断地产生着大量数据记录,随着 T-MIS、GIS、SCADA、ERP、电力生产数据系统、电力营销数据系统、用电信息采集系统等各类

电网信息系统的广泛应用,需要新的技术手段来挖掘出深层次的规律,支撑电力系统安全、稳定运行。同时结合其他行业数据,也为工业生产、城市管理、城市规划等决策提供可靠支撑。本文以跨行业数据挖掘标准流程为出发点,介绍了数据挖掘的方法以及电力数据挖掘结果对电力内部及城市运行中各领域的应用。

2 数据挖掘流程和方法

数据挖掘(data mining, DM)可以应用在各个不同的领域。数据挖掘是一个利用各种分析工具在海量数据中发现模型和数据间关系的过程,使用这些模型和关系可以进行预测,寻找数据间的潜在关联,发现忽略因素。



2.1 数据挖掘流程

CRISP-DM(cross-industry standard process for datamining, 跨行业数据挖掘过程标准)为跨行业的数据挖掘提供了一个完整周期的综合展示,形成闭合环,是不断反馈、不断完善的过程,具体包括6个阶段:商业理解、数据理解、数据准备、建立模型、模型评估和模型部署,如图1所示。

(1)商业理解:从商业的角度了解项目的要求和最终目的,以目的为切入点,并将这些目的与数据挖掘的定义以及结果结合起来。

(2)数据理解:对可用的数据进行评估和数据标准的选取。

(3)数据准备:对可用的原始数据进行一系列的组织和清洗,使之达到建模需求。

(4)建立模型:在这个阶段可能需要选择和应用不同的建模技术,同时校准模型参数。该阶段主要建立的模型包括关联规则、分类与预测、聚类、异常检测等。

(5)模型评估:该阶段重点挖掘结果与商业目的的吻合度。

(6)模型部署:模型部署可以是生成一份报告或者是一个可重复的数据挖掘过程。

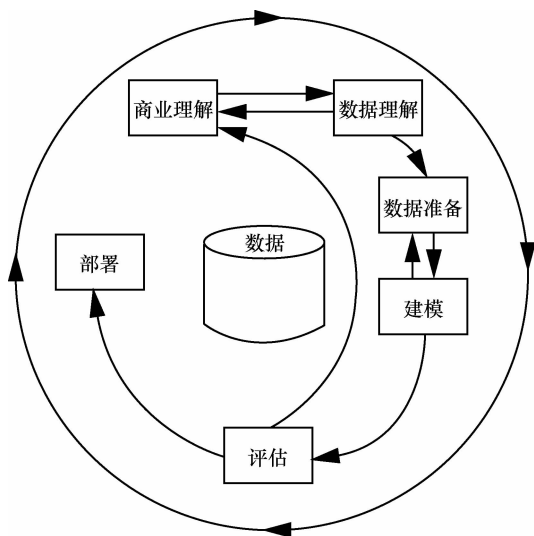


图1 跨行业数据挖掘标准流程(CRISP-DM)

2.2 数据挖掘方法

数据挖掘涉及的学科领域和方法很多,以下是几种常用方法。

- 统计分析。统计分析提供了一种基于多维度、大量历史数据的判别方法,有回归分析、方差分析等。统计分析是根据已有数据对事件运行规律的探索,推

断出因果线性关系,计算出线性回归方程。统计分析的优势在于对有确定影响因素的事件,有相对准确的影响因子效力分析。

- 神经网络。神经网络是模仿人脑神经网络的结构和工作机制而建立的一种计算模型,包括前馈式网络、反馈式网络和自组织网络模型。神经网络的优势是其具有良好的自组织、自适应、并行处理、分布存储和高度容错等特性。
- 决策树。决策树一般都是自上而下生成的。每个决策或事件都可能引出两个或多个事件,导致不同的结果,把这种决策分支画成图形,很像一棵树的枝干,故称决策树。决策树对比神经网络的优点在于可以生成一些规则,对解决多维数据的分类问题很有效,但随着分支数的增多,管理难度会加大。
- 粗糙集。粗糙集理论认为知识是有粒度的,利用相对核的概念进行知识关联性分析。该理论能够在缺少相关数据的先验知识的情况下,基于数据的分类能力,解决模糊或不确定性数据的分析处理。
- 可视化。可视化提供给用户一个人机交互的平台,使数据展示更加清晰、简洁。

3 电力数据挖掘应用

在电力行业,随着智能电网的深入发展,信息通信系统作为城市的“神经网络”与电网生产管理深度融合,可获取的电力数据包括电力基础类数据和生产营销类数据,基础类数据包括GIS信息、用户基础信息、电力线路及设备监测数据等,生产营销类数据包括发、输、变、配、用、调度环节可监测到的电流、电压、有功、无功、电量等方面的数据以及用户用电负荷、交易电价等。充分利用这些基于电网的实时和历史数据,可提供大量高附加值的服务,实现更科学的需求侧管理。在电网内部,不仅能够从生产、营销、管理等领域提高服务质量,还能确保电网的安全可靠运行;在电网之外,无处不在的电网数据与行业外数据整合,为城市管理、服务、能源、环境和产业发展等方面提供决策支撑。

3.1 电力数据挖掘在系统内部的应用

近年来,数据挖掘在电力系统中的应用取得了一定的成果,从电网设施建设规划、动态安全评估、电力负荷预测、系统故障诊断及调度选择等方面开展挖掘分析。

(1) 电网设施建设规划

基于可视化技术,将用电情况与GIS结合,以街区为

单位,展现出电力大数据地图,这个地图不仅为城市和电网规划提供了直观有效的负荷数预测依据,也可以依据展示出的停电频率较高、过载较为严重的街区进行电网设施的优先升级改造,配套辅助设施的部署,对于用电量较大、基础设施完善的地区,优先考虑智能电网示范工程子项建设。

(2)动态安全评估

动态安全分析是分析在系统受到大扰动后,对重新回归稳定运行状态的必要预防和补救措施选择。基于决策树方法,对发电调度的动态安全性进行测试,包括对暂态稳定、电压稳定和频率稳定的测试。决策树方法可以通过提取规则来获得必要的控制措施和最优化配置,使电力系统得到安全的动态控制。

(3)电力负荷预测

电力负荷预测是电力调度系统的重要考核指标,预测结果直接影响用户侧用电及电力部门运营的经济型。基于决策树方法,充分考虑影响电力负荷的影响因素,如节假日调整因子、用户历史用电记录、季节气候变化等其他一些相关属性,将其运用于日负荷预测。另外,基于粗糙集方法,针对短期负荷预测具有明显周期性特点的用能区域,使用历史数据特征推理方法对短期负荷进行预测。

(4)电力系统故障诊断

电力系统故障诊断是通过利用有关电力系统及其保护装置的广泛知识和继电保护等信息来识别故障的元件位置(区域)、类型和误动作的装置,其中最主要的是故障元件的识别问题。基于粗糙集方法,实现中小型电力系统和变电站的故障诊断,同时依据输电系统故障发生的时空关联性,运用优化相似原理,诊断高压输电线路系统故障;基于决策树方法,以较快的速度实现变电站故障诊断及自动获取。

(5)优化调度选择

电力系统的经济调度是以电网安全运行调度为基础,调节降低电力传输过程中的线损为目标。基于模糊理论,通过运用风电场的电力系统动态经济调度模型,解决由于风电场输出功率预测难、波动性大给传统机组经济调度带来的困难。

3.2 电力数据挖掘的跨领域应用

电力数据除对电力内部业务作支撑以外,在其他领域也能发挥巨大的作用,涉及城市的多业务多领域,从基础设施、公共服务、城市管理、能源使用以及产业发展等方面

全方位地支撑着智慧城市的发展,电力通信网搭建起城市的“神经网络”,使城市运行更高效、居民生活更便捷、产业发展可持续。

(1)城市基础设施智能化

城市建设规划管理:在城市规划领域,分布式能源、应急灾害指挥系统及输变电设备状态监测装置可以提供空气温度、湿度、风速、气压、噪声等环境监测指标,结合电网GIS数据,可对一定区域环境展开多角度监测,并根据区域地理环境,为具体区域全方位提供数据监测指标,配合城市环境对标数据,对某一区域属性做出明确标识,为城市规划提供数据支持,通过数据指标可反映出某一区域是否宜居、是否适合建立工业园区或者应侧重绿化建设公园等。

(2)公共服务普惠化

交通安全管理:IBM曾与电力公司和汽车公司合作,通过对电动汽车充电站、充电桩等电动汽车充电基础设施的数据收集以及电池资产信息的管理数据,分析电动汽车应在何时、何地获取动力及其对电力供应的影响等基本问题。基于汽车的电池电量、汽车的位置、一天中的时间以及附近充电站的可用插槽等,建立预测模型,将这些数据与电网的电流消耗及历史功率使用模式结合起来,通过分析来自多个数据源的巨大实时数据流和历史数据,确定司机为汽车电池充电的最佳时间和地点,并为充电站的设置点规划提供数据依据。

(3)能源利用高效化

清洁能源综合管理:综合分析对外部环境监测的海量数据,给用户提供最优的方式实现清洁能源发电及合理利用,平衡电能供应和需求,实现清洁能源与电网的有效互动,使清洁能源的利用更高效。

能源的高效利用:通过对企业实时及历史数据的整理分析,能够分析预测企业用电情况。如2012年8月,海盐多邦印刷有限公司力调电费呈现上升趋势,高于常规数值。电力部门数据监测系统监测异常后,通知企业,经检验发现企业的无功补偿设备已损坏,为企业减小了后续损失。

(4)城市管理精细化

区域群体行为分析:美国加州大学洛杉矶分校的研究者就根据“大数据”理论设计了一款“电力地图”,将人口调查信息、电力企业提供的用户实时用电信息和地理、气象等信息全部集合在一起,制作了一款加州地图。该图以街区为单位,展示每个街区在当下时刻的用电量,甚至还可以将这个街区的用电量与该街区人的平均收入和建筑物



类型等相比照,从而得出更为准确的社会各群体的用电习惯信息。并据此对小区、楼宇的入住情况进行监测和管理,进行群体空间行为分析。如能够监测识别区域房屋的入住率和空房率,将有助于相关监管部门对区域行业及群体行为特征的调查、取证和管理。

(5) 产业发展现代化

- 精细化产业管理。根据耗能比重,划分企业监控级别,结合企业行业分布,分析高耗能产业,为企业转型和高新技术改造提供决策支撑和依据。
- 产业要素集聚情况分析。通过用电负荷实时测量及统计等,反映出城市产业运行状态及产业类型分布,通过用能和机械比例数据,反映出生产要素的集聚情况。
- 产业园区经济生产分析。通过将工作时段的用电量曲线和峰谷平不同时间段电价曲线进行耦合,结合企业实际用工情况,计算出可接受的最小电费总额及相应用工方案,为企业提供更加经济合理的开工时段建议,并以报表、折线图展示。
- 生产故障监测管理。通过对电压、电流、谐波等基本电力数据的计量,对生产过程中损耗异常的那部分线路进行管理和改造,实现能效管理。
- 节能空间管理。电力系统提供全国、各行业、各地区的能效对标库,涵盖国内外典型企业、先进企业或者大型能耗企业能耗数据,随着接入企业的增加逐步完善系统标杆库,为各单位提供最全的能效对标库,支持企业根据实际情况选择对标标杆。用电企业将自身的单位产值能耗、总能耗水平与标准(国家标准、区域内行业标准)进行对标分析,将企业内部不同生产线、车间、班组的产品单位产值能耗与标准进行对比分析,以 GIS 图形式展现不同企业的能耗等级。在对数据分析的基础上对节能量、节能收益等信息进行预测并以图形形式展示。分析内容包括预期年度 GDP 节能量及节能收益、预期年度地区节能量及节能收益、预期年度行业节能量及节能收益、预期年度用电客户节能量及节能收益等。

4 结束语

本文介绍了数据挖掘技术以及各种技术在电力行业内的应用。与传统的统计数据相比,数据挖掘具有对潜在问题和规律更高的预见性、计算能力和管控能力。电力

系统的数据挖掘除了对系统内发、输、变、配、用、调度等各个环节进行有效支撑和辅助决策以外,与系统外部数据的整合和挖掘是智能电网对智慧城市支撑的重要体现,从数据层面辅助城市实现智慧和绿色的特点,数据挖掘不仅能够为政府管理、人们生活和产业发展决策作支撑,也可通过反馈机制不断收集数据,更新数据库,提升数据挖掘的可用性、确定性,进一步为城市发展提供支撑和助力。

参考文献

- 1 Han J W, Kamber M. Data mining concepts and techniques. Canada: Morgan Kaufmann Publishers, 2006
- 2 Sforza M. Datamining in a power company customer database. Electric Power Systems Research, 2000, 55(3): 201~209
- 3 陈星莺,张晓花,瞿峰等. 数据挖掘在电力系统中的应用综述. 电力科学与技术学报,2007(9)
- 4 曾碧波,黄小红,赵健伍等.TPS SCADA 系统中基于电能分析的数据仓库模型设计. 工业控制计算机,2009(9)
- 5 周广惠. 对企业用电量分析与预测的数据挖掘. 装备制造技术,2009(9)

[作者简介]



耿亮,男,北京国电通网络技术有限公司策划中心产品创新部经理,主要研究方向为智能电网关键技术。



吴燕,女,北京国电通网络技术有限公司策划中心产品创新部二级经理,主要研究方向为智能电网大数据存储与挖掘。



孟宪楠,女,北京国电通网络技术有限公司策划中心产品创新部产业分析师,主要研究方向为智能电网支撑智慧城市体系。

(收稿日期:2013-08-01)