

基于系统学习算法的电价执行规范性管理

李跃华¹, 卢峰¹, 黄益军¹

(国网浙江长兴县供电有限公司, 浙江 湖州 313100)

Standard management of electricity price implementation based on system learning algorithms

LI Yuehua¹, LU Feng¹, HUANG Yijun¹

(State Grid Zhejiang Changxing County Power Supply Co., Ltd., Huzhou 313100, China)

摘要:现行电价体系存在着电价差,客户高价低接的违约用电行为侵害着供电企业的合法利益。传统的人工逐户用电检查手段不仅效率低下,而且因检查人员自身原因存在着“人情电”、“利益电”等“小微”权利腐败问题。利用营销系统和采集系统海量实时用电数据,通过数据稽查模型的建立,形成了电价精准闭环管理机制。

关键词:电价执行规范性管理;电价获利空间;居民生活电价;农排电价;自学习

Abstract: There is a price gap in the current electricity price system. The violation of contract of high price and low connection by customers violates the legitimate interests of power supply enterprises. The traditional manual household-by-household power inspection method is not only inefficient, but also has the problem of “human power” and “interest power” corruption due to the inspectors own reasons. The precise closed-loop management mechanism of electricity price is formed by using the massive real-time electricity data of marketing system and acquisition system, and establishing the data inspection model.

Key words: standardized management of electricity price; electricity price profit space; resident living electricity price; electricity price of farm drainage; self-learning

0 引言

随着电力体制改革持续推进、一般工商业连续降价,公司利润增长空间被不断压缩。国网、省公司在今年“两会”上提出了建设泛在电力物联网,提升经营效率效益的工作要求^[1]。电价目前分成大工业、一般工商业、农业、居民生活4大类,由于电价政策中的“交叉补贴”问题,各类电价之间有着鲜明的电价差。部分高电价客户通过种种手段,使自己的所用电价全部或部分执行较低电价,形成高价低接,严重违反了国家的电价政策,也侵害了供电企业的合法权益。同时,供电部门的部分工作人员也获得了权力寻租的空间,产生了“人情电”、“利益电”的问题,滋生了腐败。

针对上述问题,从分析历史高价低接类违约用

电案例的数量分布和电价差获利空间入手,确定以农排电价和居民生活电价为研究主题。依托营销系统和采集系统的海量客户用电数据,从用电量、用电时间分布、特殊事件或关键节点影响等3个维度寻找农排用电、居民生活用电与其他电价用电之间的特异性,形成了农排用电和居民用电被侵占两大类的电价稽查理论算法模型。同时,引入“自学习”功能,即通过系统对某类电价的典型用户用电数据不断监控分析,系统模仿生物的学习功能自动对该类电价的特异性和变化趋势进行总结分析,找出该类电价近期的用电特性并同时转化成系统算法,实现电价监控管理分析的人工智能化(AI)。

1 电价管理现存问题

传统的电价管理模式一般采取装表前认真勘察,用电过程中强化稽查,发现异常后立刻处理^[2-3]。但随着社会在不断发展,用电客户数量不断增多,现场用电环境错综复杂,客户用电性质频繁变化,电价执行规范性问题的重点不断转移。在这样的形势下,供电部门仅依靠有限的用电检查人员现场排查,对客户利用电价差非法获利的行为根本无法形成有效的遏制。同时现行的电价监控体系中如

收稿日期:2019-08-01;修回日期:2019-09-03

基金项目:国家自然科学基金资助项目(51677076);国家电网公司总部科技项目“电力体制改革形势下的输电网规划方法研究”

This work is supported by National Natural Science Foundation of China(No.51677076); Science and Technology Project of State Grid Corporation “Study on the planning method of transmission and distribution network under the form of power system reform”

居民大电量、农排12个月连续有电量等手段过于片面单一,一方面误报率高,严重浪费供电部门本来就不充裕的用电检查力量,另一方面非常容易被针对性规避,造成监控手段失效。目前的营销系统和采集系统里拥有着大量的客户用电数据,却因为缺少相关的电价判断算法而无法被有效利用。通过海量的客户用电数据,找出各类电价的用电特性,建立相应用电稽查模型,最终形成一个全面、实时、高效、准确、有自学习功能的客户电价管理监控系统,为电价管理工作提供有效的抓手。

2 电价监控管理系统

根据电价易被侵占的农排用电和居民生活用电出发,从营销系统和采集系统的客户用电数据为主要数据来源,设计相应电价稽查模型,实现电价实时监控管理。

2.1 系统设计框架

第一阶段:通过人工分析历年来高价低接类违约用电实际案例,结合营销系统和采集系统客户用电数据,总结各类电价的特异性规律,转化为系统电价稽查算法。

第二阶段:开发系统的自学习功能,通过系统自主对某类电价的用电数据不断监控分析,系统模仿生物的学习功能自动对该类电价的特异性和变化趋势进行总结分析,找出该类电价近期的用电特性并同时转化成系统算法,实现电价监控管理分析的人工智能化(AI)。

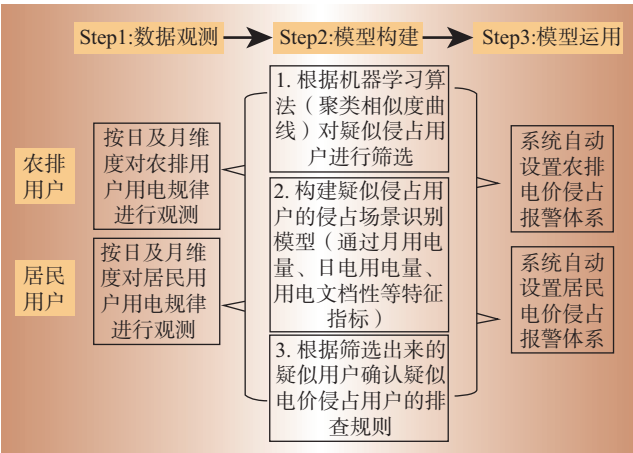


图1 框架图

Fig. 1 Framework diagram

2.2 电价获利空间分析

侵占农排、侵占居民生活电价获利空间分析分别如图2和图3所示,居民生活用电阶梯差分析如图4所示。

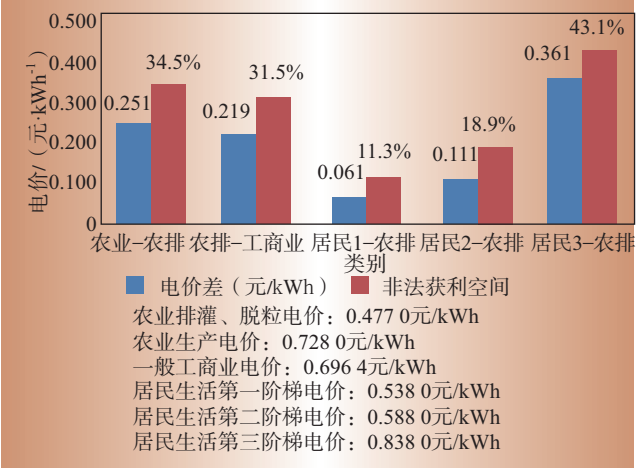


图2 侵占农排电价获利空间分析(低压)

Fig. 2 Analysis of profit space from embezzling electricity price of farm drainage (low voltage)

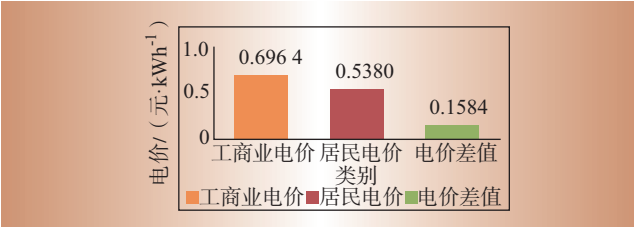


图3 侵占居民生活电价获利空间分析(低压)

Fig. 3 Analysis of profit space from embezzling resident living electricity price (low voltage)

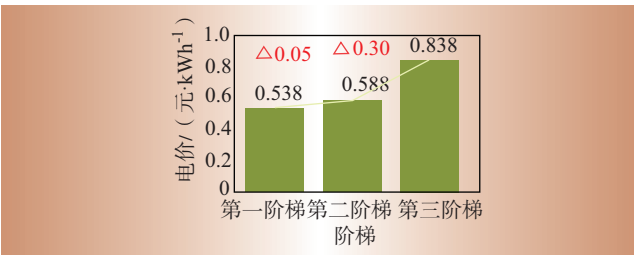


图4 居民生活用电阶梯差分析(低压)

Fig. 4 Ladder difference analysis of residential electricity consumption (low voltage)

2.3 基于用电信息采集设备的台区识别技术

(1) 正常农排用电数据分析

农排用电用于粮食灌溉和抗洪排涝,根据浙江省实际情况,农排用电基本集中在每年4—9月份,其他月份用电量接近于零,且用电月日电量不平衡,气温及天气情况对其影响极大。

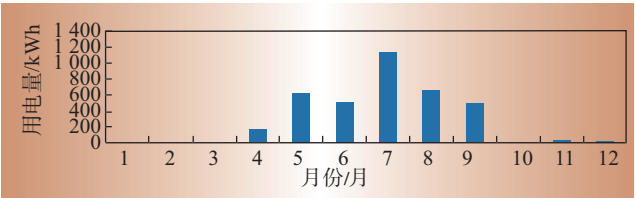


图5 典型农排用电年度数据曲线(月电量分布)

Fig. 5 Annual data model of typical farm drainage (monthly electricity distribution)

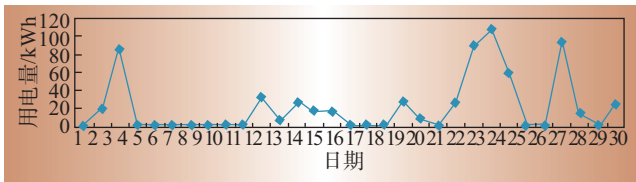


图6 典型农排用电月度数据模型(日电量分布)
Fig. 6 Monthly data model of typical farm drainage (distribution of daily electricity)

(2) 其他用电与农排用电数据特性对比分析
其他用电与农排用电数据特性对比分析如图7所示。

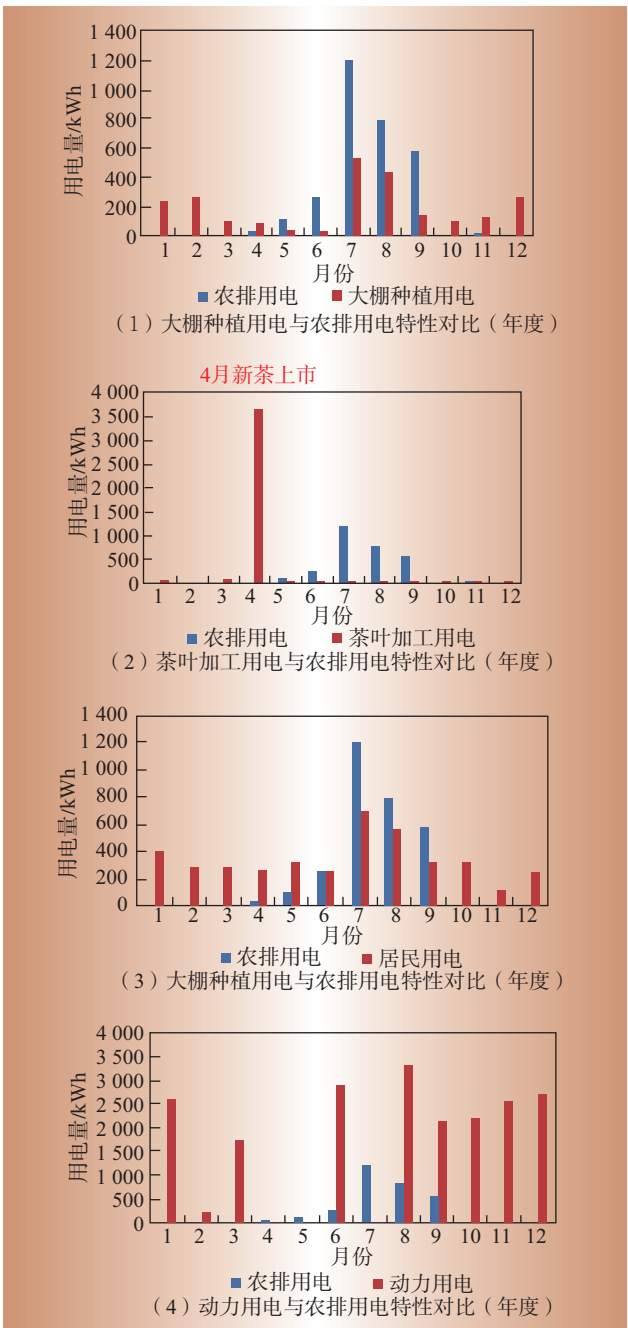


图7 其他用电与农排用电数据特性对比
Fig. 7 Comparisons of data characteristics between other electricity consumption and farm drainage

(3) 按月观测年农排用户用电规律

农排用户月用电量规律拟从以下4个方面开展观测:① 总体农排用户的月用电量分布,平均到户的月用电量分布;② 农排用户的用电高峰月份的用户占比;③ 划分农排用户的各月用电区间,分析各月各用电区间用户占比,观测各月份农排用户明细到户的主要用电量区间;④ 农排用户连续用电月份数分布。

(4) 按日观测农排用户用电规律

农排用户日用电量规律拟从以下2个方面开展观测:① 总体农排用户各月各日总用电量分布,平均到户的各日用电量分布;② 农排用户各月连续用电的日天数分布。

2.4 农排电价稽查模型构建

(1) 结合数据观测及业务经验构建指标

近年各月份用电量,近年连续用电月份数,日用电量稳定程度(方差),月份最大用量,最小用电量,月连续用电天数

(2) 运用机器学习算法区分正常农排电价用户及疑似侵占农排电价用户

1) 聚类,自动把用户分成K类,再整理合并成2类用户(正常农排电价用户,疑似侵占农排电价用户),常见的五大聚类算法有:K-Means(K均值)、均值漂移、基于密度的聚类(DBSCAN)、用高斯混合模型(GMM)的最大期望(EM)聚类、凝聚层次聚类,以下介绍K-Means聚类的具体算法及步骤。

K-Means 算法是一种简单的迭代型聚类算法,采用距离作为相似性指标,从而发现给定数据集中的 K 个类,且每个类的中心是根据类中所有值的均值得到,每个类用聚类中心来描述。对于给定的一个包含 n 个 d 维数据点的数据集 X 以及要分得的类别 K ,选取欧式距离作为相似度指标,聚类目标是使得各类的聚类平方和最小,结合最小二乘法和拉格朗日原理,聚类中心为对应类别中各数据点的平均值,同时为了使得算法收敛,在迭代过程中,应使最终的聚类中心尽可能的不变,具体算法如下:

① 首先选择一些类/组,并随机初始化它们各自的中心点。中心点是与每个数据点向量长度相同的位置。这需要我们提前预知类的数量(即中心点的数量);② 计算每个数据点到中心点的距离,数据点距离哪个中心点最近就划分到哪一类中;③ 计算每一类中中心点作为新的中心点;④ 重复以上步骤,直到每一类中心在每次迭代后变化不大为止。也可以多次随机初始化中心点,然后选择运行结果最好的一个。

2) 确定农排标杆用户,将其他用户与农排标杆

用户对比,根据用电量时间序列指标及其他特征指标计算相似度,把与标杆用户相似度高的用户划分成正常农排用户,把与标杆用户相似度低的用户划分成疑似侵占农排电价用户。计算相似度算法有:

① 从距离判断相似性:闵可夫斯基距离、标准化欧氏距离、马氏距离、巴氏距离;② 相关性指标:皮尔逊系数、杰卡德相似系数与距离;③ 度量轨迹相似度:LCSS(longest comon subsequence)、frechet distance、DTW(dynamic time warping)、edit distance。

(3) 疑似农排电价侵占用户的侵占场景识别模型构建

通过机器学习算法区分正常农排电价用户及疑似侵占农排电价用户后,将疑似侵占农排电价用户单独剥离出来,结合业务经验及相关用电监控指标,基于特殊月份用电情况分析,连续用电的行为分析,与天气的关联分析,构建疑似农排电价侵占用户的侵占场景识别模型。

1) 基于特殊月份用电情况分析定位炒茶、鱼塘打水侵占场景。

结合业务经验,对农排电价每年春节前1—2个月内出现日大电量,主要为鱼塘打水用电(农业生产电价)侵占农排电价,因此需要结合春节时间,找出疑似用户中春节前1—2个月内出现日大电量的用户(日大电量需要通过数据观测,制定相关指标定义日电量达到多少度定义为日大电量用户),标记鱼塘打水侵占用户。对农排电价每年3—4月份(新茶上市期间)出现大电量,主要为炒茶加工用电(农业生产电价)侵占农排电价。因此分析疑似用户3—4月的日用电量情况,并结合炒茶标杆用户的用电规律,运用相似度指标,在疑似用户中找出3—4月的日用电量异常的,及与炒茶标杆用户的用电规律相似的用户,标记为炒茶侵占用户。

2) 基于连续用电的行为分析定位居民生活用电和动力用电侵占场景。

根据业务经验,非农排用电月份(4—9月份)连续2个月以上出现电量,日电量连续且较为平均,主要为)居民生活用电(居民生活电价)侵占农排电价;连续2个月基本每日均出现大电量,且日电量较为平均主要为动力用电(一般工商业电价)侵占农排电价。因此需要结合业务经验及前期的数据观测制定连续用电月份数,连续用电月份用电量波动率,连续用电月份月均用电量等相关指标连续用电的行为指标。根据连续用电的行为指标及居民生活用电,动力用电标杆用户的用电规律,找出疑似用户中居民生活用电侵占农排电价及动力用电侵占农排电价的用户。

3) 与天气的关联分析。

闷热天气日电量突增的情形可能为鱼塘增氧用电侵占农排电价,因此可结合每日天气相关指标与日电量指标进行关联分析,找出关联度极高的用户,标记为鱼塘增氧用电(农业生产电价)侵占农排电价用户。剩余其他疑似未标记的用户进一步分析用电规律,探索业务特点,确定是否为农排侵占用户及侵占场景。

2.5 居民生活电价数据分析

(1) 结合数据观测及业务经验构建指标

居民生活用电客户数量巨大,占整体客户数比例极高,用电量较其他用电相对较少且比较平均,气温对其用电量影响较大,如冬季严寒取暖和夏季高温抗暑。同时,因阶梯电价的存在,低压居民阶梯三电价高出阶梯一电价0.3元/kWh,电价上涨幅度达55.8%,对居民用电成本影响极大。典型居民生活用电年度数据曲线和月度数据曲线分别如图8和图9所示。

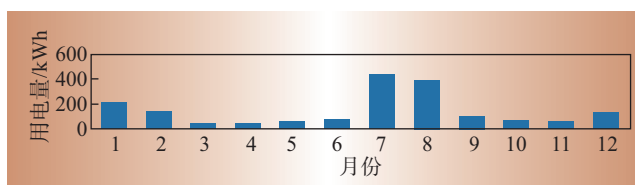


图8 典型居民生活用电年度数据曲线(月电量分布)

Fig. 8 Annual data model of typical residents electricity consumption (monthly electricity distribution)



图9 典型居民生活用电月度数据曲线(日电量分布)

Fig. 9 Monthly data model of typical residents electricity consumption (distribution of daily electricity)

(2) 其他用电与农排用电数据特性对比分析

一般工商业及其他用电与典型居民生活用电特性如图10所示。

(3) 按月观测近年度居民用户用电规律

居民用户月用电规律拟从以下2个方面开展观测:① 总体居民用户的每月用电量分布,平均到户的月用电量分布;② 划分居民用户各月的月用电量区间,观测各月各区间用电量占比;③ 居民用户的用电高峰月份分布。

(4) 按日观测居民用户用电规律

居民用户日用电规律拟从以下2个方面开展观测:① 居民用户年度各月各日用电量分布,主要的日用电水平区间;② 居民用户每日用电量波动情况。

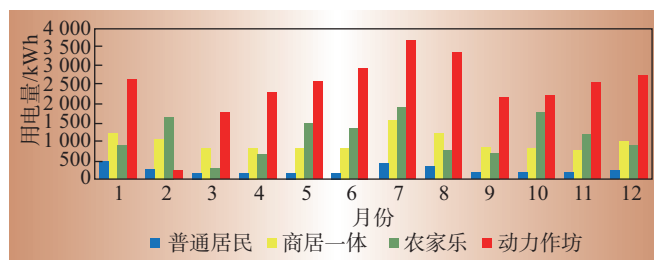


图10 一般工商业及其他用电与典型居民生活用电特性对比(年度)

Fig. 10 Comparison of electricity characteristic between general industry and commerce and other electricity consumption and typical residents electricity consumption (annual)

(5) 按负荷观测居民用户用电规律

居民用户负荷用电规律拟从以下3个方面开展观测:① 根据客户负荷特性判断用电设备性质是否符合居民生活;② 根据充电电价专用表计客户负荷性质判断是否符合电动汽车负荷曲线;③ 判断客户负荷峰值是否符合居民生活用电。

2.6 居民电价稽查模型构建

(1) 结合数据观测及业务经验构建指标

近年各月份用电量,日用电量稳定程度(方差),月用电量稳定性(方差),各季度用户电量波动情况,年度中最大用电量月份,年度中最小用电量月份,累计用电量超4 800 kWh月份。运用机器学习算法区分正常居民生活电价用户及疑似侵占正常居民生活电价用户

1) 聚类,自动把用户分成 K 类,再整理合并成2类用户(正常居民生活电价用户,疑似侵占居民生活电价用户)。农排部分已有算法的详细介绍,该部分不再重复介绍。

2) 确认一批典型的居民生活电价用户,将其他用户与这一批标杆用户对比,根据用电量时间序列指标及其他特征指标计算相似度,把与标杆用户相似度高的用户划分成正常居民生活电价用户,把与标杆用户相似度低的用户划分成疑似侵占正常居民生活电价用户。计算相似度算法如下。① 从距离判断相似性:闵可夫斯基距离、标准化欧氏距离、马氏距离、巴氏距离;② 相关性指标:皮尔逊系数、杰卡德相似系数与距离;③ 度量轨迹相似度:LCSS(longest comon subsequence)、frechet distance、DTW(dynamic time warping)、edit distance。

(2) 疑似居民生活电价侵占用户的侵占场景识别模型构建

通过机器学习算法区分正常居民生活电价用户及疑似侵占居民生活电价用户后,将疑似侵占居民生活电价用户剥离出来,结合常识单独对居民生

活电价用户进行进一步的筛选判断(如年度用电量及低的用户筛选等),缩小疑似用户范围。

制定相关用电行为的监控指标,从日用电量,月用电量与月度的关系特征,日用电量与日期的关系特征,构建模型识别其他电价侵占居民生活电价场景。

进一步制定居民合表用电、居民充电设施用电的用电负荷特性分析及普通居民生活用电能表计在年度4 800度阈值的电量突变属性,构建模型识别规避居民生活电价阶梯电费场景。

(3) 确认疑似居民生活电价侵占用户的排查规则

根据排查出来的居民生活电价侵占用户,再进一步分析其用电规律,并对比正常居民生活电价用户的用电规律,制定疑似居民生活电价侵占用户的排查规则。制定规则后并进一步根据规则进行疑似用户查找,检验规则的有效性。

3 电价自学习功能开发

3.1 广义电价自学习功能开发

根据营销系统内对客户电价的设置,对海量客户用电数据分析,通过系统学习自动生成各类电价的特异性算法规则,并根据规则形成系统电价稽查模型。

3.2 狭义电价自学习功能开发

根通过设置相对数目的标杆电价客户群,通过分析标杆客户群用电数据,获得相应电价的特异性算法规则,并根据规则形成系统电价稽查模型。

(1) 狭义电价自学习功能培养

① 营销系统根据客户电价类别、行业类别等标签进行细化分类(如:XX用户、一般工商业、商业、小区商居一体、便利杂货店);② 细化分类后进行人工现场核实确认;③ 调取该类客户的用电数据,并对同一类型进行统计分析或联动稽查;④ 找出该类电价特异性并转化为系统算法。

(2) 狭义电价自学习功能正确性验证

① 用生成的系统算法对存量用电数据进行对比核查;② 根据系统算法找出电价异常用户,并生成异常用电清单;③ 作为系统学习对象的标杆用电客户数据与原算法出现较大差异,生成异常用电清单。

4 结论

本文探讨基于系统学习算法的电价执行规范性管理,解决在电价差利益驱动下客户进行高价低接的违约用电问题。该课题的稽查模型已在长兴

县供电公司逐项投入应用。2018年11月,长兴公司根据课题农排电价异常中部分子项稽查模型进行突击排查中,共发现84户电价执行异常,完成和电价修正并补收差额电费2.38万元,违约使用电费4.76万元,现场排查准确率达到60%以上。通过研究获得以下结论:

(1) 成功利用大数据实现电价监控管理实时系统化

利用客户用电大数据资源,形成实时化的电价监控管理系统。系统通过各类电价稽查系统模型,实时对营销系统和采集系统内的客户用电数据进行监控排查,找出与实际电价不符的违约用电行为。

(2) 对电价类违约用电精确打击,提升了供电企业经营效益

充分利用营销系统和采集系统的各类用电数据,建立算法模型,实现疑似用户精准定位,不仅有效打击了各类高价低阶的违约用电行为,保障了供电企业合法权益,同时有效降低了电价稽查工作的成本,契合国网公司建设“泛在电力物联网”提升经营绩效的发展战略。

(3) 构建了电价常态稽查机制,强化了电力职工队伍建设

根据系统检测数据,公司启动了常态稽查、反馈和二次核查闭环常态化管理机制,不仅有效打击了高价低接的违约用电行为,遏制“人情电”、“利益

电”,有效规范“小微”权力,打造一支风清气正的员工队伍。D

参考文献:

- [1] 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(4):927-935.
SONG Yaqi, ZHOU Guoliang, ZHU Yongli. Present status and challenges of big data processing in smart grid [J]. Power System Technology, 2013, 37(4):927-935.
- [2] 黄海涛,张粒子,乔慧婷,等. 基于变密度聚类的居民阶梯分段电量制定方法[J]. 电网技术,2010,34(11):111-116.
HUANG Haitao, ZHANG Lizi, QIAO Huiting, et al. A method to determine step-shaped electricity consumption levels for residential area based on variable-density clustering [J]. Power System Technology, 2010, 34(11):111-116.
- [3] 刘小渊,庄华松. 电力营销稽查管理中存在的问题及策略探究[J]. 中国高新技术企业,2014(31):162-163.
LIU Xiaoyuan, ZHUANG Huasong. Problems and strategies in electric power marketing audit management [J]. China High Tech Enterprises, 2014(31):162-163.

(责任编辑 邵海雯)