

新电改背景下电力用户信用风险预警模型与评价方法

余 培¹,刘其辉¹,石 城²

(1. 华北电力大学 新能源电力系统国家重点实验室,北京 102206;

2. 国网上海市电力公司 市北供电公司,上海 200072)

Warning model and user credit risk evaluation method under the background of new power grid reform

YU Pei¹, LIU Qihui¹, SHI Cheng²

(1. State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China; 2. Shibe Electric Supply Company, State Grid Shanghai Municipal Electric Power Company, Shanghai 200072, China)

摘要:在新电改背景下,正确合理地评估电力用户的信用风险及提出预警,是决定供电企业进一步发展的重要因素。为了进一步地开展电力用户履约信用评价,帮助供电企业权衡电力交易时的风险水平,研究了电力用户信用风险评价原则和指标,并从该指标入手,基于随机森林算法构建了电力用户的信用风险预警模型,从企业市场风险、财务风险、信用风险、运营风险和结算管理风险对电力用户的信用风险进行了评价,有望完善电力市场交易机制,促进资源合理高效分配,使得电力行业在信用体系的保障下健康发展。

关键词:新电改;信用风险;预警模型;随机森林算法

Abstract: In the background of the new power reform, correct and reasonable assessment of the credit risk of power users and early warning are important factors that restrict the further development of power supply enterprises. In order to carry out credit evaluation of power users and help power supply enterprises to balance the risk level of power trading, based on random forest algorithm, the evaluation principles and indicators of power users are studied according to the evaluation indicators of credit risk in power market. The early warning model of power users evaluates the credit risk of power users from enterprise market risk, financial risk, credit risk, operational risk and settlement management risk. It is expected to improve the trading mechanism of electricity market and promote the rational and efficient allocation of resources, help power industry develop healthily under the protection of credit system.

Key words: new power reform; credit risk; warning model; random forest algorithm

0 引言

自20世纪末以来,世界各地掀起了电力市场化的改革热潮,供电企业作为发电环节和用电环节的桥梁,能够预测消费者的负荷变化趋势,同时提供满足想应电量和电能质量需求的商品。鉴于电能的特殊性,目前电力市场交易形式主要为赊销的方式,这种方式既为供电企业带来了巨大的用户市场,同时供电企业也承担了非常大的信用风险^[1]。近些年,由于客户经营不善或信用破产而拖欠电费新闻屡见不鲜,而这都将导致供电企业的电费收入没能得到保证,资金流转不充分,影响了供电企

业的经营规划和投资效率。因此,正确合理地评估和管控用户信用风险及预警,对供电企业地正常经营有着不可忽视的影响^[2]。为了能够降低用户未能正常履约给供电企业带来的经营风险,合理、综合地评估客户履约信用成为供电企业提高经营效率所面临的一个重要议题。

在我国新电改形势下,供电企业有必要选取更能综合反映用户履约水平的信用指标,在原有的缴费、安全信用评价的基础上进行体系扩充,以便更全面、系统地评价自身面临的信用风险。目前,国内学者参考国外电力交易市场经验,通过用户报价水平、缴费行为等因素,结合不通风建模方法评价用户信用水平。文献[1]综合分析了市场清除问题和最优报价策略问题组成的两层不确定约束规划模型下不同公司竞价策略及背后的不确定因素;文献[3]和文献[4]分析了国外售电侧开放市场情况下的实施成效,并对我国未来的售电公司策

收稿日期:2019-09-31;修回日期:2019-11-25

基金项目:国家重点研发计划项目(2018YFB0904000);国家电网公司科技项目(SGJB0000TKJS1801242)

This work is supported by National Key Research and Development Program of China (No.2018YFB0904000); Science and Technology Project of State Grid(No.SGJB0000TKJS1801242)

略进行了预测;文献[5]从缴费行为对用户信用度进行了分析;文献[6]根据层次分析法和德尔菲法建立了信用评价模型,分析了贵州电力市场内不同交易主体的信用等级。

在此基础之上,期望能够在用户违约现象发生之前,通过构建履约风险评价体系、加强用户信用分析与评估,并根据不同信用水平的用户针对性地制定不同的营销策略。这种将履约风险后问题解决,转化为风险前预防的做法,将违约问题消除在萌芽状态。本文构建了用户的多层信用指标体系,通过层次分析法解构了各指标所占权重,并采用随机森林算法结合历史信用水平搭建了预警模型,对用户的信用水平进行了预测,在新电改背景下,对用户履约信用合理评价并预警,能够帮助供电企业与用户开展电力交易时的风险水平,提高企业的经营效率,完善当前的电力市场体系,促进电力企业和社会的和谐发展,并能够为其它行业提供借鉴。

1 新电改背景下电力用户信用风险评价

1.1 电力用户信用风险综合评价流程

一般来说,完整的企业风险预警管理流程主要分为目标设定、风险识别、构建风险评价指标体系、构建风险预警模型、设置预警区间、建立风险防范机制6个步骤。而电力市场信用风险是指在电力市场信用交易过程中,由于各种不确定因素的作用,交易的一方在一定时间内违约的可能性以及另一方造成可能损失的程度,包括违约的概率、可能损失的数量以及损失的易变性3个方面内容,其中可能损失的程度处于最重要的位置。

电力行业的信用综合评价基础是建立完善的电力行业的信用评级指标体系。对于这方面的研究,国内外总结了影响信用评级的众多因素。为确保电力企业的评级结果的公平、客观、准确性,有必要建立一套符合我国目前电力市场实际的指标体系,有效、公正的评级指标原则,对电力企业信用评级的准确性及高效性具有重要的现实意义。

1.2 电力用户信用风险评价指标及权重计算方法

个体电力用户信用评价指标包括5个维度,即市场风险、财务风险、信用风险、运营风险以及结算管理,每个维度的指标均有一定的含义,分别是:① 企业市场风险:即企业或用户面临的市场环境风险;② 企业财务风险:即用户的盈利能力以及偿债能力;③ 企业信用风险:即企业信用、企业注册信息、企业高管信用、电力用户是否实施了危害市场的策略;④ 企业运营风险:合同执行、合作方满意度、过往事

故情况;⑤ 结算管理风险:重点包括企业的电费缴纳情况、违约金缴纳情况等。

电力用户信用评价指标权重设计采用层次分析法,而层次分析法是将待评价的因素分解成目标层、准则层和指标层等层次。因此本文在此基础上进行定性结合定量的分析方法^[7]。

一般的层次结构可以分成3层,包括为最高层、中间层和最底层。最高层是目标层,只有一个元素。它作为评价对象的预定目标,是层次结构中的目标层。中间层中包括了达到目标的每一个中间环节,中间层又可以包括多个层次。最底层是实现目标和准则的具体步骤,是具体定性和定量的元素。

由于层次结构只反映了元素之间的关系,无法反映元素之间的重要性。因此,需要两两比较元素之间的重要性,最后确定每个元素在准则层的占比。构造判断矩阵,可以反映专家对各元素相对重要性的认识,一般采用1—9及其倒数的标度方法来构造矩阵。

本文设 m 个电力市场成员,每个成员有 n 个评价指标,由此形成的指标矩阵为

$$D=(x_{ij})_{m \times n} \quad i=1,2,3,\dots,n \quad (1)$$

式中: x_{ij} 为第 j 个评价对象的第 i 个指标取值。

(1) 计算每行元素乘积

$$M_i = \prod_{j=1}^n x_{ij} \quad j=1,2,\dots,n \quad (2)$$

(2) 计算每行方根

$$\overline{W}_i = \sqrt[n]{M_i} \quad i=1,2,\dots,n \quad (3)$$

(3) 计算判断矩阵特征向量

$$W_i = \frac{\overline{W}_i}{\sum_{j=1}^n \overline{W}_j} \quad i=1,2,\dots,n \quad (4)$$

(4) 计算判断矩阵最大特征值

$$\lambda_{\max} = \sum_{i=1}^n \frac{(AW)_i}{nW_i} \quad (5)$$

此外,还需要对判断矩阵进行一致性检验,一致性是指判断矩阵中的元素满足如下关系

$$x_{ij} = x_{ik}/x_{jk} \quad i,j,k=1,2,3\cdots \quad (6)$$

当判断矩阵的阶数 $n=1,2$ 时,判断矩阵具有完全一致性。当判断矩阵的阶数 $n>2$ 时,运用随机一致性比率 CR 检验其一致性

$$CR = \frac{CI}{RI} \quad (7)$$

当 $CR<0.10$ 时,认为判断矩阵具有很好一致性,否则就要调整矩阵,使之具有很好一致性。

在式(7)中, CI 为一致性指标

$$CI = \frac{\lambda_{\max} - n}{n - 1} \quad (8)$$

式中: $\lambda_{\max}$ 为判断矩阵的最大特征根, n 为判断矩阵

的阶数。

RI 作为随机一致性指标,是指多次(通常为 500 次以上)重复进行随机判断矩阵的特征值计算之后取算数平均数得到的一致性指标。随机一致性指标 RI 取值范围如表 1 所示,表中 RI 取值是 1—15 阶重复计算 1 000 次后得出的平均随机一致性指标,并据此得出了不同指标在最终评价结果中所占的权重。

表1 随机一致性指标 RI 取值范围Table 1 Ranges of random consistency indicator RI

阶数 n	1,2	3	4	5	6	7	8
RI	0	0.52	0.89	1.12	1.26	1.36	1.41
阶数 n	9	10	11	12	13	14	15
RI	1.46	1.49	1.52	1.54	1.56	1.58	1.59

2 电力用户信用风险预警模型

随机森林(random forest)算法是一种基于分类树以及随机空间子理论的自学习算法,可以用于分类以及回归。在用于回归时,利用组合回归树的方法,使得预测结果表现很好且不容易过拟合,提高了预测精度。该算法不仅可以用于处理大数据,而且预测速度较快,因此本文采用随机森林算法建立预测模型。

根据前文对电力用户企业信用风险影响因素的分析和信用风险指标体系的构建,可以通过历史数据通过信用调查、数据记录和专家评估等方式获得各个评价指标数据,由于输入数据的量纲不一致,而且预测模型中输入输出的数据单位不一致,因此对输入的数据需要进行归一化处理,归一化后数据不存在量纲,且取值范围在0—1之间,归一化处理的公式为

$$x_{\text{nor}} = \frac{x_{\text{source}} - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \quad (9)$$

式中: x_{nor} 为归一化结果; x_{source} 为原始数据; x_{max} 和 x_{min} 分别为原始数据的最大值和最小值。在得到归一化的数据之后, 便可以利用C均值模糊聚类算法进行聚类处理, 通过最优聚类数目的判定, 为后续预测做铺垫。

同时,为了降低过拟合、随机误差对预测结果的影响,一般利用bootstrap方法进行训练集抽取^[8],bootstrap法是一种不放回的抽样方法,通过随机抽样生成决策树的样本数据。即输入的训练集有可能多次出现在同一颗决策树的训

练集中,也有可能从未出现。最终利用CART决策树算法分别由上至下训练每一棵树,直到满足要求后结束,并汇总输出结果。之后通过测试集算出每一个叶子结点的权重并储存。

在对信用风险预警时,通过 bootstrap 算法抽取的到训练集及其特征后,选取 CART 算法作为回归决策树的生成算法,因此节点的分裂依据为最小均方差,其计算方法见下式

$$m = \min_{\underline{A}, s} [\min_{c_1} \sum_{x_i \in D_1(A, s)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in D_2(A, s)} (y_i - c_2)^2] \quad (10)$$

式中: s 为当前节点的所有训练集; A 为当前节点被抽取出的特征集。训练集 s 根据特征 A 划分成为子集 D_1, D_2 , 通过遍历 A 的取值, 算得子集 D_1, D_2 输出值 y_i 的最小均方差之和, 最后遍历所有的属性得到最小均方差的取值 m , 以及对应的属性、对应的属性的值, 得到该节点的生长信息。而后对生成的每一个子节点重复上述过程, 直到达到终止条件。终止条件设置为: ① 达到设置的树深度 d ; ② 节点上的样本少于最少样本数 s ; ③ 最小均方差达到阈值 m_0 。

在生成决策树后,利用类似的方法,对抽样得到的所有训练集进行训练,便形成随机森林,随机森林的预测模型就此建立。

根据前面对随机森林的预测模型的建立过程的介绍,总结出随机森林预测步骤,随机森林预测流程图如图1所示。

随机森林预测模型的搭建步骤如下:

(1) 首先利用C均值模糊聚类方法对原始数据处理,得到相似企业信用水平后,将其分为两个部分:训练集 D 与测试集 S ,记 D 的样本容量为 N ,记 D 中样本的属性容量为 M 。

(2) 从训练集 D 中利用 bootstrap 重抽样的方法抽取样本容量也为 N 的训练集, 重复 K 次, 便可以得到 K 个训练集 $\theta_1, \theta_2, \dots, \theta_K$, 每个训练集都可以

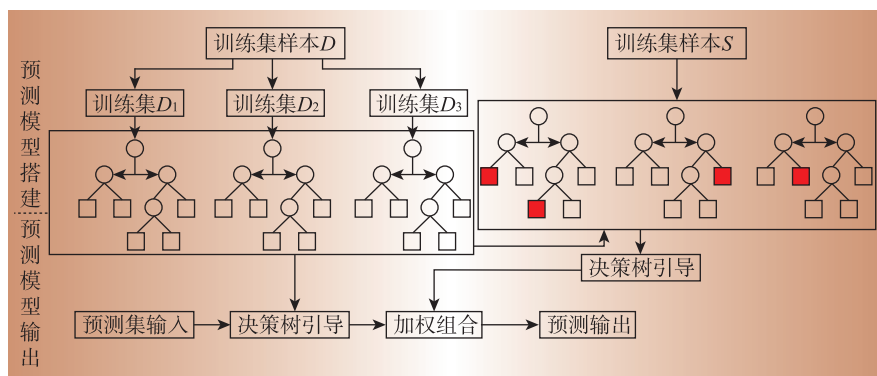


图1 随机森林预测流程图

Fig. 1 Random forest prediction flow chart

生成与之对应的决策树 $\{T_1(\theta_1), T_2(\theta_2), \dots, T_k(\theta_k)\}$, 因此 K 为随机森林的决策树数。

(3) 利用训练集 $\theta_1, \theta_2, \dots, \theta_k$ 生成相应的 CART 决策树, 在生成决策树过程中, 对于每个分裂节点, 从 M 维属性特征中不放回地抽出 m 维属性, 遍历这些属性以及属性对应的值, 计算最小均方值, 得到最优的分割判据。

(4) 训练每一棵决策树, 直到达到终止条件。

利用建立好的随机森林模型, 代入测试集数据:

(1) 输入的数据经过归一化处理后, 与相似日数据进行匹配, 代入与之相对应的相似日所形成的随机森林预测模型 $\{T_1(\theta_1), T_2(\theta_2), \dots, T_k(\theta_k)\}$ 。

(2) 分别带入模型的每一棵决策树中进行生长, 培养过程的分裂判据与建立模型类似。直至结束。之后统计每个非 0 叶子节点的权重, 权重计算公式由(11)给出

$$w_i = \frac{n(i \in l(\theta))}{N[i \in S]} \quad (11)$$

(3) 得到叶子节点的权重后, 根据(12)可以计算决策树的权重。

$$w_i = \frac{\sum w_i Y_{ii}}{\sum Y_i} \quad (12)$$

(4) 保存决策树的根与节点、权重等信息后, 随机森林预测模型建立完成。

随机森林预测模型建立完成后, 进行预测时的步骤如下:

(1) 输入的数据经过归一化处理后, 与相似日数据进行匹配, 代入与之相对应的相似日所形成的随机森林预测模型 $\{T_1(\theta_1), T_2(\theta_2), \dots, T_k(\theta_k)\}$ 。

(2) 培养每棵决策树, 培养过程的分裂判据与建立模型类似。得到每棵树的预测结果后, 利用由测试集生成的权重数据, 进行加权平均便可以得到最终的预测结果。

$$y = \sum_{i=1}^K w_i Y_i \quad (13)$$

3 基于随机森林的信用风险等级预测

运用历史数据和同期专家评分值, 对 A 典型水泥企业的信用风险时间观测数, 进行多变量时序预测模型的构建, 得到预计的下一个时间点的预测变量值。然后, 运用反映水泥企业信用特点的预测变量值进行信用风险评价, 得到预测的水泥企业在下一时间点的信用类别。针对水泥企业的信用类别, 进行相应的信用预警等级划分和信用风险防范。

采集某水泥企业的时间观测数据如表 2 所示。

各指标的变量数值越大表示信用风险越低, 违约的可能性越小。

表 2 某水泥企业的时间观测数据

Table 2 Time observations of a steel company

年份	市场 风险	财务 风险	信用 风险	运营 风险	结算 风险
2012	1.831	10.241	8.318	31.655	13.261
2013	0.806	10.885	8.278	30.879	12.712
2014	0.761	11.562	7.804	29.031	12.289
2015	0.726	12.275	7.267	27.271	11.885
2016	0.702	13.025	8.045	31.182	13.239
2017	0.688	13.815	8.450	33.240	14.048
2018	0.596	14.647	6.974	35.401	14.405

表中的数值在随机森林模型中, 称为原始数据, 为了更好地分析变量之间的相关性, 本文建立多变量随机森林预测模型进行企业信用风险指标的预测。根据 C 均值聚类算法原理, 对历史数据进行聚类。设定分类数目 $c \in [2, 30]$, 分别依据每一个 c 值都进行一次聚类操作, 根据聚类结果计算综合评价指标 J_i , 绘出其变化趋势如图 2 所示。

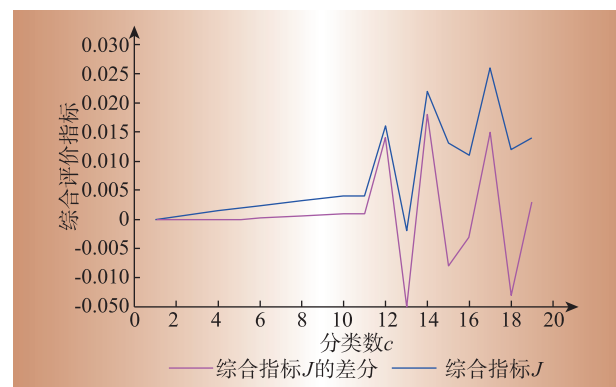


图 2 综合指标与聚类数关系图

Fig. 2 Relationship diagram of comprehensive index and cluster number

由结果可以看出, 当聚类数目 c 等于 13, 即分为 13 类时, 综合指标的一阶差分最小, 这意味着, 聚类数目 c 为 13 是最优聚类数目。各类归一化后的聚类中心如表 3 所示, 确定了聚类数目后, 便可以将历史天数做相似日归档, 同时得到各个类的聚类中心, 为相似日的判断提供依据。

本文以表 3 的数据作为历史样本, 对 2019 年 A 典型水泥企业的信用风险进行预测。在进行相似聚类后, 利用 Matlab 编程工具, 进行随机森林回归预测模型, 取决策树数目 $K=500$ 在利用训练集、测试集对随机森林模型进行搭建后, 得到 500 个决策树模型, 同时计算得出节点的权重信息。再利用建立好的模型进行回归预测, 可以得到 2019 年的预测结果如图 3 所示。

表3 各类的聚类中心(归一化后)

Table 3 Various cluster centers(after normalization)

类别	企业市场 风险	企业财务 风险	企业信用 风险	企业运营 风险	企业结 算风险
1	0.28	0.31	0.11	0.45	0.29
2	0.59	0.55	0.06	0.58	0.45
3	0.28	0.29	0.08	0.55	0.30
4	0.28	0.29	0.08	0.55	0.30
5	0.28	0.29	0.08	0.55	0.30
6	0.61	0.50	0.05	0.63	0.45
7	0.27	0.29	0.08	0.53	0.29
8	0.51	0.55	0.08	0.58	0.43
9	0.21	0.20	0.06	0.60	0.27
10	0.47	0.37	0.05	0.62	0.38
11	0.38	0.41	0.08	0.58	0.36
12	0.56	0.56	0.07	0.58	0.44
13	0.43	0.47	0.08	0.58	0.39

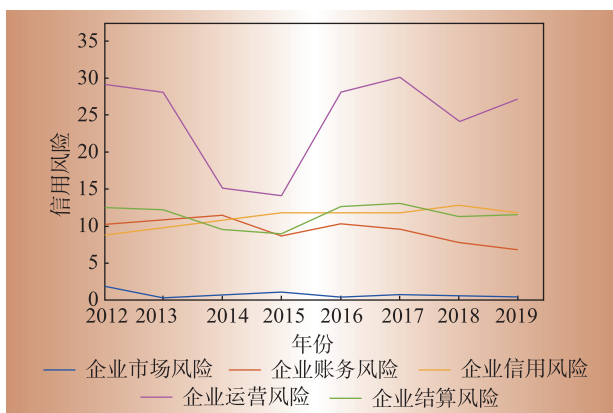


图3 A企业信用风险趋势图

Fig. 3 Corporate credit risk trend chart of A enterprise

A企业信用风险趋势如图3所示。从图3可以看出,该企业2019年的各项信用风险相比于2018年的各项指标略有下降,分析原因可知:一方面,从企业市场风险和运营风险来看,由于国内外经济下行压力较大,整体市场风险增大,不利于大型企业的原材料和加工产品进出口;另一方面,由于国家对环保节能的高度重视和去产能化的推进,转变增长方式,水泥企业面临的政策环境不容乐观,水泥企业的主要下游如建筑业、能源、造船等行业均出现需求低迷、生产萎靡的状况。而从企业财务风险、信用风险和结算风险来看,该水泥企业连续7年表现都是稳中带降,考虑到国内金融政策趋紧,企业资金周转周期延长,同时企业环保压力增大,劳动力成本增加。虽然供电企业给予了较大力度的电费优惠政策,但该水泥企业经济效益预期仍将下降,同时伴随着信用风险的上升。因此,相比于目前的研究成果,本文基于随机森林算法所提出的电力用户信用风险评价方法与预测模型,不仅能够在理论上对于电力用户的信用水平进行分析,还能在实践中对用户的信用水平进行预测,指导供电企业合理规避风险,保障电网平稳运行的同时实现利益最大化。

4 结束语

构建了用户的多层信用指标体系,通过层次分析法解构了各指标所占权重,并采用随机森林算法结合历史信用水平搭建了预警模型,综合不同维度的风险指标,通过算例对某水泥企业的信用风险进行了评价,对用户的信用水平进行了合理预测,评价结果表明本文提出的指标合理有效,信用预警模型能够预测企业的信用风险趋势,能够有效促进供电企业平稳发展,帮助电力行业在信用体系的保障下健康发展。D

参考文献:

- [1] 周冬旭,张明,朱红,等. 新电改形势下智能配电网调度互动研究应用[J]. 电力工程技术, 2018, 37(2): 89-94.
ZHOU Dongxu, ZHANG Ming, ZHU Hong, et al. Research and application of intelligent distribution network dispatching interactive under the background of electric power system reformation[J]. Jiangsu Electrical Engineering, 2018, 37(2): 89-94.
- [2] 宋健京,陈永华,杨冬梅,等. 售电侧放开后新型售电市场运营机制研究[J]. 电力需求侧管理, 2018, 20(5): 20-24.
SONG Jianjing, CHEN Yonghua, YANG Dongmei, et al. Research of new retail electricity market operation mechanism after the opening of retail electricity side[J]. Power Demand Side Management, 2018, 20(5): 20-24.
- [3] 郭鸿业,陈启鑫,夏清,等. 电力市场中的灵活调节服务:基本概念、均衡模型与研究方向[J]. 中国电机工程学报, 2017, 37(11): 3 057-3 066, 3 361.
GUO Hongye, CHEN Qixin, XIA Qing, et al. Flexible ramping product in electricity markets: basic concept, equilibrium model and research prospect[J]. Proceedings of the CSEE, 2017, 37(11): 3 057-3 066, 3 361.
- [4] DENTON M, PALMER A, MASIELLO R. Managing market risk in energy[J]. IEEE Transaction on Power Systems, 2018(2): 494-502.
- [5] 周晖,钮文洁,王毅,等. 从缴费行为分析电力客户信用度[J]. 电力需求侧管理, 2006(6): 27-29.
ZHOU Hui, NIU Wenjie, WANG Yi, et al. Analysis of clients' credit from their paying behaviors[J]. Power Demand Side Management, 2006(6): 27-29.
- [6] 王玉萍,刘磊,朱明,等. 贵州电力交易市场主体信用评级模型研究[J]. 电力需求侧管理, 2018, 20(5): 52-55, 59.
WANG Yuping, LIU Lei, ZHU Ming, et al. Research on credit rating model of players in Guizhou power trading market[J]. Power Demand Side Management, 2018, 20(5): 52-55, 59.
- [7] STREIMIKIENE D, SIKSNELYTE I. Electricity market opening impact on investments in electricity sector[J]. Renewable and Sustainable Energy Reviews, 2014, 29, 891-904.
- [8] EFRON B. Bootstrap methods: another look at the jack-knife[J]. Annals of Statistics, 1979, 7(1): 1-26.

作者简介:

余培(1993),女,贵州毕节人,研究生,研究方向为电动汽车能量管理与电力市场营销;

刘其辉(1974),男,山东济南人,博士,副教授,研究方向为新能源发电与并网技术与电力系统柔性交直流输电。

(责任编辑 陈 可)