

基于使用效率的充电设施选址评价方法研究

田贺平<sup>1</sup>, 孙舟<sup>1</sup>, 朱洁<sup>1</sup>, 陈海洋<sup>1</sup>, 赵宇彤<sup>1</sup>, 张明珠<sup>2</sup>, 陈雁<sup>2</sup>

(1. 国网北京市电力公司电力科学研究院, 北京 100075; 2. 国网信通产业集团, 北京 100085)

**摘要:** 针对电动汽车充电设施选址评价问题,通过分析充电设施使用现状及特点,提出了基于使用效率的充电设施建设选址评价模型。采用平衡数据集 SMOTE 算法生成数据样本,应用基于自助法重采样技术的随机森林算法构建模型,并应用 ROC 曲线来度量模型的分类能力与准确性。利用某市的充电设施进行了实例分析,评估候选站的使用效率作为模型应用,其结果表明,利用此模型对电动汽车公用充电设施的建设选址进行评价,准确地评估充电设施的使用率状况,提高了充电设施建设选址的针对性。

**关键词:** 电动汽车; 充电设施; 使用效率; 选址; 随机森林

**DOI:** 10.12067/ATEEE 1711060      **文章编号:** 1003-3076(2018)08-0009-08      **中图分类号:** TM910.6

1 引言

积极推动充电基础设施建设,是发展新能源汽车产业的重要保障,也是解决电动汽车充电问题的有效途径。因此充电基础设施的建设规划对电动汽车推广成效及发展规模具有显著影响,但当前充电设施建设规划以主观判断为主,存在充电设施资源投放区域、投放规模不精准及部分已投运设施影响因素多样化、使用率不平衡等问题,缺乏科学、全面的规划方法,有待利用科学分析方法发现影响因素,识别使用条件规则,进而提高充电设施使用率。

目前,国内外对于充电设施的大数据研究主要集中于电动汽车充电站规划<sup>[1-3]</sup>、充电设施数据采集<sup>[4]</sup>、电动汽车充电对配电网负荷的影响<sup>[5,6]</sup>以及充电站运营模式<sup>[7-9]</sup>与评估<sup>[10,11]</sup>等方面。文献中关于充电设施建设地址评价方法主要应用专家经验法或熵权法对指标进行量化评分,从而得到变量的取值范围。文献[12]提出一种基于 Delphi 与广义层次分析法 (General Analytic Hierarchy Process, GAHP) 相结合的电动汽车充电站选址最优策略的综合评价方法,该方法基于 Delphi 法构建电动汽车充电站地址的综合评价指标体系,采用 GAHP 法量化定性影响因素,处理不确定性灰色因素,进而可计

算得到各指标的权重。文献[13]提出一种可以考虑地理环境因素和服务半径的两步筛选法,进一步地确定充电站的建设候选地址,该方法为构建电动汽车充电站最优规划的数学模型,选取最小化规划期内充电站的总成本与网损费用之和作为优化条件。文献[12,13]均取得了一定的研究成果和结论,但为了直观评估充电设施使用效率,还需解决以下问题:①通过多方面多角度综合评价充电设施使用效率;②探究影响充电设施使用效率的因素。

本文基于某市 2016 年 3 月之前已投运充电设施的交易数据与充电设施资产数据,通过充电交易单数、充电时长以及时间占用率三方面分析现有充电设施使用现状,并根据运营管理系统中充电设施资产数据,选取充电设施属性即地理信息、运营模式、空间位置以及服务半径 3km 内其他运营商对比因素作为分析变量。对充电设施占用率进行划分评级,并采用人工合成过采样技术 (Synthetic Minority Oversampling Technique, SMOTE) 算法减弱数据的不平衡性,从而建立基于随机森林算法的充电设施建设选址评价模型。通过计算感受性曲线 (Receiver Operating Characteristics, ROC) 下方面积可知模型的分类效果较好,选取 5 个候选充电站作为模型应用的输入,其结果表明利用此模型对电动汽车公用充

**收稿日期:** 2017-11-30  
**基金项目:** 国家电网公司科技项目(52020116000J)  
**作者简介:** 田贺平(1991-),男,北京籍,工程师,硕士,研究方向为电动汽车充换电及运营管理数据分析技术;  
孙舟(1986-),男,北京籍,工程师,硕士,研究方向为电动汽车充换电及运营管理数据分析技术。

电设施的建设选址进行评价,可以准确评估充电设施使用率状况。

2 充电设施使用效率评估

为了量化充电设施使用效率,选取充电设施时间占用率作为反映充电设施使用效率的指标。应用箱线图法对充电设施时间占用率进行划分评级并建立等级指标。其中时间占用率  $R$  定义为:

$$R = \frac{\sum_i t_i}{T \times 24 \times 60} \times 100\%$$

(1)

式中,  $t_i$  为第  $i$  次充电的交易时长 (min), 交易时长由充电设施交易数据的开始时间与结束时间计算得到,反映了充电设施工作的强度;  $T$  为分析期间总交易天数,即研究课题所选取的时间窗口。

通过挖掘充电设施交易记录数据,根据充电交易单数、充电时长、时间占用率三个方面分析充电设施使用现状,并对使用效率进行评估。

计算分析期间内每个充电设施所发生的充电单数和充电时长,发现充电设施在充电交易过程中的使用现状特点如下:

(1)部分充电设施的交易单数很少(少于 100 单交易,平均每天的交易单数小于 1),利用其数据绘制直方图,如图 1 所示。

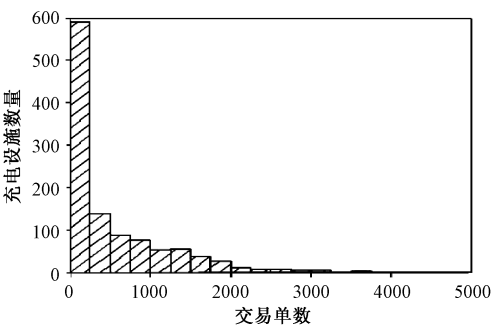


图 1 充电设施交易单数直方图  
Fig.1 Histogram of charging facility transactions

(2)部分充电设施的交易时间间隔较小(如小于 10min),即上一位用户充电结束到下一位用户开始充电的时间间隔较小。图 2 为某个充电设施的部分交易记录的时间间隔曲线图。

(3)部分充电站内每个充电桩的时间占用率相差很大,极端情况是站内最大占用率是最小占用率的数千倍,如图 3 所示。

为探究交易单数较少且充电时长较短的充电设

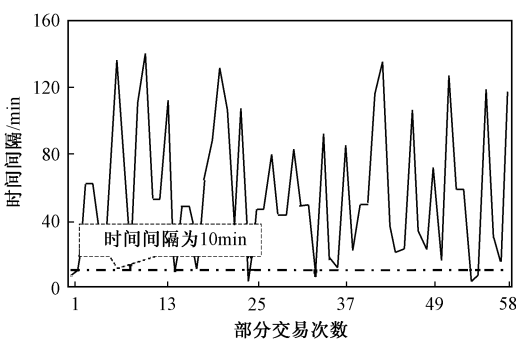


图 2 某个充电设施的部分交易记录的时间间隔  
Fig.2 Time interval of some transaction records in charging facility

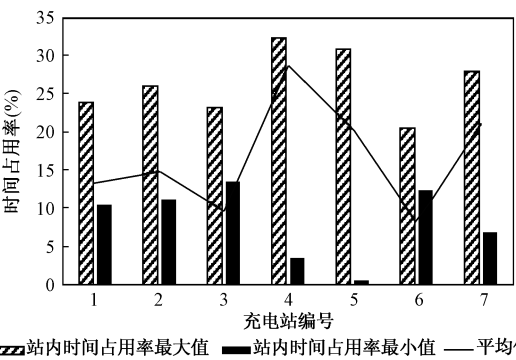


图 3 充电设施时间占用率不平衡性  
Fig.3 Imbalance charging facilities' occupancy rate

施的特点,发现这些设施的无交易天数(即该充电设施在 0:00 ~ 24:00 的时间范围内无交易记录)占分析天数的比重较大,说明该类充电设施多半处于无人问津的空闲状态,无交易天数指标体现了充电设施被使用的闲置情况。

通过计算充电设施的交易时长发现,存在同一个充电设施的连续两条交易记录的时间间隔较小(如小于 10min)的情形,即上一条交易完成至下一条交易开始的时间间隔较小,称该数据为连单数据。在实际应用中,这段较小的时间间隔包括了上一位用户拔掉充电枪和下一位用户插入充电枪进行充电的全部时间。因此根据业务逻辑,认为该情形可以体现充电桩被使用的紧张情况,使用该充电设施的用户是否需要排队等候。

针对站内充电设施时间占用率分布不均衡的情况,根据充电设施的时间占用率与充电站平均时间占用率的计算结果,发现存在充电站平均时间占用率偏高,而站内部分桩的占用率较低甚至为零的情形(见图 3 充电站 5),说明充电站存在使用效率不

平衡的状况,需现场核实是否有油车站位、充电设施故障等问题并采取相应的措施。

3 充电设施建设选址评价模型

为解决由于充电设施资源投放区域、投放规模不精准导致的使用率较低和使用率不平衡等问题,通过建立基于随机森林算法的充电设施建设选址评价模型,发现影响充电设施使用率的因素,识别使用条件规则,进而提高充电设施使用率。

现有充电设施使用率不平衡导致所采集的数据也存在不平衡性。在不平衡的数据集中,机器学习算法无法达到从少数类样本中获取足够的信息来进行准确预测的目的,因此机器学习算法对不平衡数据集进行建模时所得结果缺乏稳定性,其预测结果可能是有偏的,所得模型的预测精度也具有一定的误导性。因此机器学习算法对数据的要求是平衡数据集。

3.1 基于 SMOTE 算法的数据抽取

处理不平衡数据集时通常采用采样法,其目的为将不平衡数据修正为平衡数据。修正的原则为根据相应的模型算法调整原始数据集的样本量,使得不同类的数据比例一致<sup>[14,15]</sup>。

SMOTE 算法是行之有效且应用广泛的平衡数据集的方法,该算法的目的是生成与少数类样本相似的新数据。

平衡充电设施使用效率数据集的步骤如下:

(1)将两两样本之间的距离作为度量充电设施使用效率数据样本间相似性的依据,并根据距离矩阵确定出样本的近邻。

(2)在实际充电设施使用效率数据样本点和它近邻点的连线上随机选取一点作为生成的新样本点,从而通过生成少数类样本的随机集合来提高分类器的精度。

3.2 基于随机森林算法的分类模型构建

随机森林是一个树型决策树分类器的集合,即  $\{h(x, \beta_k), k = 1, 2, \dots\}$ , 其中元分类器  $h(x, \beta_k)$  是应用分类与回归树 (Classification and Regression Trees, CART) 算法构建的未剪枝的决策树。森林的输出采用简单多数投票法,或者是单棵树输出结果的简单平均得到,其中前者主要应用于分类模型,后者主要应用于回归模型<sup>[16-19]</sup>。

随机森林基于自助法 (boot-strap) 重采样技术,从原始训练集中有放回地随机抽取样本,以生

成  $k$  个新的训练集,然后依据这些训练集生成  $k$  颗决策树,样本数据的分类结果则由这些决策树投票决定。

在本文中,整合充电设施时间占用率与充电设施资产数据,应用随机森林算法构建充电设施选址分类模型。根据模型的决策树数量与模型误差关系图确定森林中的决策树数量,而模型分类结果由决策树结果投票决定。

3.3 基于 ROC 曲线的模型评估

为了评价充电设施选址的可靠性,需要选取评估模型性能的指标。通常分类算法采用正确率来度量分类精度,衡量模型效率。然而在不平衡数据中,采用正确率作为评价模型优良的标准时,由于少数类样本对整体精度的影响较小,不能很好地反应整体精度,所以评价结果具有一定的欺骗性,因此选择合适的评价方法对不平衡数据的分析是至关重要的<sup>[18-20]</sup>。

在不均衡数据集中,用  $P$  表示少数类,  $N$  表示多数类;  $FP$  表示多数类样本错分成少数类的数目,而  $FN$  表示将少数类样本错分成多数类的数目,同理  $TP$  和  $TN$  分别表示少数类与多数类样本被正确分类的个数,则少数类样本正确率  $TPR$  与多数类样本误判率  $FPR$  的公式为:

$$TPR = TP / (TP + FN)$$

$$FPR = TN / (FP + TN)$$

通过 ROC 曲线来衡量分类预测精度,观察模型正确地识别少数类的比例与模型错误地把多数类识别成少数类的比例之间的权衡。ROC 曲线以  $TPR$  为纵坐标,  $FPR$  为横坐标,且  $TPR$  增加的同时,  $FPR$  亦呈增长趋势,即  $TPR$  的增加以  $FPR$  的增加为代价。

ROC 曲线下方的面积 (Area Under the ROC Curve, AUC) 是模型准确率的一种度量,  $AUC$  的取值范围为  $[0.5, 1]$ , 其值越大整体分类精度就越高。一个分类性能良好的模型应该尽可能靠近图形的左上角,而一个随机猜测模型应位于连接点 ( $TPR = 0, FPR = 0$ ) 和 ( $TPR = 1, FPR = 1$ ) 的主对角线上。如果模型是完美的,那么它的  $AUC = 1$ , 如果模型是个简单的随机猜测模型,那么它的  $AUC = 0.5$ , 如果一个模型好于另一个,则它的曲线下方面积相对较大。因此,当  $AUC$  的值越接近于 1, 说明模型分类效果越好;而越接近于 0.5 时,模型分类效果越差。

本文采用 SMOTE 算法平衡数据集,对平衡后



的充电设施交易数据利用随机森林算法建立充电设施建设选址评价模型,并利用 *AUC* 指标评价模型的可靠性。

4 数据准备

数据来源于某市 2016 年 3 月之前已投运充电设施的交易数据(数据量 150 万条)、1354 个充电设施资产数据以及 1407 个其他运营商充电设施资产数据等,选取 6 个月作为分析期,且在分析期间所研究的充电设施的运行时间相同(6 个月),保证了充电设施数量、时间环境等客观条件均相同,使得分析更加客观可靠。

4.1 充电设施占用率评级

综合考虑充电设施无交易天数与时间占用率数据,根据箱线图法对时间占用率进行等级划分,根据研究问题的需求,将时间占用率依据其值大小分为 A (占用率偏低,处于较闲置状态)、B (占用率取值属中等水平,处于较正常状态)两个等级。

由于无交易天数数据体现充电设施被使用的空闲情况,而箱线图法的上四分位值或上边缘值为具有统计意义的分界值,因此考虑应用无交易天数的上四分位值(126.5)或上边缘值(182)对应的设施的占用率(分别为 6.3452%、0.008%)为划分较闲置的设施(A 等级)的依据。图 4 为无交易天数数据的箱线图。

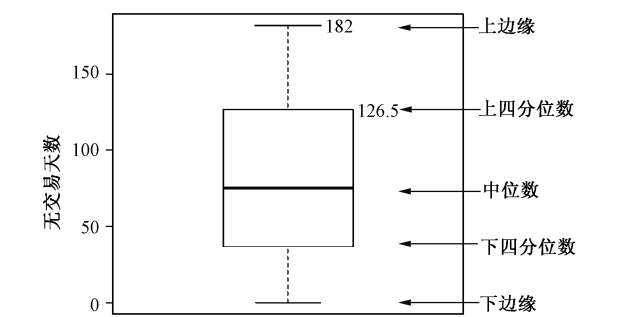


图 4 无交易天数数据的箱线图  
Fig. 4 Boxplot of no-trading data

为增强后续数据分析结果的稳健性,需要考虑到数据的平衡性,因此在本文中,将占用率为[0, 6.3452%)的充电设施划分成占用率为 A 等级,占用率为[6.3452%,60%)的充电设施设为 B 等级。

4.2 充电站运行效率评价

通过计算充电设施的占用率、不平衡性、连单数与无交易天数等数据信息,探究现有充电设施的运

行效率特点。近半年内部分充电站运行效率指标如表 1 所示。其中,不平衡性由占用率极差表示,即不平衡性等于站内最大占用率减去站内最小占用率。

表 1 近半年内部分充电站运行效率指标

Tab. 1 Operating efficiency index of some charging stations in last half year

充电站 编号	占用率 (%)	占用率 等级	不平衡性 (%)	平均 连单数	无交易 天数
1	52.1213	B	11.7418	1710	14
2	49.7741	B	0.5241	1318	10
3	49.2830	B	4.7173	1479	26
4	35.5600	B	19.2331	659	13
5	30.3006	B	3.6939	424	13
6	3.3444	A	9.6509	42	98
7	1.7429	A	3.4221	17	133
8	1.6875	A	1.5354	11	147
9	1.4733	A	0.03	14	125
10	1.4466	A	2.4947	3	167

在 B 等级充电站中,根据充电站占用率,结合占用率不平衡性与连单数据分析充电站的运行效果。充电站占用率取值越高说明充电站资源利用率越高,而不平衡性越大说明充电站资源利用越不平衡。由表 1 可知,充电站 1 与充电站 4 的占用率不平衡性较大,说明该充电站的资源利用不均衡,需运维人员管理,查看是否有油车站位、充电设施是否运转良好等;而对于占用率较大、不平衡性较小的充电站,说明用户在该站的充电需求较大,如果此时该充电站的连单情况也较严重(如充电站 2),说明充电需求呈递增的趋势,因此需要迫切增建充电桩。

对于 A 等级充电站,占用率取值偏低,无交易天数取值较大,说明大部分充电设施处于较闲置状态。A 等级充电站内同样存在由于资源利用不均衡使得不平衡性较大的情形(如充电站 6)。无交易天数取值越大,说明充电站无人问津的情况越严重,需要现场核查充电设施运行状况以及该区域的充电需求情况,从而判断是否需维护充电设施或者移建。

4.3 基于 SMOTE 算法的平衡数据集

数据集中 A、B 等级的样本数量比例为 1:2,处于较闲置状态的充电设施为少数类,在构建机器学习分类模型中,少数类样本对模型精度的贡献率较小。而实际应用中,少数类数据集通常是需要分析的重点,但对不平衡数据集建模时,模型将会更偏向于多数类数据集,而低估了少数类样本点,从而使得

模型对少数类样本的分析不充足。在本文中,经 SMOTE 算法平衡充电设施数据集后,A、B 等级样本数的比值为 1:1。

## 5 算例研究

### 5.1 模型变量选取

在本文中,选取占用率等级作为构建充电设施使用效率影响因素判别模型的因变量,根据充电设施属性选取 4 个因素作为模型的自变量:

(1)地理信息因素。来源于运营管理系统,包括 4S 店、地方委办局、地铁口、孵化器、高速路、公司营业厅、公园、供电所、机场、街道、开发区、科技园区、其他企业、桥下空间、商场广场、商务园、医院、院校、住宅小区、自然风景区等 20 个水平。

(2)其他运营商对比因素。取值范围为 $[0,1]$ ,通过地理信息系统数据,根据充电站的经纬度,计算两两充电设施间距离,确定与充电站距离小于 3km 的其他运营商的充电设施个数,根据式(2)得到其他运营商对比因素  $r$  的取值:

$$r = \frac{N}{N+n} \quad (2)$$

式中, $N$  为站内桩的个数,其值为公司运营充电站的站内桩个数; $n$  为 3km 内其他运营商桩的个数,即在该站方圆 3km 范围内的竞争对手运营商的桩个数。

(3)运营模式因素。来源于运营管理系统,根据停车场收费的不同,定义了 6 个水平,包括免费、按次收费、5 元/h 以上、5 元/h 以下、前时间段免费后时间段收费以及条件收费。

(4)空间位置因素。来源于运营管理系统,包括地面、地下、停车楼。

在充电设施建设选址评价模型中,将采用随机森林算法,分类变量为充电设施占用率等级,决策变量为地理信息、其他运营商对比、运营模式及空间位置。为检验模型的性能与泛化能力,不仅需要对模型的误差进行评估,而且需要使用一个“测试集”来测试模型对新样本的判别能力,因此采用留出法对数据集进行划分。将充电设施数据按 8:2 随机分为训练集 (train set) 和测试集 (test set),并应用 SMOTE 算法对训练集进行样本再抽取,得到平衡训练集,使得 A、B 等级样本数之比为 1:1。

### 5.2 模型参数分析

决策树数量与模型误差关系如图 5 所示。可知,模型中的决策树个数在 350 左右时模型误差会

出现较大的波动,当决策树数量大于 350 以后,模型误差逐渐趋于稳定,但仍有少许波动。因此选取森林中决策树的个数为 350,森林的输出结果由该 350 个决策树的输出结果投票所得。通过图 5 可得,总体误差在 26% 左右,但与多数类样本 (B 等级) 的误差相比,少数类样本 (A 等级) 的误差较大些。由于 B 等级的充电设施数量比 A 等级的充电设施数量多,所含信息量较全,因此 B 等级的充电设施对模型精度的贡献较大。

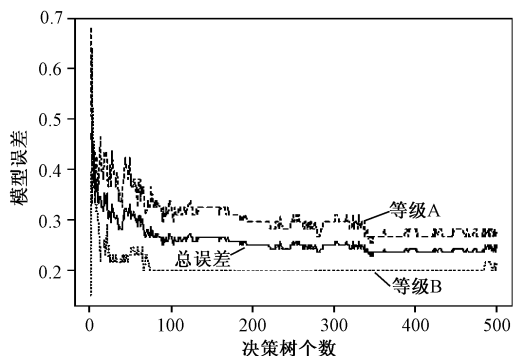


图 5 决策树数量与模型误差关系图

Fig. 5 Relationship between number of decision trees and model error

为评估充电设施建设选址评价模型的精度与可靠性,选取  $AUC$  作为模型性能的评价指标。图 6 为测试集的 ROC 曲线图,其  $AUC$  值为 0.721,由此可知该分类模型的分类效果良好。应用原始数据所建模型的  $AUC$  值为 0.646,可见平衡数据集可以有效地提高模型的分类精度。

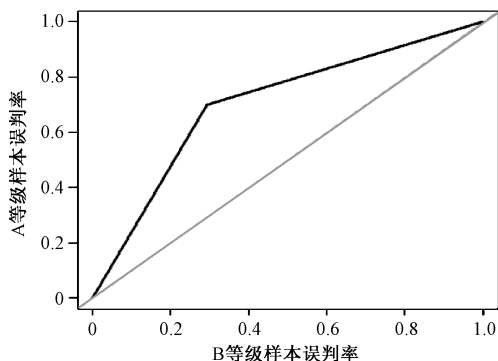


图 6 测试集 ROC 曲线

Fig. 6 ROC curve of test set

从精度平均减少值与节点不纯度的平均减少值两个维度考查变量的重要程度,结果如图 7 所示。在输出结果中,对应变量的重要值越大,则说明该变

量对于模型进行分类越重要。从图 7 可知,对充电设施使用效率影响的重要程度依次是地理信息、其他运营商对比、运营模式、空间位置,其中地理信息的重要程度最大,其他运营商对比次之,由于空间位置的重要程度很低,对模型的贡献程度可忽略不计,因此构建模型时可去除该因素。

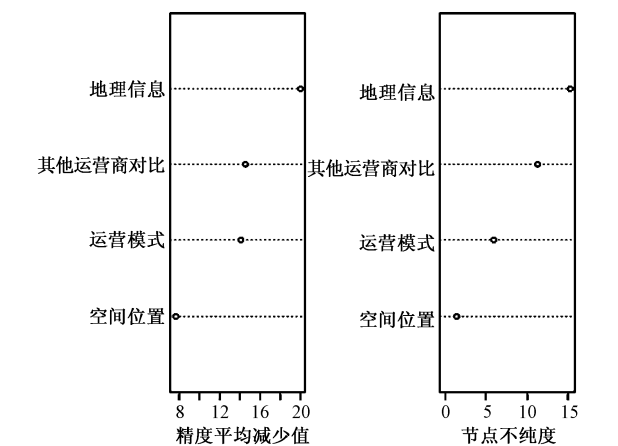


图 7 重要程度图  
Fig. 7 Importance of variables

实际中,不同的地理信息会在其影响范围内具有较固定的车流与人流信息,所以地理信息对充电设施使用效率的影响较为稳定;其他运营商对比因素具有分流用户的作用,因此其他运营商对比因素对充电设施使用效率的影响具有波动性。

5.3 模型应用

应用充电设施建设选址评价模型对 2016 年候选充电站进行占用率等级评估,其中候选站的 4 个因素已知,并根据 5.2 节建模结果,采集候选站的地理信息、其他运营商对比、运营模式 3 个维度的数据,得到候选站的“预测结果”、“建设意见”,预测结果如表 2 所示。

表 2 中,5 个候选地址均为 2016 年已投运充电站,表格最后两列给出了该充电站 2017 年上半年的平均占用率等级以及时间占用率值。由于充电设施建设选址评价模型的 AUC 值为 0.721,可知模型预测结果具有一定的可靠性。根据预测结果指导候选充电站的建设迫切程度排序,其中预测结果为 B 等级的候选站的迫切程度高于 A 等级。在相同预测结果的候选站中,考查地理信息,并结合其他运营商对比、经济效益和可实施性等条件,从而确定候选站建设的迫切程度排序,得出对新建充电站的选址问题的指导方法。

表 2 部分候选充电站选址评价结果  
Tab. 2 Evaluation results of site selection for some candidate charging stations

候选 站点 编号	地理 信息	其他 运营商 对比	运营 模式	预测 结果 (等级)	建设 意见	真实 等级	时间 占用率 (%)
1	住宅 小区	0.2	免费	B	优先 建设	B	28.20
2	商务 园	0.6667	免费	B	优先 建设	B	31.85
3	科技 园区	1	免费	A	暂缓 建设	B	17.31
4	住宅 小区	0.8889	免费	A	暂缓 建设	A	2.17
5	公园	0.9286	按次 收费	A	暂缓 建设	A	2.75

6 结论

本文将充电设施建设选址问题转化为充电设施占用率的评级与分类问题。挖掘现有充电设施的运行特点,根据其相异性,提出增建、移建、现场核查等建议;结合影响充电占用率的相关因素,针对不同占用率级别的样本不平衡问题,应用 SMOTE 算法平衡充电设施数据集,并利用 ROC 曲线进行模型评估,有效提高了随机森林算法的分类精度。模型对候选站地址进行占用率等级预测,其结果表明利用此模型可以准确评估充电设施运行效率状况,从而提高充电设施建设选址的针对性。

接下来我们将引入更广泛的影响电动汽车充电情况的数据,更加深入地分析影响充电设施使用效率的因素;对充电设施使用效率提出具体的业务应用,从而提高使用效率。

本文充电设施建设选址评价模型的结论,可结合充电负荷预测、配网容量裕度、现有充电设施布局等因素,为电动汽车充电网络的定容建设规划提供依据。

参考文献 (References):

[1] 程宏波,肖永乐,王勋,等 (Cheng Hongbo, Xiao Yongle, Wang Xun, et al.). 基于引力模型的电动汽车充电站选址规划 (Site planning of electric vehicles charging station based on gravity model) [J]. 电工电能新技术 (Advanced Technology of Electrical Engineering and Energy), 2016, 35 (5): 61-66.

- [2] 贾龙, 胡泽春, 宋永华 (Jia Long, Hu Zechun, Song Yonghua). 考虑不同类型充电需求的城市内电动汽车充电设施综合规划 (An integrated planning of electric vehicle charging facilities for urban area considering different types of charging demands) [J]. 电网技术 (Power System Technology), 2016, 40 (9): 2579-2587.
- [3] 吴健生, 李小舟 (Wu Jiansheng, Li Xiaozhou). 基于最大熵理论的多类型电动汽车充换电站规划 (Coordinated planning of multi-type electric vehicle charging stations based on maximal entropy theory) [J]. 电工电能新技术 (Advanced Technology of Electrical Engineering and Energy), 2015, 34 (9): 29-32.
- [4] 黄小庆, 陈颀, 田世明, 等 (Huang Xiaoping, Chen Jie, Tian Shiming, et al.). 电动汽车充电站规划、运行中的大数据集成应用 (Big data integration for optimal planning and operation of electric vehicle charging stations) [J]. 电网技术 (Power System Technology), 2016, 40 (3): 762-767.
- [5] 齐晓琳, 艾欣, 唐亮, 等 (Qi Xiaolin, Ai Xin, Tang Liang, et al.). 商业大楼中的电动汽车与温控负荷联合优化调度 (Joint optimal dispatch of electric vehicles and thermostatically-controlled loads in commercial buildings) [J]. 电工电能新技术 (Advanced Technology of Electrical Engineering and Energy), 2017, 36 (4): 51-58.
- [6] 和敬涵, 谢毓毓, 叶豪东, 等 (He Jinghan, Xie Yuyu, Ye Haodong, et al.). 电动汽车充电模式对主动配电网的影响 (Influence of electric vehicles charging modes on active network distribution) [J]. 电力建设 (Electric Power Construction), 2015, 36 (1): 97-102.
- [7] 吴晓康, 杨庆新, 张献, 等 (Wu Xiaokang, Yang Qingxin, Zhang Xian, et al.). 电动汽车动态充电中耦合结构研究及其效率分析 (Coil structure study and efficiency analysis for the driving wireless charging system of electric car) [J]. 电工电能新技术 (Advanced Technology of Electrical Engineering and Energy), 2016, 35 (9): 8-13.
- [8] 杨一帆, 孙瑜, 孙志杰, 等 (Yang Yifan, Sun Yu, Sun Zhijie, et al.). 考虑峰谷电价的充电站运营商和电动汽车用户经济性研究 (Study on the economical performance of charging station operators and electric vehicle users considering peak-valley electricity price) [J]. 电网技术 (Power System Technology), 2016, 40 (9): 2603-2608.
- [9] 张华栋, 杨明, 李建祥, 等 (Zhang Huadong, Yang Ming, Li Jianxiang, et al.). 一种基于单元数量控制的电动汽车充换电站电池充放电策略 (Strategy of battery charging and discharging for electric vehicle swapping station based on battery unit number) [J]. 电工电能新技术 (Advanced Technology of Electrical Engineering and Energy), 2016, 35 (8): 57-65.
- [10] 贾乐刚, 杨军 (Jia Yuegang, Yang Jun). 基于 AHP-熵权法的电动汽车充电站运行能效评估 (Running energy efficiency assessment in electric vehicle charging station based on AHP-entropy method) [J]. 电力建设 (Electric Power Construction), 2015, 36 (7): 209-215.
- [11] 黄小庆, 杨旻, 肖波, 等 (Huang Xiaoping, Yang Hang, Xiao Bo, et al.). 考虑客户满意度的电动汽车充电站运营状态评估 (Operating state evaluation considering customer satisfaction for electric vehicle charging station) [J]. 电力自动化设备 (Electric Power Automation Equipment), 2014, 34 (10): 96-101.
- [12] 冯超, 周步祥, 林楠, 等 (Feng Chao, Zhou Buxiang, Lin Nan, et al.). Delphi 和 GAHP 集成的综合评价方法在电动汽车充电站选址最优决策中的应用 (Application of comprehensive evaluation method integrating Delphi and GAHP in optimal siting of electric vehicle charging station) [J]. 电力自动化设备 (Electric Power Automation Equipment), 2012, 32 (9): 25-29.
- [13] 刘志鹏, 文福栓, 薛禹胜, 等 (Liu Zhipeng, Wen Fushuan, Xue Yusheng, et al.). 电动汽车充电站的最优选址和定容 (Optimal siting and sizing of electric vehicle charging stations) [J]. 电力系统自动化 (Automation of Electric Power Systems), 2012, 36 (3): 54-59.
- [14] 朱明, 陶新民 (Zhu Ming, Tao Xinmin). 基于随机下采样和 SMOTE 的不均衡 SVM 分类算法 (The SVM classifier for unbalanced data based on combination of RU-Undersample and SMOTE) [J]. 信息技术 (Information Technology), 2012, (1): 39-43.
- [15] 陈睿, 张亮, 杨静, 等 (Chen Rui, Zhang Liang, Yang Jing, et al.). 基于 BSMOTE 和逆转欠抽样的不均衡数据分类算法 (Classification algorithm for imbalanced data sets based on combination of BSMOTE and inverse under sampling) [J]. 计算机应用研究 (Application Research of Computers), 2014, 31 (11): 3299-3303.
- [16] 薛毅, 陈丽萍 (Xue Yi, Chen Liping). 统计建模与 R 软件 (Statistical modeling and R software) [M]. 北京: 清华大学出版社 (Beijing: Tsinghua University Press), 2007.
- [17] 黄文, 王正林 (Huang Wen, Wang Zhenglin). 数据挖掘: R 语言实战 (Data mining: R in action) [M]. 北京: 电子工业出版社 (Beijing: Electronic Industry Press), 2014.
- [18] Han Jiawei, Kamber Micheline, Pei Jian. 数据挖掘概念



与技术 (原书第 3 版) (Data mining concepts and techniques (Third Edition)) [M]. 范明, 孟小峰 (Fan Ming, Meng Xiaofeng) 译. 北京: 机械工业出版社 (Beijing: China Machine Press), 2012.

[19] 方匡南, 吴见彬, 朱建平, 等 (Fang Kuangnan, Wu Jianbin, Zhu Jianping, et al.). 随机森林方法研究综述 (A review of technologies on random forests) [J]. 统计与信息论坛 (Statistics & Information Forum), 2011, 26 (3): 32-38.

[20] 赵腾, 王林童, 张焰, 等 (Zhao Teng, Wang Lintong, Zhang Yan, et al.). 采用互信息与随机森林算法的用户用电关联因素辨识及用电量预测方法 (Relation factor identification of electricity consumption behavior of users and electricity demand forecasting based on mutual information and random forests) [J]. 中国电机工程学报 (Proceedings of the CSEE), 2016, 36 (3): 604-614.

Research on location evaluation method of charging facilities  
based on use efficiency

TIAN He-ping<sup>1</sup>, SUN Zhou<sup>1</sup>, ZHU Jie<sup>1</sup>, CHEN Hai-yang<sup>1</sup>, Zhao Yu-tong<sup>1</sup>, ZHANG Ming-zhu<sup>2</sup>, CHEN Yan<sup>2</sup>  
(1. State Grid Beijing Electric Power Research Institute, Beijing 100075, China;  
2. State Grid Information & Telecommunication Group, Beijing 100085, China)

**Abstract:** Aiming at the problem of electric vehicle charging facilities location evaluation, this paper presents a charging facilities construction site evaluation model based on the use efficiency by analyzing the current situation and characteristics of charging facilities. With the random forest algorithm that is based on boot-strap resampling technique, the data samples are generated by SMOTE (synthetic minority oversampling technique). The accuracy of classification prediction was measured by ROC (receiver operating characteristics) curve, which improves the classification accuracy of random forest algorithm. In this paper, we propose a structural analysis based on charging facilities in a city, and selected candidate stations for model application. The results show that this model can be used to evaluate the construction site of the electric charging facilities of electric vehicles, to evaluate the situation of the utilization rate of the charging facilities and to improve the targeting of the construction of the charging facilities.

**Key words:** electric vehicle; charging facilities; use efficiency; site selection; random forest