

基于电力大数据分析的接电成本预测模型

王 红¹,孙志翔²

(1. 国网江苏省电力有限公司,南京 210024;2. 国网江苏省电力有限公司 营销服务中心,南京 210019)

Electricity connection cost prediction model based on electric power big data analysis

WANG Hong¹, SUN Zhixiang²

(1. State Grid Jiangsu Electric Power Co., Ltd., Nanjing 210024, China; 2. Marketing Service Center, State Grid Jiangsu Electric Power Co., Ltd., Nanjing 210019, China)

摘要:为提升电网的运营效率和资源利用率,提高预算及成本控制的精准度,同时帮助用户制定合理的供电方案、用电策略,推进用电服务及电力营商环境不断优化,构建基于电力大数据分析的接电成本预测模型。首先,分析电力大数据中的数据类型,并采用基于MapReduce并行化处理的聚类挖掘算法,从电力系统中挖掘与接电成本相关的电力大数据,获取聚类结果。其次,通过时间序列分析法构建接电总成本预测模型,并通过多元回归方法构建接电成本影响因素预测模型,经模型预测后,获取最佳接电成本预测结果。最后,经实验验证:该模型可精准预测企业用户接入时产生的接电成本,还能够有效预测不同设备价格、不同电压等级下的接电成本变化。

关键词:电力大数据;接电成本;预测模型;并行化处理;多元回归

Abstract: In order to improve the operation efficiency and resource utilization of the power grid, improve the accuracy of budget and cost control, at the same time, help users to develop reasonable power supply plan and electricity consumption strategy, promote the continuous optimization of electricity service and electricity business environment, and build a power connection cost prediction model based on big data analysis of power. Firstly, analysis the data types in the power data, and using the MapReduce parallelization processing cluster mining algorithm, mining from the power system and the power cost related power data, obtain the clustering results. Then, through the time series analysis method build the total cost prediction model, and through the multivariate regression method to build the cost of factors prediction model, after the model prediction, get the best cost prediction results. Finally, by experiment, the model can accurately predict the power connection cost generated when enterprise users access, and can also effectively predict the change of power connection cost under different equipment prices and different voltage levels.

Key words: electric power big data; connection cost; predictive model; parallel processing; multiple regression

0 引言

2020年12月,在《国务院办公厅转发国家发展改革委等部门关于清理规范城镇供水供电供气供暖行业收费促进行业高质量发展意见的通知》(国办函〔2020〕129号)文件中,明确电力客户接入电网涉及的外部供电工程全部由电网企业或政府进行投资建设。该政策将大幅增加电网企业投资规模,推动调整电网规划建设模式,改变电网企业服务地方政府和电力客户的方式,对其经营发展具有重大影响。接电成本是指电网企业处理用户用电需求而产生的成本。这些成本包括电网企业出资用户厂区红线外的电力外线工程建设,包含投资成本、运营成本、维

护成本等^[1]。其中,投资成本主要包括电网企业建设电网设施所需的资金投入,包括变电站、输电线路、配电线路等的建设成本^[2];运营成本包括电网企业日常运营所需的人员、设备、物资等的费用支出^[3];维护成本则是指电网企业为保障电网设施安全、稳定运行而进行的维修、检测、更换等方面的费用;监管费用则是指电网企业为遵守国家电力监管机构的规定而进行的费用支出^[4]。电网企业接电成本能够直接影响市场竞争力和优质服务能力。因此,电网企业需要通过提高效率、降低成本等多种手段来控制接电成本,并通过精细化管理来实现企业的盈利和可持续发展^[5-6]。通过接电成本预测,可以帮助电网企业更加准确地了解电力市场的需求和供给情况,进而调整电力生产和输送计划,强化全面预算管理,提高电网的运营效率和资源利用率。为此,接电成本预测具有重要的意义和价值,不断优化电力营商环境。

收稿日期:2023-03-10;修回日期:2023-06-01

基金项目:国家电网有限公司科技项目(B710D0208XLI)

目前有较多学者对成本、开销预测进行研究,文献[7]研究企业成长性“高维云”预测模型,该方法能够有效预测企业成长趋势,实现企业开销预测,但该方法在预测过程中存在一定的误差,与实际结果产生较大差异;文献[8]研究中长期用电量预测模型,该方法可预测电力用户的用电量情况,为电网企业提供合理的预测数据。

电力大数据是指在电力生产、传输和配电等各个环节中产生的大量数据,这些数据涉及到电力设备运行状态、能源消耗、负荷特征等多方面内容。因此,应用电力大数据的可以提高电力系统的效率、降低成本、优化供需平衡。本文结合电力大数据,构建基于电力大数据分析的接电成本预测模型,为电网企业提供可靠的成本预测方式。

1 接电成本预测模型设计

1.1 电力大数据应用分析

电力大数据是目前一种较为新型的技术,随着电网企业的飞速应用发展,智能电网与信息技术实现了深度融合,使智能电网得到了不断的优化。目前,电网企业已逐步建设高质量发展的信息集成平台,通过数据集中式的处理,使电力数据更加清晰直观。通过以下内容,分析电力大数据的具体产生来源。

(1) 输电环节:电网输电侧电力数据,包含设备数据、输电线路故障数据等。

(2) 变电环节:包含电力调度数据与变电设备运行与维护信息等。

(3) 配电环节:包含用户负荷数据、用电数据、用户用电需求数据等。

(4) 调度环节:包含调度决策数据、运行控制数据以及安全评估数据等。

(5) 其他环节:对于用户内部工程,同样受到变压器、电缆、铜材等价格波动影响,需要同步考虑并纳入计算。

通过上述分析可知,电力大数据中存在较多种类型,这些数据有不一样的特性,为此需要有效的方式对其进行挖掘,才能够实现后续接电成本预测。

1.2 基于 MapReduce 并行化处理的聚类挖掘算法

本文利用 MapReduce 并行化软件架构,并结合电力大数据聚类挖掘算法,从海量电力大数据中挖掘可用的信息。通过 MapReduce 可以将数据集分解为较小子集,并采用并行方式对每一子集进行处理。而应用基于 MapReduce 与聚类算法进行电力大数据挖掘时,主要通过两步骤实现:第一步通过 K-means 算法初始化聚类中心,同时将数据样本划

分为相同大小的数据块;第二步启动 MapReduce 软件中的 Map 与 Reduce 任务,对 K-means 算法进行并行处理,直至获取聚类结果。传统聚类算法在进行聚类中心初始化时更具有随机性^[9],容易导致聚类结果不精准,因此本文对初始化聚类中心方法进行优化,通过优化后的 K-means 算法实现电力大数据聚类。

选取 k 个样本设为初始化聚类中心点,并将这 k 个样本存放至分布式文件系统(hadoop distributed file system, HDFS)上,将其作为全局变量。设电力大数据聚类样本数据集为 $D = \{d_i | d_i \in R, i = 1, 2, \dots, n\}$,通过表示 k 个聚类中心 c_1, c_2, \dots, c_k ,并对其作出如下定义。

(1) 随机从数据集中选取两个 n 维向量 $d_i = (d_{i1}, d_{i2}, \dots, d_{in})$ 与 $d_j = (d_{j1}, d_{j2}, \dots, d_{jn})$,两者之间可通过式(1)描述其欧氏距离

$$\text{dist}(d_i, d_j) = \sqrt{(d_{i1}, d_{j1})^2 + (d_{i2}, d_{j2})^2 + \dots + (d_{in}, d_{jn})^2} \quad (1)$$

(2) 设数据样本点中心为 $O(d_i, d_j)$,可通过式(2)对其进行计算

$$O(d_i, d_j) = \left(\frac{d_{i1}, d_{j1}}{2}, \frac{d_{i2}, d_{j2}}{2}, \dots, \frac{d_{in}, d_{jn}}{2} \right) \quad (2)$$

(3) 设样本点两者的平均距离为 a_{verage} ,并通过式(3)表示

$$a_{\text{verage}} = \frac{\sum \text{dist}(d_i, d_j)}{c_n^2} \quad i, j = 1, 2, \dots, n \quad (3)$$

该距离是指全部样本中,每两个样本之间距离之和除以样本点 n 的组合数。

输入为样本数据集 D 、聚类数量 k 、样本点距离 $\text{dist}(d_i, d_j)$ 的数据矩阵 D_1 、终止条件 ε 、集合 A 以及聚类中心集合 R 。

输出为按照终止条件,输出 k 个初始聚类中心结果。

(1) 对数据集 D 中的样本点距离 $\text{dist}(d_i, d_j)$ 进行计算,并将其保存在矩阵 D_1 中。

(2) 初始化集合 A 与集合 R ,向集合 A 中存储最小距离样本点,同时向集合 R 中存入中心 O_i ,将其作为首个聚类中心。

(3) 计算矩阵 D_1 次小距离的样本点中心,同时计算该中心与 O_i 之间的距离,并与 a_{verage} 进行对比,若低于 a_{verage} ,则将该样本存入 A 中,之后计算第 3 距离小的样本点,反复进行步骤(3)操作,若该中心大于 a_{verage} ,则将其存入集合 R 中。

(4) 直至集合 R 内的聚类样本中心数量为 k 个,完成 K-means 算法数据聚类。

(5) 启动 MapReduce 软件并行化处理聚类过程,直至获取全部聚类电力大数据。

根据聚类结果和关联规律,从原始数据中提取与接电成本相关的特征。将每个数据点与最近的聚类中心关联起来,并提取聚类中心的负荷模式、天气条件、价格水平等与接电成本相关的信息特征作为描述数据点的特征。

1.3 接电成本预测模型构建

将聚类提取的特征作为输入变量,接电成本作为目标变量,建立预测模型,预测模型可以更好地捕捉数据集中的特征和结构,提高预测的准确性和可解释性。

1.3.1 模型假设

在构建接电成本预测模型之前具体做出如下假设:

(1) 假设短期内,电网企业的经营范围、成本核算标准等相关条件不会出现大幅变动,若出现变动,也会快速恢复以往状态^[10]。

(2) 电网企业的日常电力大数据全面、真实,不存在遗漏、错误问题。

(3) 关于接电成本的相关数据均能够及时获取。

(4) 影响电网接电成本的相关因素不具有随机性,而是确定性变量,设接电成本影响因素为 X_t , 则 X_t 的存在不会是偶然性事件。

(5) 影响接电成本的相关因素范围较广,不同因素之间不一定存在关联性。

(6) 接电成本的影响因素与随机影响并不存在关联性。

(7) 变压器、电缆、铜材等物价,不出现大幅度波动。

1.3.2 接电成本预测模型设计

(1) 总体接电成本预测模型

在原始总体节点成本时间序列生成时,通过引入额外的自变量考虑不同技术措施或影响因素对接电成本的影响,捕捉到不同技术措施或影响因素对接电成本的影响程度,并帮助提高模型的预测准确性。本文假设总体接电成本为 B , 在电力大数据条件下,为保障数据的平稳性,可通过统计检验 (augmented dickey-fuller test, ADF) 可验证时间序列数据的平稳性,通过如下3种形式构建该模型

$$B_t = \theta B_{t-1} + \varepsilon_t \quad (4)$$

$$B_t = \alpha + \theta B_{t-1} + \varepsilon_t \quad (5)$$

$$B_t = \alpha + \beta_t + \theta B_{t-1} + \varepsilon_t \quad (6)$$

式中: t 为时间趋势项; ε_t 为误差项; α 为漂移项; β 、 θ 为参数。

此时设定经聚类后的电力大数据为不同年份单位接入容量成本的平稳数据,不需要再次进行稳定处理,则可通过时间序列预测分析法差分自回归移动平行模型 (autoregressive integrated moving average model, ARIMA), 对第 t 期的总体接电成本预测

模型进行构建,具体如下

$$B_t = \varphi_1 B_{t-1} + \varphi_2 B_{t-2} + \varphi_p B_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \quad (7)$$

式中: $\{B_t\}$ 为处于第 t 期、 $t-1$ 期直至 $t-q$ 期的随机误差项,依次为 ε_t 、 ε_{t-1} 、 \dots 、 ε_{t-q} , 该误差项是指具有独立性的白噪声序列,即 $VAR(\varepsilon_t) = \sigma^2$, 同时 $E(\varepsilon_s \varepsilon_t) = 0 (s \neq t)$ 。通过迭代法确定 φ_i 的值,假设 φ_i 的初始值是小于1的随意值,之后通过不断代入计算结果,直至收敛,均可通过计算机完成计算。

在电力大数据环境下,时间序列预测分析法具有较高的可行性,通过该方法,可以有效获取自身的变化规律,还能够分析干扰项对规律的影响,可对成本预测提供较大的帮助。假设当前已收集到电网企业前 N 期的成本数据,同时,随机影响因素也会对成本项造成影响,例如铜材、气候等因素,这些随机因素干扰下,会使得每期接电成本不固定,而在长期条件下,这些随机因素相对固定,偏差也会呈现稳定状态,因此这些随机因素符合随机误差项 ε_t 的特点。

在应用该成本预测模型时,可基于电力大数据分析,从电网企业随时调取相关数据,引入到预测模型中,可获取未来接电成本的发展趋势,通过这一预测,还能够及时管控企业效益。

(2) 接电成本影响因素预测模型

由于电网企业接电成本组成复杂,其中包括一些关键动因,例如人员、资产、技术与营业规模等,这些动因均可逐一分析出对接电成本造成的影响因素,具体如表1所示。

表1 接电成本影响因素分析

Table 1 Analysis of influencing factors of connection cost

具体动因	具体影响因素	电力大数据中相关数据
资产动因	气候、材料费用、客户规模、资产规模以及技术人员占比等	电网企业的设备投资、固定资产、运维成本等
技术动因	技术工艺、新标准、新设施等	电网企业的技术水平、研发投入、创新能力等
人员动因	员工素质、员工流动性、部门规模等	电网企业的员工数量、技术人员的专业背景、培训记录等
营业规模动因	用电结构、区域人口密度、区域面积等	电网企业的营业规模、服务范围、用户数量等
电力接入工程	设计费、物资费、施工费等	接电容量、接电方式、施工记录等
用户内部工程	设计费、物资费、施工费等	用电设备信息、用电行为等

采用多元回归方法构建电网企业接电成本影响因素预测模型,假设接电成本的影响因素为 X_t , 从中挑选 k 个变量,此时接电成本影响因素预测模

型可通过式(8)表示

$$B_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \mu_i \quad (i = 1, 2, \dots, k) \quad (8)$$

式中: β_i 为偏相关系数, 在其他变量受到控制时, X_i 提高或降低一个单位对总成本产生的干扰, 此时, $E(u_i) = 0, E(u_i u_j) = \sigma^2 (i \neq j)$; u_j 为干扰项。可通过最小二乘法获取回归系数 β , 通过矩阵方式对其进行表示

$$\begin{Bmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{Bmatrix} = \begin{bmatrix} 1 & X_{11} & \dots & X_{1k} \\ 1 & X_{21} & \dots & X_{2k} \\ \vdots & \vdots & & \vdots \\ 1 & X_{n1} & \dots & X_{nk} \end{bmatrix} \begin{Bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{Bmatrix} + \begin{Bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{Bmatrix} \quad (9)$$

并将总成本记为 $B = X\beta + u$, 此时对式(10)进行计算

$$\min \sum e_i^2 = \sum (B - \hat{B})^2 = \sum [B_i - (\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \dots + \hat{\beta}_k X_{ki})]^2 \quad (10)$$

对式(10)求偏导, 使其为0, 即 $\frac{\partial(\sum e_i^2)}{\partial \beta_i} = 0$ 。在进行上述计算求解时, 需应用估计系数, 为此, 通过矩阵方式表示估计系数: $\beta_i = (X^T X)^{-1} X^T Y$, 其中, X^T 是指 X 的转置。

根据上述分析与接电成本相关的动因, 本文利用柯布-道格拉斯生产函数, 构建模型 $B = QL^\alpha K^\beta S^\delta$, 其中, 人员数量为 L , 固定资产原值为 K , 营业规模为 S , 其主要是指营业收入, 技术为 Q 。由于该模型不属于线性模型, 通过对其进行转化, 使其成为线性模型, 转化形式如下

$$\ln(B) = \ln(Q) + \alpha \ln(L) + \beta \ln(K) + \delta \ln(S) \quad (11)$$

此时本文可利用多元回归方法, 分析影响接电成本的关键因素。对比式(9)和式(11)可以看出, 式(9)主要从微观视角进行分析, 而式(11)是从宏观角度进行分析。利用宏观与微观角度的大数据, 并通过加权平均将总体接电成本预测模型和多元回归预测模型结果进行整合, 得到最终的接电成本预测结果。实现接电成本影响因素的深入研究, 使预测结果更加精准。

2 实验分析

2.1 实验设置

为了验证本文模型的有效性, 模拟的具体步骤如下:

(1) 选取某电网企业中的数据作为电力大数据分析所使用的基础数据, 包括电力供应情况、电缆接入成本、接入容量。具体数据如表2所示。

表2 基础数据

Table 2 Basic data

基础数据	结果
电力供应情况/MW	1 000
电缆接入成本	投资成本
	运营成本
	维护成本
接入容量/kVA	[10, 20, 30, 40, 50]

将电力供应情况、电缆接入成本等作为自变量, 将接入成本作为因变量, 收集一定数量的样本数据进行模型训练和验证, 构成回归预测参数辨识的样本数据集。当多个用户接入电网同一节点, 共用同一配电线路或设施时, 根据每个用户的负荷容量占比来分摊接电成本。

(2) 对初始化聚类中心方法进行优化, 通过优化后的 K-means 算法实现电力大数据聚类。使用 MapReduce 软件并行化处理聚类过程, 直至获取全部聚类电力大数据。

(3) 在电力大数据条件下, 为保障数据的平稳性, 采用时间序列法构建总体接电成本预测模型。同时, 采用多元回归方法构建电网企业接电成本影响因素预测模型, 完成接电成本预测模型设计。

(4) 基于设计的预测模型, 通过以用户接入 10 kV 开关站方式和以电缆接入的方式进行模拟, 判断本文模型的有效性。

2.2 以用户接入 10 kV 开关站方式进行模拟

通过上述测算标准, 以用户接入 10 kV 开关站为例, 开展实验分析工作。10 kV 开关站包括开关站内外电气设备、平面布置及建筑物基础结构; 与开关站相关的防火、通风、防洪、防潮、防尘、防毒、防小动物和低噪声等设施。假定实验条件如表3所示。

表3 接电成本预测实验条件

Table 3 Experimental conditions for electrical connection cost prediction

条件名称	数据值
海拔高度/m	≤1 000
环境温度/°C	-30 ~ +40
月平均最高温度/°C	35
污秽等级	Ⅲ级
地震烈度	按 7° 设计, 地震加速度为 0.1 g, 地震特征周期为 0.35 s
洪涝水位	站址标高高于 50 年一遇洪水水位和历史最高内涝水位, 不考虑防洪措施; 设计土壤电阻率: 不大于 100 Ω·m
地基	地基承载力特征值取 $f_{ak} = 150$ kPa, 无地下水影响
腐蚀	地基土及地下水对钢材、混凝土无腐蚀作用

目前实际工作中常用10 kV开关站典型设计共两个方案,技术方案组合见表4。

表4 10 kV开关站典型设计技术方案组合

Table 4 Typical portfolio of technical solutions for a 10 kV sample set characteristic index switching station

方案	电气主接线	10 kV 进出线回路数	设备选型
1	单母线分段(两个独立单母线)	2进(4进),6至12回馈线	金属铠装移开式或气体绝缘金属封闭式
2	单母线三分段	4进,6至12回馈线	金属铠装移开式

方案1主要技术原则:采用单母线分段或两段独立的单母线接线,10 kV进线2回或4回,馈线12回,采用金属铠装移开式或气体绝缘金属封闭式开关柜,采用电缆进出线。在其他影响因素设定相同的情况下,针对不同配电容量的接电成本进行详细分析。分析结果如图1所示。

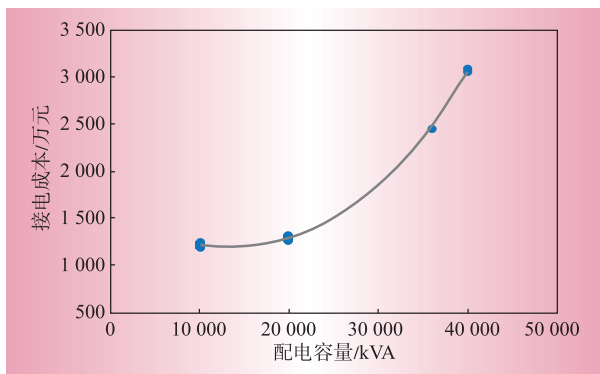


图1 方案1下不同配电容量的接电成本曲线

Fig. 1 Access cost curve of different distribution capacities in scheme 1

方案2主要技术原则:单母三分段接线,10 kV进线4回,馈线12回,采用金属铠装移开式开关柜,户内双列布置,采用电缆进出线。针对不同配电容量的接电成本进行详细分析,分析结果如图2所示。

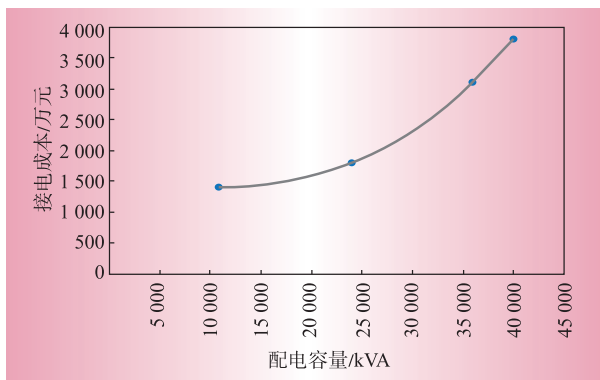


图2 方案2下不同配电容量的接电成本曲线

Fig. 2 Access cost curve of different distribution capacities in scheme 2

2.3 以电缆接入的方式进行模拟

将本文总体接电成本预测模型应用至电网企业省级电网中,对当地电力大数据进行采集,获取大量与接电成本相关的电力数据,应用这些数据进行接电成本预测分析,验证本文模型效果。在这些电力大数据中,主要包含以下数据内容,如表5所示。

表5 接电成本预测相关数据

Table 5 Related data of connection cost forecasting

数据名称	数据类型
专变接入数据	数据记录
专变报装容量	数据记录
专变典型设计	行业规范
设备故障记录	数据记录(实时更新)
设备采购信息	数据记录
设备价值信息	数据记录

应用本文模型,选取某城市中心区域以电缆接入的方式进行模拟,对接电成本进行预测,预测不同报装容量情况下,接入电网所需的成本,并对比实际成本,验证本文模型的应用性。每个不同报装容量分别进行10次接电成本预测,并根据得到的预测接电成本与实际接电成本进行预测误差计算,将10次的预测误差计算结果的平均值作为不同报装容量的预测误差值结果。具体分析结果如表6所示。

表6 专变接电成本预测(电缆)

Table 6 Forecast of special transformer connection cost (cable)

报装容量/ kVA	预测接电成本/ (万元·m ⁻¹)	实际接电成本/ (万元·m ⁻¹)	预测误差/%
500	0.140 1	0.135 5	3.42
630	0.162 9	0.159 9	1.89
800	0.204 0	0.191 1	6.78
1 000	0.241 6	0.235 5	2.59
1 600	0.456 4	0.465 3	-1.91
1 890	0.607 7	0.592 1	2.64
2 845	1.112 3	1.119 3	-0.62
3 800	1.802 8	1.783 2	1.10
4 120	1.999 1	1.996 3	0.14

根据表6可知,当报装容量逐渐增加时,单位距离的接电成本随之上升,这是由于容量增加造成设备投资的增加,通过本文模型的预测和实际比较可知,模拟出来的数据离散程度与实际用户接电成本基本相符,最大误差6.78%,均在合理范围内,因为本文模型在构建之前,利用聚类中心的特征来描述数据点的特征,提高了预测模型的准确性和可解释性。因此,本文模型具有较高水平的接电成本预测效果,可有效预测专变接电时的成本情况。

采用本文模型预测专变以电缆接入时,所产生的接电成本,分析结果如图3所示。

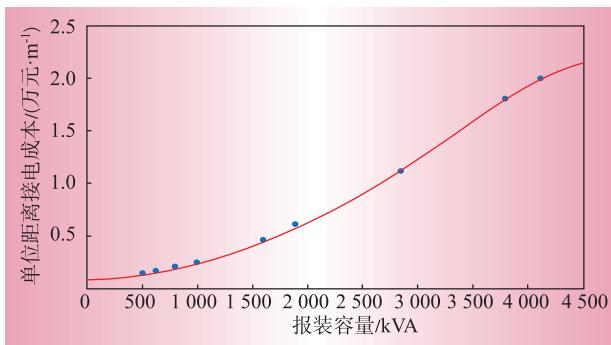


图3 专变接电成本分析曲线

Fig. 3 Analysis curve of special transformer connection cost

由以上数据可以判断出,误差在可接受的范围内,用该方法构建接入成本的费用量化模型是可行的,且该模型具有准确性。

3 结束语

本文研究基于电力大数据分析的接电成本预测模型设计,通过聚类方法挖掘出与接电成本相关的电力大数据后,通过总成本预测模型与接电影响因素预测模型,实现精准的接电成本预测。实验结果表明,所构建模型具有较好的预测效果,在未来研究过程中,可基于现有预测模型继续进行扩展,构建送电成本预测方法,使预测内容更加全面。D

参考文献:

[1] 万磊,陈成,黄文杰,等. 基于BRB和LSTM网络的电力大数据用电异常检测方法[J]. 电力建设,2021,42(8):38-45.
WAN Lei, CHEN Cheng, HUANG Wenjie, et al. Detection method of power big data based on BRB and LSTM network[J]. Electric Power Construction, 2021, 42(8): 38-45.

[2] 朱州. 基于大数据分析的电力客户服务需求预测[J]. 沈阳工业大学学报,2020,42(4):368-372.
ZHU Zhou. Power customer service demand forecast based on big data analysis[J]. Journal of Shenyang University of Technology, 2020, 42(4):368-372.

[3] 王永明,陈宇星,殷自力,等. 基于大数据分析的电力用户行为画像构建方法研究[J]. 高压电器,2022,58(10):173-179,187.
WANG Yongming, CHEN Yuxing, YIN Zili, et al. Research on the construction method of power user behavior portrait based on big data analysis[J]. High-voltage Electrical Appliances, 2022, 58(10):173-179, 187.

[4] 程晓磊,王鹏,王渊,等. 基于大数据的配电网故障诊断预测模型设计[J]. 电气传动,2022,52(2):61-66.

CHENG Xiaolei, WANG Peng, WANG Yuan, et al. Design of fault diagnosis and prediction model of power distribution network based on big data[J]. Electrical Transmission, 2022, 52(2):61-66.

[5] 刘玉娇,宋坤煌,王向. 基于电力大数据的经济景气指数分析[J]. 电信科学,2020,36(6):166-171.
LIU Yujiao, SONG Kunhuang, WANG Xiang. Analysis of the economic climate index based on the big data of electricity [J]. Telecommunication Science, 2020, 36(6):166-171.

[6] 许绘香,曹敏,马莹莹. 基于大数据分析的非线性网络流量组合预测模型[J]. 沈阳工业大学学报,2020,42(6):670-675.
XU Huixiang, CAO Min, MA Yingying. A combined prediction model for nonlinear network traffic based on big data analysis [J]. Journal of Shenyang University of Technology, 2020, 42(6):670-675.

[7] 吴永清,罗贤春,文庭孝. 基于专利大数据的企业成长性“高维云”预测模型构建及实证研究[J]. 现代情报,2020,40(3):38-46.
WU Yongqing, LUO Xianchun, WEN Tingxiao. Construction and empirical research of the “high-dimensional cloud” prediction model for enterprise growth based on patent big data[J]. Modern Intelligence, 2020, 40(3):38-46.

[8] 王新刚,朱彬若,顾臻. 基于综合能源计量大数据的中长期用电量预测[J]. 中国电力,2021,54(10):211-216.
WANG Xingang, ZHU Binruo, GU Zhen. Medium- and long-term electricity consumption forecast based on comprehensive energy metering big data [J]. China Electric Power, 2021, 54(10):211-216.

[9] 王欢,李鹏,曹敏,等. 基于CNN_BiLSTM的长短期电力负荷预测方法[J]. 计算机仿真,2022,39(3):96-103.
WANG Huan, LI Peng, CAO Min, et al. Long and short-term power load forecasting method based on CNN_BiLSTM[J]. Computer Simulation, 2022, 39(3):96-103.

[10] 莫文火,陈碧云. 基于邻域关系矩阵的电力大数据增量式属性约简研究[J]. 高压电器,2020,56(8):192-197,204.
MO Wenhua, CHEN Biyun. Incremental attribute reduction study of electric power big data based on neighborhood relation matrix [J]. High-voltage Electrical Appliances, 2020, 56(8):192-197, 204.

作者简介:

王红(1973),女,山东泰安人,硕士,高级经济师,研究方向为电力市场营销、电力需求侧管理等;

孙志翔(1992),男,通信作者,江苏连云港人,硕士,研究方向为电力系统及其自动化、电力市场营销等。

(责任编辑 于丽芳)