



基于大数据的电能替代潜力用户识别模型

夏怀健, 张文, 贾素锦, 林海英, 贺新房

(北京中电普华信息技术有限公司, 北京 100085)

摘要: 实施电能替代对推动我国节能减排的发展有着重要的作用。首先, 介绍了同行业用户聚类模型, 然后研究了电能替代潜力用户识别模型, 最后选取某行业用户进行实例验证。结果表明, 该模型可有效预测用户电能替代潜力。

关键词: 电能替代; 特征分析; 神经网络; 潜力识别

中图分类号: TM72.1

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2019030

Power substitution potential user identification model based on big data

XIA Huaijian, ZHAGN Wen, JIA Sujin, LIN Haiying, HE Xinfang

Beijing China-Power Information Technology Co., Ltd., Beijing 100085, China

Abstract: The implementation of electric energy substitution plays an important role in promoting the development of energy conservation and emission reduction in China. Firstly, the user clustering model of the same industry was introduced, then the user identification model of electric energy substitution potential was studied. Finally, an industry user was selected for instance verification. The results show that the model can effectively predict the user's energy substitution potential.

Key words: power substitution, feature analysis, neural network, potential identification

1 引言

近年以来, 污染问题引起全社会对环境保护的高度关注, 为改善环境质量、调整能源消费结构, 国家电网公司适时提出实施“以电代煤、以电代油、电从远方来”的发展战略, 推动了电能替代产业的迅速发展。在此背景下, 研究如何快速、高效地挖掘电能替代潜力将是未来电力营销市场分析决策的

重要内容。参考文献[1]对电能替代潜力进行量化, 并构建了多情景下的电能替代潜力模型, 实现对不同情景下电能替代量的有效预测; 参考文献[2]以北京市能源消费现状为基础, 建立了电能替代其他能源的评价指标体系, 对北京市主要终端能源未来需求展开预测研究; 参考文献[3]以甘肃省为研究对象, 从用煤、用油等能源结构及特点角度展开研究, 阐述了甘肃省电能替代潜力用户的大致方向。这些

理论研究推动了电能替代的快速发展,对电能替代的潜力分析展开深层次的理论研究。

由于电能替代在我国还处于起步阶段,许多理论研究偏重以区域省份为研究对象,而针对潜力用户层面的理论研究相对较少。

本文依托于电能服务管理平台电能替代用户大数据,选取同种行业内用户 96 点负荷作为电能替代用户特征分析模型的特征指标,利用模糊聚类分析模型对同行业用户进行灵活分类,基于指标个数以及用户模糊聚类类别,设计基于粒子群优化的神经网络识别模型,以聚类分组的平均年替代电量作为识别用户的电能替代潜力值,实现对电力大用户电能替代潜力分类,以此评估用户的电能替代潜力。为电力行业相关管理部门挖掘电能替代潜力用户、判断电能替代市场变化、制定相应的电能替代决策提供一定的理论支撑。

2 同行业用户聚类模型

同行业用户聚类模型研究利用用户 96 点负荷数据作为特征指标,采用模糊 C 均值算法对电能替代实施用户大数据进行模糊聚类,形成不同的聚类分组,其模型的具体步骤如下所示。

(1) 特征指标选取

用户的负荷随时间在不断发生变化,受到多种因素的影响,具有连续性和周期性的特点。不同类型的用户具有不同的负荷特点,因此,该模型以同种行业内用户 96 点负荷功率作为电能替代用户特征分析模型的特征指标^[4]。

(2) 确定隶属度矩阵

基于电能服务管理平台电能替代用户大数据,选取 n 个同行业用户,构建样本空间矩阵,设矩阵为 X ,即:

$$X = \begin{pmatrix} x_{11}, x_{12}, \dots, x_{1p} \\ x_{21}, x_{22}, \dots, x_{2p} \\ \vdots \quad \vdots \quad \ddots \quad \vdots \\ x_{n1}, x_{n2}, \dots, x_{np} \end{pmatrix} \quad (1)$$

将 n 个样本分为 c ($2 \leq c \leq n$) 类。令 $U=(u_{ik})_{c \times n}$ 为隶属度矩阵, u_{ik} 表示第 k 个样本 x_k 的隶属度。对隶属度矩阵归一化,同一数据集的隶属度总和为 1^[5-6],即:

$$\sum_{i=1}^c u_{ik} = 1, \quad k=1, 2, \dots, n \quad (2)$$

(3) 定义目标函数

设 $W=\{v_1, v_2, \dots, v_c\}$ 为聚类中心,其中, $v_i = \{v_{i1}, v_{i2}, \dots, v_{ip}\}$ 。定义目标函数 V 为^[7]:

$$V(U, W) = \sum_{i=1}^c \sum_{j=1}^N u_{ik}^m d_{ik}^2 \quad (3)$$

其中, $d_{ik} = \|x_k - v_i\|$, 表示为第 i 聚类中心与第 k 个特征向量的欧氏距离。

当 u_{ii} 满足式 (4) 时:

$$u_{ij} = \frac{1}{\sum_{k=1}^c (d_{ij} / d_{kj})^{2/(m-1)}} \quad (4)$$

可保证定义目标函数 V 取得最小值。

(4) 聚类有效性检验

分别采用同类离散度 Q_1 、中心离散度 Q_2 和平均隶属度 Q_3 对聚类有效性进行检验,计算式如下:

$$Q_1 = \sqrt{\frac{1}{c} \sum_{i=1}^c \sum_{k \in i} d_{ik}^2} \quad (5)$$

$$Q_2 = \sqrt{\frac{1}{(c-1)c} \sum_{i \neq k} d_{ik}^2} \quad (6)$$

$$Q_3 = \frac{1}{n} \sum_j U_{\max, j} \quad (7)$$

其中, Q_1 表示同一聚类之间中心的距离,该指标越小,表示分类效果越好; Q_2 为类与类之间的差别,该指标越大,表示不同聚类之间的差别越明显; Q_3 为聚类之间的清晰度,该指标越大,表示分类效果越清晰。以上述检验指标为基础,定义综合有效性指标 Q ,即:

$$Q = -\omega_1 Q_1 + \omega_2 Q_2 + \omega_3 Q_3 \quad (8)$$



其中, ω_1 、 ω_2 、 ω_3 为偏好矩阵系数。

通过上述分析可知, 综合有效性指标 Q 越大, 则分类效果就越好, 即可确定用户聚类的数目^[8]。

3 电能替代潜力用户识别模型

BP 神经网络是一种多层前向网络, 具有高度自学和自适应能力, 但存在收敛速度慢、易陷入局部最小值等缺点。粒子群优化算法具备在复杂、多峰值的数据中有效且迅速寻找全局最优解的优点, 可有效弥补 BP 神经网络的不足, 进而提升神经网络识别模型的整体性能^[9-10]。

(1) 神经网络识别模型原理

BP 神经网络由输入层、隐含层和输出层组成, 识别模型结构如图 1 所示, 其中, W_{ij} 、 W_{jk} 分别表示输入层 i 到隐含层 j 的权值、隐含层 j 到输出层 k 的权值^[11]。

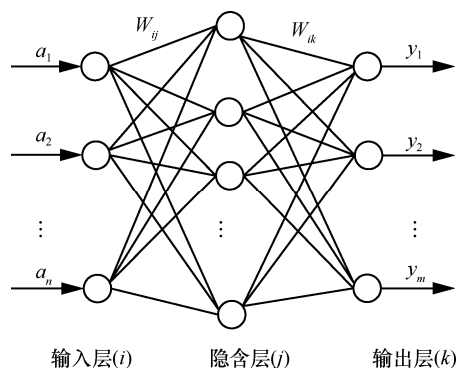


图 1 神经网络识别模型结构

模型以用户 96 点负荷功率作为输入, 根据用户聚类类别数确定识别模型个数, 并以聚类分组的平均年替代电量作为输出结果。

(2) 神经网络模型优化

假设 BP 神经网络隐含层激活阈值为 σ_j , 输出层激活阈值为 σ_k , 将神经网络模型权值和阈值组成一个群落 U , 即 $U = [W_{ij}^1, \theta_j^1, W_{jk}^2, \theta_k^2, \dots]$, 其中, $i=1, \dots, n; j=1, \dots, q; k=1, \dots, m$, 则粒子优化群在群落中的搜寻空间维数为:

$$S = (q \cdot n + q) + (m \cdot q + m) \quad (9)$$

若群落 U 由 n 个粒子组成, 假定第 i 个粒子的空间位置为 $\eta_i = (\eta_{i1}, \eta_{i2}, \dots, \eta_{is})$, 存在最佳位置为 $\xi_i = (\xi_{i1}, \xi_{i2}, \dots, \xi_{is})$, 粒子飞行速度为 $v_i = (v_{i1}, v_{i2}, \dots, v_{is})$, 整个群落搜索到最佳位置为 $\tau_i = (\tau_{i1}, \tau_{i2}, \dots, \tau_{is})$, 每个粒子的位置按照式 (10) 进行更新^[12-14]:

$$\begin{cases} v_{is}^{k+1} = \omega v_{is}^k + c_1 r_1 (\xi_{is}^k - \eta_{is}^k) + c_2 r_2 (\tau_{is}^k - \eta_{is}^k) \\ \eta_{is}^{k+1} = \eta_{is}^k + v_{is}^{k+1} \end{cases} \quad (10)$$

其中, k 为迭代次数; ω 表示惯性权重; c 表示加速因子, 一般取值为 2.05; r 表示在区间 [0,1] 均匀分布的随机数。

通过对优化后神经网络进行反复训练, 即可求得 BP 神经网络最优权值 W 和最优阈值 θ ^[15-16]。

4 实例分析

基于电能服务管理平台, 选取若干电能替代实施用户, 并进行模糊 C 均值聚类, 根据上述计算式, 即可得到聚类为 4~7 时的有效性指标。聚类有效性指标见表 1。

表 1 聚类有效性指标

指标	聚类为 4	聚类为 5	聚类为 6	聚类为 7
Q_1	0.78	0.7	0.56	0.45
Q_2	0.47	0.49	0.46	0.44
Q_3	0.83	0.79	0.71	0.71
Q	0.417	0.431	0.408	0.405

由表 1 可知, 聚类有效性最大值为 0.431, 即聚类分为 5 类时较为合适。

由于聚类分类中的用户均为电能替代已实施用户, 基于电能服务管理平台项目统计, 即可得知项目年替代电量。假设聚类分类组用 A~E 表示, 基于聚类用户的年替代电量, 计算每个分组的平均年替代电量, 作为识别用户的电能替代潜力值, 见表 2。

表 2 聚类分组的电能替代潜力值

聚类分类	A 类	B 类	C 类	D 类	E 类
潜力电量/万 kW·h	8.5	26.5	1.08	16.5	3.8

通过上述分析可知, 潜力用户识别模型的指标为用户 96 点负荷数据, 用户聚类类别数为 5, 可构建神经网络识别模型, 即识别模型以用户 96 点负荷数据作为输入, 以用户聚类类别数 5 作为模型输出个数, 输出结果以聚类用户分组的平均替代电量值进行表示。

基于电能服务管理平台用户大数据对优化后神经网络进行反复训练, 即可求得 BP 神经网络最优权值 W 和最优阈值 θ , 实现模型的有效识别。选取 60 个同行业用户对潜力识别模型进行验证, 利用 MATLAB 工具对相关数据进行验证分析, 并与传统的 BP 神经网络模型进行对比, 其结果如图 2 及表 3 所示。

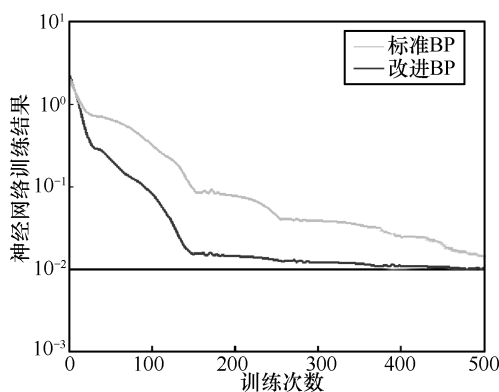


图2 两种 BP 神经网络训练结果对比分析

表3 模型识别结果分析

聚类分类	A 类	B 类	C 类	D 类	E 类
潜力电量/万 kW·h	8.5	26.5	1.08	16.5	3.8
潜力用户数/个	17	9	11	17	6

由图 2 以及表 3 可得到以下结论。

(1) 标准 BP 神经网络收敛速度较慢, 振荡幅度较大; 优化后的 BP 神经网络相对标准 BP, 收敛速度较快, 且振荡较小, 其训练效果明显优于标准 BP。

(2) A 类与 D 类潜力用户都有 17 个, 其用户潜力替代电量分别为 8.5 万 kW·h、16.5 万 kW·h; B 类用户共有 9 个, 其用户潜力替代电量为 26.5 万 kW·h;

C 类用户共有 11 个, 其用户潜力电量为 1.08 万 kW·h; E 类用户共有 6 个, 其用户潜力电量为 3.8 万 kW·h。

5 结束语

电能替代的潜力用户识别模型实现了用户电能替代潜力有效预测, 能够将用户分成不同类别的潜力对象, 有助于业务人员缩小筛选范围, 有针对性重点开展电能替代工作, 同时, 为相关部门制定相应的电能替代市场推广策略提供一定的理论依据。

参考文献:

- [1] 孙毅, 周爽, 单葆国, 等. 多情景下的电能替代潜力分析[J]. 电网技术, 2017(1): 118-124.
SUN Y, ZHOU S, SHAN B G, et al. Analysis of power substitution potential under multiple scenarios[J]. Power System Technology, 2017(1): 118-124.
- [2] 尹航. 节能减排环境下电能替代其他能源评价方法研究[D]. 北京: 华北电力大学, 2013.
YIN H. Research on evaluation method of electric energy substitution for other energy in energy saving and emission reduction environment[D]. Beijing: North China Electric Power University, 2013.
- [3] 刘俊, 徐辉, 刘亚平, 等. 甘肃省用能特点及电能替代潜力用户分析[J]. 电力需求侧管理, 2016, 18(1): 44-49.
LIU J, XU H, LIU Y P, et al. The analysis of energy consumption characteristics and electric power alternation potential in Gansu province[J]. Power Demand Side Management, 2016, 18(1): 44-49.
- [4] 郑海雁, 金农, 季聪, 等. 电力用户用电数据分析技术及典型场景应用[J]. 电网技术, 2015, 39(11): 3147-3153.
ZHENG H Y, JIN N, JI C, et al. Analysis technology and typical scenario application of electricity big data of power consumers[J]. Power System Technology, 2015, 39(11): 3147-3153.
- [5] 张洪艳. 模糊 C 均值聚类算法及应用[J]. 科技资讯, 2014(5): 178-179.
ZHANG H Y. Fuzzy C-means clustering algorithm and its application[J]. Science & Technology Information, 2014(5): 178-179.
- [6] 毛罕平, 张艳诚, 胡波. 基于模糊 C 均值聚类的作物病害叶片图像分割方法研究[J]. 农业工程学报, 2008, 24(9): 136-140.
MAO H P, ZHANG Y C, HU B. Segmentation of crop disease leaf images using fuzzy C-means clustering algorithm[J]. Transactions of the Chinese Society of Agricultural Engineering, 2008, 24(9): 136-140.



- 2008, 24(9): 136-140.
- [7] 李雷, 罗红旗, 丁亚丽. 一种优化的模糊 C 均值聚类算法[J]. 计算机技术与发展, 2009, 19(12): 71-73.
- LI L, LUO H Q, DING Y L. A novel FCM clustering algorithm[J]. Computer Technology and Development, 2009, 19(12): 71-73.
- [8] 张粒子, 蔡学文, 鲁余, 等. 面向错峰潜力分析的典型用户筛选[J]. 电力系统保护与控制, 2013(11): 146-151.
- ZHANG L Z, CAI X W, LU Y, et al. Peak shifting potential analysis-oriented typical consumers screening[J]. Power System Protection and Control, 2013(11): 146-151.
- [9] 张键锋, 王劲. 基于文本挖掘与神经网络的音乐风格分类建模方法[J]. 电信科学, 2015, 31(7): 80-85.
- ZHANG J F, WANG J. Modeling method of music style classification based on text mining and neural network [J]. Telecommunications Science, 2015, 31(7): 80-85.
- [10] 何勇, 项利国. 基于模糊聚类的 BP 神经网络模型研究及应用[J]. 系统工程理论与实践, 2004(2): 79-83.
- HE Y, XIANG L G. Study and application of BP neural network model based on fuzzy clustering[J]. Systems Engineering—Theory & Practice, 2004(2): 79-83.
- [11] 杨斐, 王坤明, 马欣, 等. 应用 BP 神经网络分类器识别交通标志[J]. 计算机工程, 2003, 29(10): 120-121.
- YANG F, WANG K M, MA X, et al. Application of BP neural network classifier for road traffic sign recognition[J]. Computer Engineering, 2003, 29(10): 120-121.
- [12] 高艳霞, 李禹生. 基于粒子群优化算法的 BP 神经网络在图像识别中的应用[J]. 武汉轻工大学学报, 2006, 25(4): 35-38.
- GAO Y X, LI Y S. Image recognition by bp neural networks model based on particle swarm optimization[J]. Journal of Wuhan Polytechnic University, 2006, 25(4): 35-38.
- [13] 余华, 黄程韦, 金赞, 等. 基于粒子群优化神经网络的语音情感识别[J]. 数据采集与处理, 2011, 26(1): 57-62.
- YU H, HUANG C W, JIN Y, et al. Speech emotion recognition based on particle swarm optimizer neural network[J]. Journal of Data Acquisition & Processing, 2011, 26(1): 57-62.
- [14] 张婷婷, 张武雄, 裴冬, 等. 基于自回归模型与神经网络模型的车流量预测对比[J]. 电信科学, 2016, 32(2): 55-59.
- ZHAGN T T, ZHANG W X, PEI D, et al. Comparison of vehicle flow prediction based on autoregressive model and neural network model [J]. Telecommunications Science, 2016, 32(2): 55-59.
- [15] 李伟, 何鹏举, 杨恒, 等. 基于粗糙集和改进遗传算法优化 BP 神经网络的算法研究[J]. 西北工业大学学报, 2012, 30(4): 601-606.
- LI W, HE P J, YANG H, et al. An effective backpropagation algorithm for optimizing BP neural network based on rough set and modified genetic algorithm[J]. Journal of Northwestern Polytechnical University, 2012, 30(4): 601-606.
- [16] 赵会敏, 雒江涛, 杨军超, 等. 集成 BP 神经网络预测模型的研究与应用[J]. 电信科学, 2016, 32(2): 60-67.
- ZHAO H M, LUO J T, YANG J C, et al. Research and application of integrated BP neural network prediction model [J]. Telecommunications Science, 2016, 32(2): 60-67.

[作者简介]



夏怀健(1988—), 男, 北京中电普华信息技术有限公司工程师, 主要研究方向为电力规划与电能替代。



张文(1973—), 男, 北京中电普华信息技术有限公司工程师, 主要研究方向为电力营销、智能配网规划。



贾素锦(1976—), 女, 北京中电普华信息技术有限公司高级工程师, 主要研究方向为电力市场与节能服务。

林海英(1976—), 女, 北京中电普华信息技术有限公司高级工程师, 主要研究方向为电力营销、负荷预测、电力咨询。

贺新房(1979—), 男, 北京中电普华信息技术有限公司工程师, 主要研究方向为电力营销。