

市场竞争机制下用户用电行为特性辨识模型

李 扬, 陈玉辰, 王子健, 蒋浩然

(东南大学 电气工程学院, 江苏 南京 210096)

摘要: 研究市场环境下的用户用电行为, 有利于优化电力市场机制, 为市场的数据化运营奠定基础, 对电网的安全稳定运行有着积极的作用。首先, 针对不同用户在电力市场环境下对电能价格和需求响应政策等的反应机制, 从购电潜力、电价敏感度、需求响应潜力3个方面构建市场行为评价指标体系, 在对初始用户样本进行指标量化的基础上, 采用二次聚类法确定用户初始类别; 然后, 基于学习向量量化(LVQ)神经网络和自组织映射(SOM)神经网络, 提出具有类别增量学习功能的自适应辨识模型; 最后, 基于某地市的实际用户数据进行模型验证。算例结果表明, 所提自适应辨识模型的辨识结果准确, 能有效地辨识得到新的用户类型, 同时在更新辨识模型的速度上也有较大的优越性。

关键词: 电力市场; 用电行为; 竞争机制; 自适应辨识; LVQ神经网络; SOM神经网络; 类别增量学习; 模型

中图分类号: TM 714; F 426.61

文献标志码: A

DOI: 10.16081/j.epae.201910015

0 引言

随着电力体制改革的深入进行, 用户在电力市场中的地位逐步提高。研究市场环境下的用户购电行为辨识模型, 一方面可充分挖掘用户在电力市场中的可调节能力, 对电网的安全稳定运行起到积极的作用; 另一方面有利于电力供应方针对不同的用户采取定制服务, 提高服务质量。文献[1-2]采用自组织映射(SOM(Self-Organizing Maps)神经网络、K均值法、模糊C均值聚类等方法对实际用户负荷曲线进行划分; 文献[3]利用基于密度的空间聚类算法提取用户的典型用电模式和行为特征, 在计算效率和聚类有效性上具有较大的优越性; 文献[4]提取用户的价值特征并采用近邻传播法对特征值进行聚类。上述负荷分析方法大多从时间维度出发对不同电力用户的负荷曲线或用电特征进行聚类分析, 此类方法对用户购电行为的分析和预测具有很强的适应性, 但不能全面地评估用户在市场竞争机制下的用电行为。文献[5]基于SOM神经网络算法得到居民用户的需求响应类型及响应等级, 但随着电力市场的进一步深入, 用户对市场的响应特征可能发生较大的变化, 新的用户类型随之出现。针对样本中出现新类型样本的问题, 文献[6]基于近邻算法, 通过比较“新样本-标准样本匹配度”与匹配度阈值来判断是否出现新类别; 文献[7]将支持向量机算法和随机森林算法相结合, 对于分类过程中新增加的类

别, 采用类别增量学习机制进行再学习, 不需要重新学习整个随机森林; 文献[8]提出一种基于SOM和概率神经网络(PNN)的增量式学习算法, 采用SOM对每类训练数据进行学习并以训练后SOM的原型向量作为此类别的模式神经元来构建PNN, 有效地提高了再学习效率。

结合以上研究, 本文以购电能力、电价敏感度和需求响应潜力为用户用电行为特征体系指标, 基于学习向量量化(LVQ)神经网络和SOM神经网络建立对电力市场深入发展下新类型用户出现的情况具有适应能力的用电行为自适应辨识方法。

1 用户市场用电行为评价指标体系

只有合理的用户市场用电行为评价指标才能科学地对市场环境下的用户进行分类, 从而充分挖掘用户的潜在价值, 帮助电力供应方判定用户类别、制定营销策略。综合考虑用户在对电价、激励政策的反应机制, 从购电潜力、电价敏感度、需求响应潜力3个方面构建用户市场用电行为评价指标体系。

1.1 购电潜力的特征指标

购电潜力采用同比用电增长率和环比用电增长率这2个指标来分析该用户的购电潜力。其中, 同比变化以年为时间尺度, 用于表征该用户的长期发展趋势; 环比变化以月为时间尺度, 用于表征该用户的近期发展趋势^[4]。计算方法如下:

$$S_{11} = (Q_i - Q_{i-1}) / Q_{i-1} \times 100\% \quad (1)$$

$$S_{12} = (Q_i - Q_{i-12}) / Q_{i-12} \times 100\% \quad (2)$$

其中, S_{11} 为同比用电增长率; S_{12} 为环比用电增长率; Q_i 、 Q_{i-1} 、 Q_{i-12} 分别为该用户在 i 月、 $i-1$ 月和 $i-12$ 月的用电量。

1.2 电价敏感度的特征指标

电价敏感度的量化是根据消费者心理学原理,

收稿日期: 2019-04-30; 修回日期: 2019-09-01

基金项目: 国家自然科学基金资助项目(动态可控负荷参与电力系统调频辅助服务理论与方法研究)(51777030)

Project supported by the National Natural Science Foundation of China(Research on the Theory and Method of Power System Frequency Regulation Service Considering Dynamic Controllable Load)(51777030)

引入负荷转移率的概念,定义负荷转移率为实施峰谷电价(TOU)后用户负荷从高电价时段向低电价时段转移量与高时段电价负荷之比。假设负荷转移率与峰平、峰谷、平谷之间的电价差是成比例的。根据大量的社会调查数据,基于负荷转移率的用户响应模型可以近似拟合成分段线性函数^[9](将其绘制在坐标系中,横坐标为各时段之间的电价差,纵坐标为用户的反应度)。各时段之间的响应数学模型为:

$$\mu_{fp} = \begin{cases} 0 & 0 \leq \Delta P_{fp} \leq a_{fp} \\ K_{fp}(\Delta P_{fp} - a_{fp}) & a_{fp} < \Delta P_{fp} \leq b_{fp} \\ \mu_{fpmax} & \Delta P_{fp} > b_{fp} \end{cases} \quad (3)$$

$$\mu_{fg} = \begin{cases} 0 & 0 \leq \Delta P_{fg} \leq a_{fg} \\ K_{fg}(\Delta P_{fg} - a_{fg}) & a_{fg} < \Delta P_{fg} \leq b_{fg} \\ \mu_{fgmax} & \Delta P_{fg} > b_{fg} \end{cases} \quad (4)$$

$$\mu_{pg} = \begin{cases} 0 & 0 \leq \Delta P_{pg} \leq a_{pg} \\ K_{pg}(\Delta P_{pg} - a_{pg}) & a_{pg} < \Delta P_{pg} \leq b_{pg} \\ \mu_{pgmax} & \Delta P_{pg} > b_{pg} \end{cases} \quad (5)$$

其中, μ_{fp} 、 μ_{fg} 、 μ_{pg} 分别为峰时段到平时段的转移率、峰时段到谷时段的转移率、平时段到谷时段的转移率; μ_{fpmax} 、 μ_{fgmax} 、 μ_{pgmax} 分别为峰时段到平时段的转移率上限、峰时段到谷时段的转移率上限、平时段到谷时段的转移率上限,即饱和区对应的纵坐标; ΔP_{fp} 、 ΔP_{fg} 、 ΔP_{pg} 分别为峰时段电价 P_f 与平时段电价 P_p 之差、峰时段电价 P_f 与谷时段电价 P_g 之差、平时段电价 P_p 与谷时段电价 P_g 之差; a_{fp} 、 a_{fg} 、 a_{pg} 为各分段线性响应曲线上的死区拐点; b_{fp} 、 b_{fg} 、 b_{pg} 为各分段线性响应曲线上的饱和区拐点; K_{fp} 、 K_{fg} 、 K_{pg} 为各分段线性响应曲线的斜率,即电价差处于第二段时的计算因子。

基于式(3)~(5)所示3类响应数学模型,各时段的拟合负荷可表示为:

$$L(k) = \begin{cases} L(k_0) + \mu_{fg} L_f N_f / N_g + \mu_{pg} L_p N_p / N_g & k \in T_g \\ L(k_0) - \mu_{pg} L_f + \mu_{fp} L_f N_f / N_p & k \in T_p \\ L(k_0) - \mu_{fp} L_f - \mu_{fg} L_f & k \in T_f \end{cases} \quad (6)$$

其中, $L(k_0)$ 、 $L(k)$ 分别为TOU实施前、后平时段第 k 小时的负荷; N_f 、 N_p 、 N_g 分别为峰时段、平时段、谷时段的小时数; T_f 、 T_p 、 T_g 分别为峰时段、平时段、谷时段集合; L_p 、 L_f 分别为TOU实施前平时段的平均负荷、峰时段的平均负荷。

参考文献[10]中的电价敏感度量方法,将某个用户电价改变前、后的多组相似日(日期类型和气候情况均相似)的负荷数据代入式(6),采用最小二乘法求解得到该用户的 μ_{fp} 、 μ_{fg} 、 μ_{pg} ,将其分别作为该用户的电价敏感度 S_{21} 、 S_{22} 、 S_{23} 。

1.3 需求响应潜力的特征指标

需求响应潜力特征指标体系包括周休负荷 S_{31} 、周休负荷下降率 S_{32} 、错时负荷 S_{33} 、峰谷差率 S_{34} 、避峰负荷 S_{35} 、最大负荷与温度相关性 S_{36} 、需求响应成

本 S_{37} ^[11],其具体计算方法分别见式(7)~(13)。

$$S_{31} = \bar{P}_{Weekday} - \bar{P}_{Weekend} \quad (7)$$

$$S_{32} = (\bar{P}_{Weekday} - \bar{P}_{Weekend}) / \bar{P}_{Weekday} \quad (8)$$

$$S_{33} = P_{peak} - \min \{ P_{advance}, P_{delay} \} \quad (9)$$

$$S_{34} = (P_{max} - P_{min}) / P_{max} \quad (10)$$

$$S_{35} = P_{peak} - P_{safe} \quad (11)$$

$$S_{36} = \text{cov}(P_{max}, \bar{t}) / (\sigma(P_{max})\sigma(\bar{t})) \quad (12)$$

$$S_{37} = C_{total} / W_{total} \quad (13)$$

其中, $\bar{P}_{Weekday}$ 和 $\bar{P}_{Weekend}$ 分别为用户工作日和周末的负荷平均值; P_{peak} 为全省高峰时刻用户的负荷值; $P_{advance}$ 、 P_{delay} 分别为全省高峰时刻提前 t_1 小时和推迟 t_2 小时的用户负荷值, t_1 、 t_2 根据电网高峰时段持续的时间来确定,高峰时段认为是负荷值大于95% P_{peak} 所持续的时段; P_{max} 、 P_{min} 分别为用户典型日负荷曲线的最高负荷和最低负荷; P_{safe} 为保证安全生产的安保负荷; $\text{cov}(\cdot)$ 、 $\sigma(\cdot)$ 分别为协方差函数、标准差函数; P_{max} 为用户日最高负荷序列; $\bar{t}$ 为用户的日平均气温序列; C_{total} 为用户的年生产总值; W_{total} 为用户的年总用电量。

1.4 特征指标量化模型

购电潜力 S_1 、电价敏感度 S_2 、需求响应潜力 S_3 由专家打分法确定,具体方法如式(14)所示。

$$S_j = \sum_i \alpha_{ji} \frac{S_{ji}}{S_{ji}^*} \quad j=1,2,3 \quad (14)$$

其中, α_{ji} 为专家打分结果; S_{ji}^* 为 S_{ji} 的基准值。

2 用户用电行为自适应辨识模型

2.1 自适应辨识流程

建立基于LVQ神经网络算法的自适应辨识模型^[12-13],具体步骤如下。

(1)获取用户的历史数据并进行用电行为特征量化:选取已有用户的历史负荷数据、电价数据和需求响应项目参与情况,计算用电行为特征。

(2)根据已有用户样本的用电行为特征确定初始类别:基于步骤(1)中的特征量化值,采用二次聚类法对已有用户样本进行聚类,确定初始类别。

(3)根据新用户的用电行为特征进行辨识:根据步骤(2)的聚类结果建立分类器,根据新用户的用电特征值,利用分类器将该用户辨识为初始类别中的某一类。

(4)检验新用户的辨识结果,确定新类候选点:根据步骤(2)中初始类别的类间距离、类外距离和步骤(3)的辨识结果,判断新用户是否属于已有类别。

(5)新类识别:对步骤(4)中的新类候选点进行新类识别。

(6)进行类别增量学习并更新辨识模型:根据步骤(5)中的新类结果进行分类器的类别增量学习,更新新辨识模型。

2.2 基于二次聚类法的初始类型提取

初始类别确定的具体方法为二次聚类法,即一次聚类采用系统聚类法对负荷特性进行分类;二次聚类采用模糊C均值聚类法,聚类中心由初次系统聚类结果提供。将系统聚类法、模糊C均值聚类法相结合,既可避免模糊C均值聚类法对初始参数的敏感性,又能取得准确客观的聚类效果。具体步骤如下:

(1) 读取样本数据,设置初始类别数为1,最大类别数为 $2 \ln N$;

(2) 采用系统聚类法对样本数据进行聚类,得到类中心,将类中心结果作为模糊C均值聚类法的初始聚类中心,对样本数据进行模糊C均值聚类,得到聚类结果;

(3) 采用 Xie-Beni 有效性指标对聚类结果进行有效性验证,保存 Xie-Beni 指标;

(4) 当类别数小于最大类别数时,将类别数增加1并转至步骤(2),否则转至步骤(5);

(5) 比较 Xie-Beni 指标,选择 Xie-Beni 指标最大者为最优类数;

(6) 若存在某个样本自成1类,则将其视为特殊数据,提取此类并将最优类数减去相应类数,取剩下的样本按步骤(2)处理,得到最终的聚类结果。

2.3 基于LVQ神经网络的行为辨识

选取 LVQ 神经网络算法作为辨识模型的分器。LVQ 神经网络是一种训练竞争层的有监督学习方法的前向神经网络,在模式识别、优化领域有广泛的运用。LVQ 神经网络的结构见图1,其由输入层、竞争层、线性输出层组成。输入层与竞争层间采用全连接的方式,竞争层与线性输出层间采用部分连接的方式。竞争层神经元与线性输出层神经元的值只能是0、1。当某个输入模式被送至网络时,与输入模式距离最近的竞争层神经元被激活,神经元的状态为1,而其他竞争层神经元的状态均为0。因此,与被激活的神经元相连的线性输出层神经元的状态也为1,而其他线性输出层神经元的状态均为0。

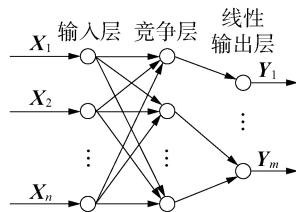


图1 LVQ神经网络的结构

Fig.1 Structure of LVQ neural network

以已有用户样本的特征值(购电潜力、电价敏感性、需求响应潜力)为输入量,类别经过0-1编码后作为输出量训练LVQ神经网络。训练完成后,将新用户的特征值输入LVQ神经网络,将该用户辨识至初始类别中的某一类。辨识模型的外部结构见图2。

为了保证神经网络的训练速度、收敛性,以类别

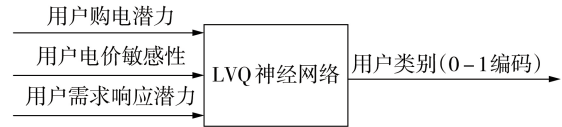


图2 辨识模型的外部结构

Fig.2 External structure of identification model

数的2倍作为LVQ神经网络的神经元个数,当神经元个数较少时,选取LVQ₁为训练算法(仅有一个“获胜”神经元),当神经元个数较多时,选取LVQ₂为训练算法(引入“次获胜”神经元);采用多次训练的方式来减小网络训练的随机性,设置训练次数为10次,单次训练的迭代次数为M,最小误差为s,当网络误差小于s时,中断训练并将该网络作为分类器,否则持续训练,选择10次训练中误差最小的网络作为分类器。具体的LVQ神经网络训练流程见图3。

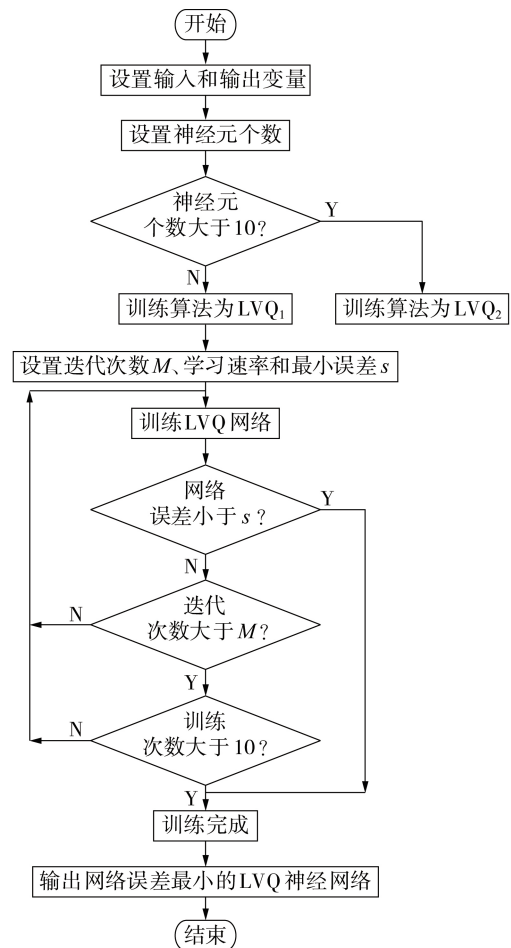


图3 LVQ神经网络训练流程

Fig.3 Training process of LVQ neural network

2.4 新类别训练

LVQ神经网络是一种前向有监督的神经网络,输入任意一组样本,其都能将该样本对应至已有类别的某一类中;然而考虑到电力用户的主观随机性,且电力市场还在发展阶段,用户对市场的反应极有可能随着市场环境的变化而发生较大的变动,新的

用户类型随之产生。当待分类数据中出现客观新类后,分类所遵循的独立同分布假设不能被满足,在识别出新类别并重新训练分类器前,那些应该被分到新类别的样本将会全部被错误分类。LVQ 神经网络缺乏自适应新类型的能力,因此需在采用 LVQ 神经网络进行用户辨识后,加入新类别的检验和训练。

本文将新类别定义为偏离已有类别且足够稠密的样本簇^[14],即需要同时满足“偏离已有类”和“稠密”这 2 个条件,此处定义类外距离 D_{out} 来衡量各类别之间的整体偏离程度,定义类内距离 D_{in} 来衡量类别内部的平均稠密程度,计算公式分别如式(15)和式(16)所示。将类内距离和类外距离的平均值定义为距离阈值 D_B ,如式(17)所示。

$$D_{out} = \left[\sum_{i \neq j} (1 - \beta_{ij}) \|X_i - X_j\|^2 \right] / \sum_{i \neq j} (1 - \beta_{ij}) \quad (15)$$

$$D_{in} = \left(\sum_{i \neq j} \beta_{ij} \|X_i - X_j\|^2 \right) / \sum_{i \neq j} \beta_{ij} \quad (16)$$

$$D_B = (D_{in} + D_{out}) / 2 \quad (17)$$

其中, X_i, X_j 分别为用户 i 、用户 j 的特征值; β_{ij} 为用户 i 和用户 j 是否处于同一类别的表征值,当用户 i 和用户 j 处于同一类别时有 $\beta_{ij} = 1$,反之则 $\beta_{ij} = 0$ 。

用户样本 a 与其辨识类别 C 之间距离 D_{a-C} 的具体计算方法如式(18)所示。

$$D_{a-C} = \left(\sum_{a \neq j} \beta_{aj} \|X_a - X_j\|^2 \right) / \sum_{a \neq j} \beta_{aj} \quad (18)$$

判断 D_{a-C} 与 D_B 的大小关系,当 $D_{a-C} \geq D_B$ 时,则用户 a 为新类别的候选点。

在确定新类别的候选点之后,需要对新类别进行判定。考虑到本项目所用分类器 LVQ 神经网络的竞争性原理和辨识模型的继承性,采用 SOM 神经网络作为新类别的识别算法。SOM 神经网络是一种无监督竞争式学习网络,能够识别环境特征并自动聚类,在用户负荷特征聚类分析中得到了广泛的应用。SOM 神经网络由输入层和输出层组成,输入层中的每一个神经元通过可变权值与输出层中的各神经元相连,输出神经元形成一个二维平面阵列。具体的新类别训练流程如图 4 所示。

2.5 类别增量学习

辨识模型从创建到不断更新是一个连续的过程,而相邻步骤中对分类器模型的更新可以视为一个类别增量学习的过程,可以发现在相邻分类器之间有大量重复的样本。若对所有样本进行重新学习和训练,将会造成大量的重复计算,对运算空间和速度造成较大的浪费。因此,可以考虑由扬弃的继承来实现分类器对样本变化的自适应更新。

获取初始分类器训练所得 LVQ 神经网络神经元 ω'_1 和竞争层到输出层的权值 ω'_2 ,获取新类别训练

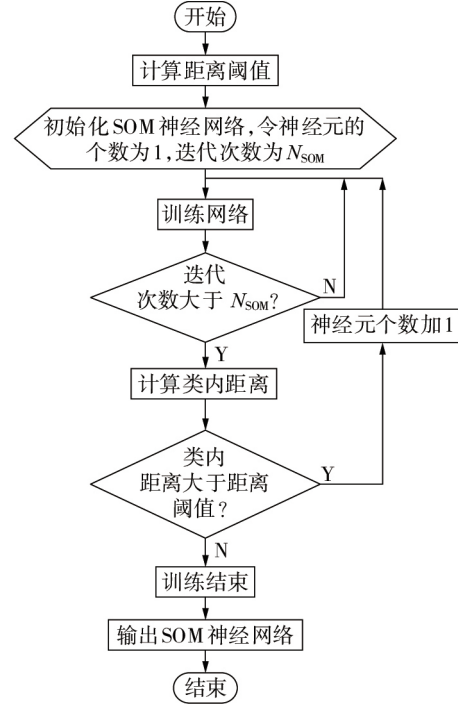


图4 新类别的训练流程

Fig.4 Training process of new category

模型所得 SOM 网络神经元 ω'_1 和竞争层到输出层的权值 ω'_2 ,将 ω'_2, ω'_1 组合为 ω_{10} , ω'_2, ω'_2 组合为 ω_{20} ,并将其分别作为待训练 LVQ 神经网络的初始神经元和竞争层到输出层的权值初值,计算方法见式(19)。

$$\omega_{10} = \begin{bmatrix} \omega'_1 \\ \omega'_1 \end{bmatrix}, \quad \omega_{20} = \begin{bmatrix} \omega'_2 & 0 \\ 0 & \omega'_2 \\ 0 & \omega'_2 \end{bmatrix} \quad (19)$$

以 ω_{10}, ω_{20} 为网络初值,设置训练次数为 10,单次训练的迭代次数为 50,学习速率为 0.2,最小误差为 0.01,当网络误差小于 0.01 时,中断训练并将该网络作为分类器,否则持续训练并选择 10 次训练中误差最小的网络作为分类器,即更新后的辨识模型。

3 算例分析

基于某地市的实际用户数据进行算例分析,提取某地市 341 家工商业用户的数据,采用 2.2 节的二次聚类法将 341 家用户分为 6 类,分类结果见图 5。

在此基础上对自适应辨识模型进行分析和验证。设置新用户中无新类别、新用户中有 1 个新类别和新用户中有 2 个新类别这 3 种场景。

3.1 场景 1:新用户中无新类别

设置新用户数分别为 10、30、50、70、90、110 户,从 314 家用户中随机抽取各类别用户作为新用户,剩余用户作为已有用户训练 LVQ 神经网络,采用训练后的网络对新用户进行辨识。辨识结果见表 1。

从辨识结果可看出,平均辨识正确率均维持在 94.22%~96.14%,采用 LVQ 神经网络进行用户辨识

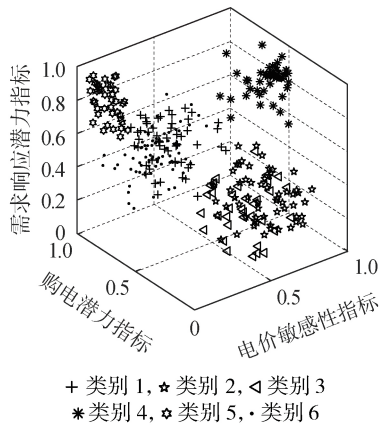


图5 二次聚类法的分类结果

Fig.5 Classification results of two-step clustering method

表1 场景1的辨识结果

Table 1 Identification results of Scene 1

新用户数/户	已有用户数/户	试验组数	平均辨识错误用户数/户	平均辨识正确率/%
10	304	20	0.5	95.00
30	284	20	1.6	94.67
50	264	20	2.7	94.60
70	244	20	2.7	96.14
90	224	20	6.0	94.22
110	204	20	6.3	94.27

具有较高的准确性。由表1可知,随着新用户数的增加,平均辨识正确率先增后降,当新用户数为70户时平均辨识正确率达到最高;当新用户数小于70户时,试验基数较小导致了较高的辨识错误率;当新用户数大于70户时,由于已有用户样本的大幅减少,辨识正确率有所降低。

3.2 场景2:新用户中有1个新类别

从314家用户中提取出某一类别的全部用户,另外抽取其他类别的部分用户作为新用户,剩余用户作为已有用户训练LVQ神经网络,采用训练后的网络对新用户进行辨识。为了避免算例的偶然性,依次抽取第1类至第6类的用户作为新用户。对新类别候选点的识别正确率进行统计,结果见表2。

由表2可以看出,当采用本文设置的距离阈值识别新类别的候选点时,除将“第1类用户和其他部分用户”作为新用户的情景A的识别正确率为91.07%外,其他情景的识别正确率均在95%及以上,识别效果良好。

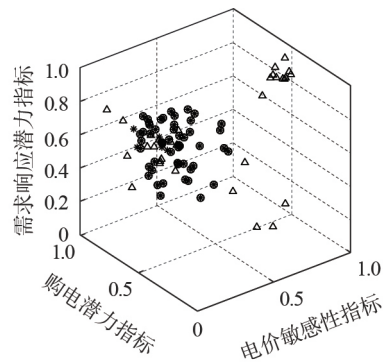
图6为情景A的辨识结果。分析图6所示结果可以发现,第1类边缘处存在5个用户与第6类用户过于接近,且随机提取新用户时提取了部分第6类靠近类中心的用户,导致类稠密度一定程度地下降,从而进一步导致识别正确率略有下降。

采用SOM神经网络从识别出的新类别候选点中训练新类别,按照新类别训练的流程,首先设置神经元个数为1,即将新类别候选点视为1个类别,计

表2 场景2的辨识效果

Table 2 Identification results of Scene 2

情景	新用户范围	新类别用户数/户	距离阈值	识别正确率/%	新类别数/个
A	第1类和其他部分用户	56	0.309 1	91.07	1
B	第2类和其他部分用户	59	0.289 0	98.33	1
C	第3类和其他部分用户	35	0.287 8	100	1
D	第4类和其他部分用户	47	0.303 2	100	1
E	第5类和其他部分用户	40	0.308 4	95.00	1
F	第6类和其他部分用户	77	0.301 0	96.10	1



*新用户中所属的新类别, △不属于非新类别的新用户
○识别出的新类别候选点

图6 情景A的辨识结果

Fig.6 Identification results of Case A

算其类内距离均远小于距离阈值,因此6种情景的新类别数目均为1,符合场景设置。

将SOM神经网络与LVQ神经网络的神经元组合作为新LVQ网络训练的神经元初始值,进行类别增量学习,记录训练时间。将不考虑神经元的继承而是直接将所有用户数据输入LVQ神经网络进行训练的训练时间与本文方法进行对比,结果见表3。

表3 场景2的训练时间

Table 3 Training time of Scene 2

情景	新用户范围	训练时间/s	
		考虑神经元的继承	不考虑神经元的继承
A	第1类和其他部分用户	4.233 4	25.929 5
B	第2类和其他部分用户	0.178 2	12.048 6
C	第3类和其他部分用户	10.912 1	26.479 4
D	第4类和其他部分用户	0.183 4	11.617 2
E	第5类和其他部分用户	0.204 4	8.908 7
F	第6类和其他部分用户	8.664 4	76.340 4
平均训练时间/s		4.062 7	26.887 3

从表3中可以看出,在场景2的6种情景下,考虑、不考虑神经元继承的平均训练时间分别为4.0627、26.8873 s,后者达到前者的6.62倍,证明采用本文所提方法进行类别增量学习可有效减少运算时间。

3.3 场景3:新用户中有2个新类别

从314家用户中任意提取出某2类的全部用户作为新用户,剩余用户作为已有用户训练LVQ神经网络。采用训练后的网络对新用户进行辨识,辨识结果如表4所示。

表4 场景3的辨识效果

Table 4 Identification results of Scene 3

新用户范围	新类别用户数	识别为新类别用户数	距离阈值	识别正确率/%	新类别数
第1,2类	115	111	0.2723	96.52	2
第1,3类	91	88	0.2709	96.70	2
第1,4类	103	102	0.2559	99.03	2
第1,5类	96	94	0.2624	97.92	2
第1,6类	133	133	0.2493	100	2
第2,3类	94	94	0.2259	100	2
第2,4类	106	106	0.2438	100	2
第2,5类	99	99	0.2502	100	2
第2,6类	136	135	0.2654	99.26	2
第3,4类	82	82	0.2424	100	2
第3,5类	75	75	0.2489	100	2
第3,6类	112	111	0.2640	99.11	2
第4,5类	87	87	0.2667	100	2
第4,6类	124	124	0.2590	100	2
第5,6类	117	116	0.2654	99.15	2

由表4可得,在场景3的15种情景下,新类别用户识别正确率均在96.50%以上,其中有12种情景的识别正确率在99%以上,且当采用SOM神经网络训练新类别时,15种情景下均提取出2个新类别,与场景设置一致。可见,本文所提新类别辨识模型在“新用户中有2个新类别”的场景下有很好的适应性。

分别统计考虑和不考虑神经元继承情况下分类器的训练时间,结果见表5。从表5中可发现,在场景3中,考虑神经元继承的训练时间依旧远远少于

表5 场景3的训练时间

Table 5 Training time of Scene 3

新用户范围	训练时间/s	
	考虑神经元继承	不考虑神经元继承
第1,2类	4.4485	16.2162
第1,3类	11.3954	32.2853
第1,4类	1.6685	19.7713
第1,5类	1.0612	13.6426
第1,6类	0.1957	21.6302
第2,3类	0.1812	51.9725
第2,4类	1.7251	8.8279
第2,5类	4.4052	26.9121
第2,6类	1.1372	4.6459
第3,4类	8.4732	20.8017
第3,5类	1.0548	10.7420
第3,6类	0.9505	7.3146
第4,5类	0.9085	4.9125
第4,6类	1.7652	11.3763
第5,6类	0.2301	9.2946
平均训练时间/s	2.6400	17.3564

不考虑神经元继承的训练时间,与表3的结果一致,验证了类别增量学习模型在运行时间上的优越性。

4 结论

本文计及电力市场深入发展新用户类别出现的情况,建立了具有类别增量学习能力的用电行为集群特性自适应辨识模型。基于某地市实际用户数据的算例分析结果表明:

(1)基于LVQ神经网络算法的自适应辨识模型在各场景中对于属于已有类别的用户辨识正确率达到100%,辨识结果准确;

(2)基于距离阈值的新类别候选点确定方法可正确隔离出属于新类别的用户;

(3)在更新辨识模型的过程中,考虑前后神经网络中神经元的继承性可有效减少运算时间,在一定程度上避免LVQ神经网络训练的随机性,具有较大的优越性。

参考文献:

- [1] 李智勇,吴晶莹,吴为麟,等. 基于自组织映射神经网络的电力用户负荷曲线聚类[J]. 电力系统自动化,2008,32(15):66-70,78.
LI Zhiyong, WU Jingying, WU Weilin, et al. Power customers load profile clustering using the SOM neural network[J]. Automation of Electric Power Systems, 2008, 32(15): 66-70, 78.
- [2] 白雪峰,蒋国栋. 基于改进K-means聚类算法的负荷建模及应用[J]. 电力自动化设备,2010,30(7):80-83.
BAI Xuefeng, JIANG Guodong. Load modeling based on improved K-means clustering algorithm and its application[J]. Electric Power Automation Equipment, 2010, 30(7): 80-83.
- [3] 苏适,李康平,严玉廷,等. 基于密度空间聚类和引力搜索算法的居民负荷用电模式分类模型[J]. 电力自动化设备,2018,38(1):129-136.
SU Shi, LI Kangping, YAN Yuting, et al. Classification model of residential power consumption mode based on DBSCAN and gravitational search algorithm[J]. Electric Power Automation Equipment, 2018, 38(1): 129-136.
- [4] 魏小曼,余昆,陈星莺,等. 基于Affinity propagation和K-means算法的电力大用户细分方法分析[J]. 电力需求侧管理,2018,20(1):15-19,35.
WEI Xiaoman, YU Kun, CHEN Xingying, et al. Analysis of power large user segmentation based on Affinity propagation and K-means algorithm[J]. Power Demand Side Management, 2018, 20(1): 15-19, 35.
- [5] 孙彦萍,李虹,杨文海,等. 基于SOM需求响应潜力的居民用户优化聚合模型[J]. 电力建设,2017,38(7):25-33.
SUN Yanping, LI Hong, YANG Wenhai, et al. Optimized aggregation model for resident users based on SOM demand response potential[J]. Electric Power Construction, 2017, 38(7): 25-33.
- [6] 叶青,卢梓豪,周洁,等. 基于近邻分类的增量学习分类算法研究[J]. 计算机工程与应用,2016,52(20):154-157.
YE Qing, LU Zihao, ZHOU Jie, et al. Research on incremental learning classification algorithm based on near neighbor[J]. Computer Engineering and Applications, 2016, 52(20): 154-157.

- [7] 薛峰,朱强,林楠. 结合改进的SVM和随机森林算法车标分类识别[J]. 计算机工程与设计,2017,38(12):3425-3430.
XUE Feng, ZHU Qiang, LIN Nan. Vehicle logo classification method based on SVM and random forest[J]. Computer Engineering and Design, 2017, 38(12): 3425-3430.
- [8] 戚湧,胡俊,於东军. 基于自组织映射与概率神经网络的增量式学习算法[J]. 南京理工大学学报(自然科学版),2013,37(1):1-6.
QI Yong, HU Jun, YU Dongjun. Incremental learning algorithm based on self-organizing map and probabilistic neural network[J]. Journal of Nanjing University of Science and Technology (Nature Science), 2013, 37(1): 1-6.
- [9] 阮文骏,王蓓蓓,李扬,等. 峰谷分时电价下的用户响应行为研究[J]. 电网技术,2012,36(7):86-93.
RUAN Wenjun, WANG Beibei, LI Yang, et al. Customer response behavior in time-of-use price[J]. Power System Technology, 2012, 36(7): 86-93.
- [10] 钱程. 基于用户用电行为建模和参数辨识的需求响应应用研究[D]. 南京:东南大学,2016.
QIAN Cheng. Research on the application of demand response based on modeling of electrical behavior and parameter identification[D]. Nanjing: Southeast University, 2016.
- [11] 徐青山,丁一帆,颜庆国,等. 大用户负荷调控潜力及价值评估研究[J]. 中国电机工程学报,2017,37(23):6791-6800,7070.
XU Qingshan, DING Yifan, YAN Qingguo, et al. Research on evaluation of scheduling potentials and values on large consumers[J]. Proceedings of the CSEE, 2017, 37(23): 6791-6800, 7070.
- [12] 王珂. 基于自适应FCM和LVQ神经网络的负荷特性分类[J]. 电气自动化,2014,36(5):55-56,93.
WANG Ke. Load characteristic classification based on self-adaptive FCM and LVQ neural network[J]. Electrical Automation, 2014, 36(5): 55-56, 93.
- [13] 刘文轩,严凤,田霖,等. 基于LVQ神经网络的配电网故障定位方法[J]. 电力系统保护与控制,2012,40(5):90-95.
LIU Wenxuan, YAN Feng, TIAN Lin, et al. LVQ neural network approach for fault location of distribution network[J]. Power System Protection and Control, 2012, 40(5): 90-95.
- [14] 谢茂强,黄亚楼,殷爱茹,等. 基于距离尺度学习的新类识别方法[J]. 模式识别与人工智能,2009,22(1):47-52.
XIE Maoqiang, HUANG Yalou, YIN Airu, et al. New class recognition method based on distance metric learning[J]. Pattern Recognition and Artificial Intelligence, 2009, 22(1): 47-52.
- [15] 李娟,王宇平. 基于样本密度和分类误差率的增量学习矢量量化算法研究[J]. 自动化学报,2015,41(6):1187-1200.
LI Juan, WANG Yuping. An incremental learning vector quantization algorithm based on pattern density and classification error ratio[J]. Acta Automatica Sinica, 2015, 41(6): 1187-1200.
- [16] 祁兵,程媛,武昕. 基于Fisher有监督判别的非侵入式居民负荷辨识方法[J]. 电网技术,2016,40(8):2484-2490.
QI Bing, CHENG Yuan, WU Xin. Non-intrusive household appliance load identification method based on Fisher supervised discriminant[J]. Power System Technology, 2016, 40(8): 2484-2490.

作者简介:



李 扬

李 扬(1961—),男,江苏泰兴人,教授,博士,主要研究方向为需求侧管理、电力市场等(E-mail:li_yang@seu.edu.cn);

陈玉辰(1994—),女,江苏东台人,硕士研究生,研究方向为需求侧管理(E-mail:chenyuchen1204@163.com);

王子健(1992—),男,安徽宿松人,硕士研究生,主要研究方向为电力市场、能源互联网等(E-mail:569606029@qq.com);

蒋浩然(1995—),男,江苏盐城人,硕士研究生,主要研究方向为负荷特性分析、数据挖掘等(E-mail:913452561@qq.com)。

Identification model of users' electricity consumption behavior characteristics under market competition mechanism

LI Yang, CHEN Yuchen, WANG Zijian, JIANG Haoran

(School of Electrical Engineering, Southeast University, Nanjing 210018, China)

Abstract: The study of users' electricity consumption behavior in the market environment is conducive to optimizing the electricity market mechanism, laying a foundation for data operation of the market, and playing a positive role in the safe and stable operation of the power grid. Firstly, according to the response mechanism of different users to electricity price and demand response policy in the electricity market environment, the market behavior evaluation index system is constructed from three aspects of electricity purchase potential, electricity price sensitivity and demand response potential. Based on the quantification of the initial user samples, the two-step clustering method is applied to determine the initial user categories. Then, an adaptive identification model with the function of category incremental learning is proposed based on LVQ (Learning Vector Quantization) neural network and SOM (Self-Organizing Maps) neural network. Finally, the proposed model is verified based on the actual users' data of a certain city. The results show that the proposed adaptive identification model can give accurate identification results, and can effectively identify new user categories. At the same time, the speed of updating the identification model has great advantages.

Key words: electricity market; electricity consumption behavior; competition mechanism; adaptive identification; LVQ neural network; SOM neural network; category incremental learning; models