

DOI: 10.13382/j.jemi.2017.12.020

采用 FCM 聚类与改进 SVR 模型的窃电行为检测

康宁宁 李川 曾虎 李英娜

(昆明理工大学 信息工程与自动化学院 昆明 650504)

摘要:针对窃电手段多样、隐蔽性强、窃电检测效率有待提高等问题,首先采用模糊 C 均值(FCM)聚类算法构造不同的用户负荷特征曲线,通过待测负荷曲线与相应特征曲线作对比初步确定疑似窃电用户;其次,采用粒子群算法优化的支持向量机回归模型对疑似窃电用户的用电行为进行检测。实验证明,所用方法缩小了窃电检测的范围、克服了窃电样本少的影响,改善了窃电检测的效率,并且窃电检测的均方误差和平均绝对误差分别提高了 0.005 1 和 0.034。

关键词:窃电检测;负荷曲线;FCM 聚类分析;粒子群算法;支持向量机回归算法

中图分类号: TM73; TP311 **文献标识码:** A **国家标准学科分类代码:** 520.20

Electric larceny detection using FCM clustering and improved SVR model

Kang Ningning Li Chuan Zeng Hu Li Yingna

(School of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650504, China)

Abstract: Aiming at the variety of electric larceny means, the efficiency of electric larceny detection remains improvement. Firstly, the fuzzy C mean clustering algorithm is used to construct different load characteristic curves of the user, and the suspicious electric larceny user is preliminarily determined by comparing the curves to be detected with the corresponding characteristic curve. Secondly, the particle swarm optimization support vector machine regression model is adopted to detect the behavior of suspected power stealing users. The experiments show that this method can reduce the range of electricity larceny detection and overcome the influence of less electricity larceny samples, improve the efficiency of electricity larceny detection, and increase the mean square error and average absolute error by 0.0051 and 0.034 respectively.

Keywords: electricity larceny detection; load curve; FCM clustering analysis; particle swarm optimization algorithm; support vector regression algorithm

0 引言

窃电不仅威胁着电网的安全运行,而且给国家经济造成重大损失,因此窃电检测的工作十分重要。传统的窃电检测主要依靠人工进行,这种方式不仅需要大量的人力资源,增加了电网公司运营成本^[1,2],而且检测效率较低,滞后于窃电行为的发生,存在取证难的问题;随着数据挖掘和机器学习等技术的发展,智能检测的方法不断涌现。

目前的检测方法主要有两种:1)根据统计方法来计

算电阻、线损、非技术损失等来判断^[3-5];2)采用数据挖掘和机器学习技术,如贝叶斯网络、构造决策树、检测离群点等^[6-8],用来发现不正常的用电模式。前一种方法虽然可确定窃电行为的发生,但不能给出确定的窃电用户;后一种方法窃电辨识率依靠大量的窃电样本训练,因此存在不足。针对这些问题,本文提出先采用模糊 C 均值(fuzzy C-means, FCM)算法对用户的用电行为进行分析,先划定疑似窃电用户,再通过粒子群优化算法-支持向量回归(PSO-SVR)模型对其进行窃电行为判定,SVR 可有效克服窃电样本少的影响。

1 采用 FCM 聚类的用户负荷曲线分类

用户负荷曲线是反映用户在一段时间内的负荷变化规律的曲线,因为用户某一天的用电负荷曲线并不能综合反映用户的日用电行为,因此文中采用了加权平均的方法来获取每个用户典型的日负荷曲线。

1.1 FCM 聚类

FCM 聚类是基于划分的聚类算法,通过隶属度来判断样本所属的类,隶属度表示一个对象 x 隶属于集合 A 的程度,通常记作 u ,其取值范围是 $[0, 1]$, $u = 1$ 表示完全属于集合 A 。

设数据集 $X = \{x_1, x_2, \dots, x_n\}$, 将其划分为 c 个簇, $V = \{v_1, v_2, \dots, v_c\}$ 表示聚类中心。

目标函数如下:

$$J(U, V) = \sum_{i=1}^n \sum_{j=1}^c u_{ij}^m d_{ij}^2 \quad (1)$$

式中: m 为模糊加权指数, u_{ij} 为第 i 个样本隶属于第 j 个簇的隶属度, d_{ij} 为第 i 个样本与第 j 个簇的中心的距离。

$$\text{隶属度矩阵 } U = \begin{bmatrix} u_{11} & \dots & u_{1c} \\ \vdots & \ddots & \vdots \\ u_{n1} & \dots & u_{nc} \end{bmatrix}, u_{ij} \text{ 表示第 } i \text{ 个样本}$$

属于第 j 个簇的隶属度,其计算公式如下:

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

其中,约束条件为 $u_{ij} \in [0, 1]$ 且 $\sum_{k=1}^c u_{ik} = 1$, d_{ik} 为第 i 个样本与第 k 个簇中心的距离。聚类中心的计算公式如下:

$$v_i = \frac{\sum_{j=1}^n u_{ij}^m x_j}{\sum_{j=1}^n u_{ij}^m} \quad (3)$$

该算法主要步骤如下:

- 1) 初始化模糊加权指数 m 、聚类数 k 以及初始的聚类中心;
- 2) 根据式(2)更新隶属度矩阵 U ;
- 3) 根据式(3)更新聚类中心矩阵 V ;
- 4) 根据式(1)计算目标函数 $J(U, V)$, 若两次目标函数的值之差小于最小误差则结束;否则跳转到步骤 2);
- 5) 根据隶属度矩阵将所有样本进行分类,若样本 x_i 属于第 j 个簇的隶属度最大,则该样本属于第 j 个簇。

1.2 聚类数的确定

在聚类分析中,决定聚类质量的关键是确定聚类

数^[9],常用的确定聚类数的聚类有效性指标有 DBI、以及有效性函数等^[10-11],本文采用加权和聚类有效性指标 (weighted sum cluster validity index, WSCVI) 来确定 FCM 算法的聚类数,其定义如下:

$$WSCVI = \sum_{i=1}^n \omega_i CVI_i \quad (4)$$

式中: n 为指标个数, ω_i 为第 i 个指标的权重系数, CVI_i 为第 i 个指标的值。

运用 WSCVI 确定聚类数的步骤如下:

- 1) 设置聚类数的范围,一般设置为 $k_{\min} = 2, k_{\max} = \sqrt{n}$, n 为样本总数;
- 2) 确定聚类有效性评价指标以及各个指标的权重,并将 WSCVI 指标聚类数确定的标准;
- 3) 在不同的聚类数下对数据集进行聚类;
- 4) 对聚类结果使用聚类有效性指标进行评价,计算其指标值;
- 5) 分析比较聚类结果,选择聚类有效性指标值最优情况下的聚类数目作为最佳聚类数。

本文抽取了云南省某地州供电局电能计量自动化系统中一共 439 个居民用户从 2016 年 3 月 ~ 2016 年 5 月共 3 个月份的用电负荷数据,并对负荷数据进行了预处理,实验中聚类数的范围为 2 ~ 20,初始聚类中心选择随机选择法进行初始化,模糊加权指数设置为 2,有效性评价指标选用 CH 指标、DB 指标、I 指标和 S 指标^[12-13],由于没有先验知识,所以每个指标的权重都一样,当 WSCVI 取最大值时,则该聚类结果最佳,得到 FCM 算法在不同聚类数下的 WSCVI 指标如图 1 所示。

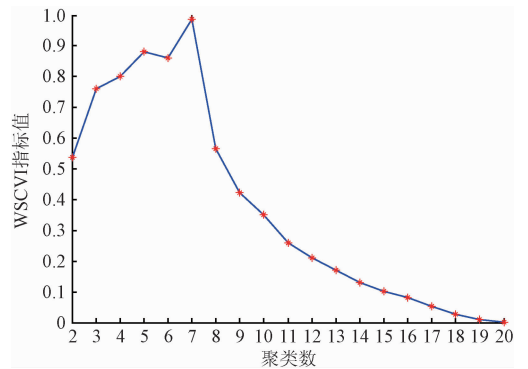


图 1 FCM 算法在不同聚类数下的 WSCVI 指标

Fig. 1 WSCVI index of FCM algorithm under different clusters

从实验的结果中可以看出,当聚类数为 7 时,FCM 算法的 WSCVI 指标值达到最优,所以最佳聚类数为 7。

1.3 用户负荷分类

本文所采用的 FCM 聚类算法对负荷曲线进行分类的流程如图 2 所示。

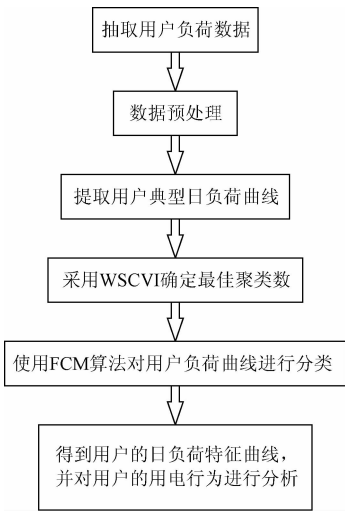


图 2 负荷分类流程

Fig. 2 Flow chart of load classification

FCM 算法进行聚类分析得到的聚类中心曲线即为该类用户的日负荷特征曲线,实验结果如图 3 所示,用户电行为分析如表 1 所示。

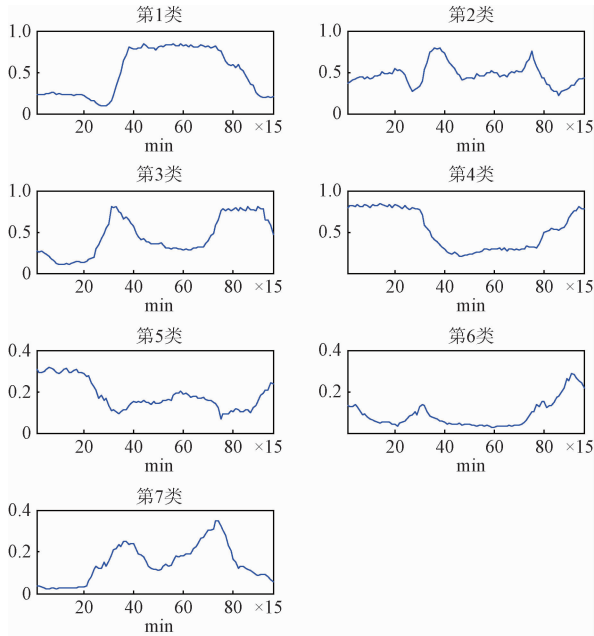


图 3 不同类别的用户日负荷特征曲线

Fig. 3 Daily load characteristic curves of different classes of users

2 疑似窃电用户筛选

通过以上的分析得到了用户的日负荷特征曲线,它反映了用户正常用电情况下的负荷变化,因此计算用户的日负荷曲线与其日负荷特征曲线的相似性就可筛选疑

表 1 用户用电行为分析

Table 1 Analysis of user electric-used behavior

类别	用电行为分析
第 1 类	用户在白天的用电量一直维持比较高的水平,从晚上 12 点后用电比较低,可能是一般的个体户。
第 2 类	用户的用电量变化呈现双峰型,可能是老人家庭和上班族,而负荷值比较高,家中大功率电器比较多。
第 3 类	用户的用电量变化也呈现双峰型,但是只是在早上和晚上出现了用电高峰,可能是老人家庭和上班族,有少量的大功率电器。
第 4 类	用户的主要的用电时间段为晚上,可能是上班族,家中大功率电器比较多。
第 5 类	用户的用电时间段为晚上,可能是上班族,家中有少量的大功率电器。
第 6 类	用户一整天的用电比较的低,晚上用电比较多,但是负荷最高峰比较小,家中几乎没有大功率电器,该用户可能是上班族。
第 7 类	用户用电呈现双峰型,但是负荷最大值很小,可能为老人家庭,家中几乎没有大功率电器。

似窃电用户;由于用户的负荷曲线是由一系列与时间顺序相关的负荷值所组成,因此可以采用基于时间序列的相似性来度量两条曲线的匹配度,常用的相似性度量有相关系数和欧氏距离。

相关系数度量侧重于考虑曲线的形状相似性,而且不受数据处理方法的影响^[14];而欧氏距离侧重于考虑值上的相似,为了兼顾曲线的值和型,本文将相关系数和欧式距离进行了综合考虑,定义两条曲线的匹配度 M 如下:

$$M = \omega_1 r + \omega_2 \ln d \tag{5}$$

式中: r 为用户考察日的负荷曲线与其负荷特征曲线的相

关系数, $r = \frac{\sum_{i=1}^{96} (x_i - \bar{x})(l_i - \bar{l})}{\sqrt{\sum_{i=1}^{96} (x_i - \bar{x})^2 \sum_{i=1}^{96} (l_i - \bar{l})^2}}$ 其中 x_i, l_i 为用

户考察日时刻 i 的负荷值, l_i 为其负荷特征曲线相应时刻的值, \bar{x}, \bar{l} 分别为以上两条曲线 1~96 个时刻点负荷值的均值; ω_1 和 ω_2 为权重系数且 $\omega_1 + \omega_2 = 1$, 本文取 $\omega_1 = \omega_2 = 0.5$; 因欧氏距离越大, 相似度越小, 为了将距离变成 0 到 1 之间的数, 所以对欧氏距离取对数, 这样距离越大, 匹配度越小, 则两条曲线间的相似性越小。根据一般情况和文中具体窃电用户样本与负荷特征曲线的对比得出匹配度的阈值为 0.8, 即匹配度小于 0.8, 将用户定为疑似窃电用户, 需对其用电行为做进一步分析。此外, 阈值可以根据具体情况进行设定。

3 PSO-SVR 模型

3.1 SVR

支持向量机(support vector machine, SVM)是 Vapnik 等人根据风险最小化原则提出的^[15-17],SVR 是建立在 SVM 原理上用于解决回归问题的,其输出结果是一个确定的值,可利用该模型得出疑似窃电用户的窃电嫌疑系数,为防止误判而没有直接将该用户分类为窃电和非窃电两类,因此文中采用支持向量机回归模型更合适。

SVR 分为线性回归和非线性回归,因窃电检测与各参量间的关系属于非线性关系,所以本文选用非线性回归的支持向量机,对于非线性回归需要将非线性映射到高维的特征空间,然后在高维空间中进行线性回归,设映射函数为 $\Phi(x)$,则回归函数为 $f(x) = w \cdot \Phi(x) + b$,转换成约束条件下的最优化问题:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (6)$$

约束条件:

$$\begin{cases} w \cdot \phi(x) + b - y_i \leq \varepsilon + \xi_i \\ y_i - w \cdot \phi(x) - b \leq \varepsilon + \xi_i^* \\ \xi_i \geq 0, \xi_i^* \geq 0 \end{cases} \quad (7)$$

则二次规划目标函数为:

$$\min \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n ((\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)(\phi(x_i) \cdot \phi(x_j))) + \varepsilon \sum_{i=1}^n (\alpha_i^* + \alpha_i) - \sum_{i=1}^n y_i (\alpha_i^* - \alpha_i) \quad (8)$$

引入核函数:

$K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$, 则该问题转换成:

$$\min \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n ((\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)K(x_i, x_j)) + \varepsilon \sum_{i=1}^n (\alpha_i^* + \alpha_i) - \sum_{i=1}^n y_i (\alpha_i^* - \alpha_i) \quad (9)$$

可以求得非线性拟合函数为:

$$f(x) = \sum_{i=1}^n K(x_i, x) (\alpha_i - \alpha_i^*) \quad (10)$$

本文选用径向基核函数:

$$K(x_i, x_j) = \exp(-\gamma(x_i - x_j)^2), \gamma > 0 \quad (11)$$

SVR 算法中,惩罚因子 C 、核函数参数 γ 和不敏感损失函数 ε 对预测的准确性有很重要的影响,因此本文选取 PSO 算法对 SVR 预测模型的参数进行优化。

实验中,SVR 模型的输入为窃电综合评价指标,包括电压不平衡率、电流不平衡率、功率因数不平衡率、台区线损率、电量峰值、电量谷值、月用电量同比和合同容量比这 8 个指标,模型的输出为疑似窃电用户的窃电检测结果,即对窃电进行评判的窃电嫌疑系数。

3.2 PSO 算法

PSO 算法来源于鸟类觅食过程的模拟和研究。它的基本思想是通过群体之间的协作和信息共享来寻找最优解。PSO 算法具有记忆特点,可以动态跟踪当前的搜索状态,调整搜索策略。

设有 n 个粒子组成的群体,在维度为 q 的空间内搜索,每个粒子表示为 $X_i = (x_{i1}, x_{i2}, \dots, x_{iq})$,每个粒子所处的位置为 $P_i = (p_{i1}, p_{i2}, \dots, p_{iq})$,每个粒子的速度为 $V_i = (v_{i1}, v_{i2}, \dots, v_{iq})$ 。在搜索过程中,需要记录每个粒子的搜索到的历史最优值 $P_i^{\text{best}} = (p_{i1}^{\text{best}}, p_{i2}^{\text{best}}, \dots, p_{iq}^{\text{best}})$ 和全局最优值 $P_g = (p_{g1}, p_{g2}, \dots, p_{gq})$ 。每个粒子的速度和位置更新如下:

$$v_{id}(k+1) = \omega v_{id}(k) + c_1 \xi (p_{id}^{\text{best}}(k) - p_{id}(k)) + c_2 \eta (p_{gd}(k) - p_{id}(k)) \quad (12)$$

$$x_{id}(k+1) = x_{id}(k) + r v_{id}(k+1) \quad (13)$$

式中: $v_{id}(k+1)$ 为第 $k+1$ 次迭代时,粒子 i 在 d 维上的速度, $x_{id}(k+1)$ 为 $k+1$ 次迭代时,粒子 i 在 d 维上的位置, ω 为保持原来速度的系数,即惯性权重, c_1 是粒子跟踪自己最优的权重系数,通常设置为 2, c_2 是跟踪全局最优的权重系数,通常设置为 2。 ξ 和 η 是 $[0,1]$ 的随机数, r 是对位置更新的时候,在速度前面加的一个系数,即为约束因子,通常设置为 1。

PSO 算法可以分为全局粒子群算法和局部粒子群算法,上面公式是描述的全局粒子群算法,局部粒子群算法的速度和位置更新如下:

$$v_{id}(k+1) = \omega v_{id}(k) + c_1 \xi (p_{id}^{\text{best}}(k) - p_{id}(k)) + c_2 \eta (p_{nd}(k) - p_{id}(k)) \quad (14)$$

$$x_{id}(k+1) = x_{id}(k) + r v_{id}(k+1) \quad (15)$$

式中: $P_n = (p_{n1}, p_{n2}, \dots, p_{nq})$ 为粒子邻域范围内找到的历史最优解。全局粒子群算法能够共享粒子群的历史最优解,因此收敛速度快,但是容易陷入局部最优解;而局部粒子群算法共享其邻域内的最优解,因此收敛速度慢,但是具有较好的全局搜索性^[18-19]。

3.3 PSO 优化 SVR 模型

在 PSO 算法中需要确定每个粒子的表示,所以在优化 SVR 模型时,每个粒子的初始位置用 $P_i = (C_i, \gamma_i, \varepsilon_i)$ 来表示,每个粒子的位置代表了一组可能的取值,在 SVR 回归预测中,可以使用均方误差来评价模型,均方误差越小,则 SVR 模型越好,因此粒子的适应度函数可以用均方误差来表示:

$$f = \frac{1}{n} \sum_{i=1}^n (y_i - y'_i)^2 \quad (16)$$

式中: y_i 为实际的值, y'_i 为回归预测得到的值。

为对 PSO-SVR 模型进行综合评价,本文采用 k-折交

叉验证的方法,取交叉验证的平均均方误差作为最后的适应度函数值。算法流程如图 4 所示。

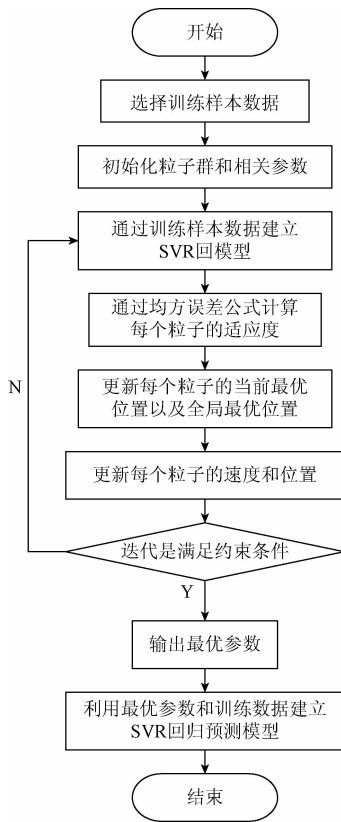


图 4 PSO-SVR 算法流程

Fig. 4 Flow chart of PSO-SVR algorithm

PSO-SVR 算法的主要步骤如下:

- 1) 选择训练样本数据;
- 2) 设置 PSO 的相关参数,如种群规模、惯性权重系数、迭代终止条件、粒子的速度和位置的范围;
- 3) 初始化粒子种群,包括每个粒子的初始位置和初始速度;
- 4) 通过训练样本数据以及粒子的位置参数建立 SVR 模型,并计算其 k-折交叉验证的平均均方误差作为每个粒子的适应度函数值;
- 5) 对每个粒子更新其当前最优位置以及粒子群的全局最优位置;
- 6) 根据式 (12) 和 (13) 更新每个粒子的速度以及位置;
- 7) 判断是否结束,若结束,则输出最优参数;否则,跳转到步骤 4);
- 8) 将最优参数带入到 SVR 模型中并通过样本数据进行训练得到 SVR 模型,此时所建立的模型即为优化后的模型。

4 实例分析

4.1 模型应用说明

对于一个待测用户的用电负荷曲线,首先与 FCM 算法得出的 7 类负荷特征曲线进行对比,通过计算匹配度来判定该用户是否为疑似窃电用户。若否,则排除其窃电嫌疑,否则该用户被列为疑似窃电用户,再采用 PSO-SVR 模型对其用电行为进行检测,并根据模型的输出值得出其窃电的嫌疑系数。

4.2 数据准备

在前面抽取得到了 439 个用户的用电数据,通过 FCM 聚类得到了 7 个类别及其日负荷特征曲线,因窃电检测时间段为一个月,所以抽取这 439 个用户 2016 年 6 月份的全部用电数据,然后通过计算每个用户典型的日负荷曲线与其日负荷特征曲线的匹配度,共筛选出 26 个疑似窃电用户,经调查发现有 3 个用户存在窃电行为。因此将这 26 个用户的用电数据作为建立窃电检测模型的样本数据,对其进行数据预处理和窃电特征数据提取,0 表示窃电嫌疑系数为 0,1 表示窃电嫌疑系数为 1,训练样本与测试样本的数据比为 4:1,训练样本数 20,测试样本数 6。

4.3 模型训练与测试

粒子种群数目设置为 50,迭代次数为 100,权重系数 c_1 和 c_2 都设置为 2,迭代过程中权重因子采用线性的策略,经 PSO 算法优化后 SVR 的参数为径向核函数的参数 γ 为 0.080 5,惩罚因子 C 为 99.860 9,不敏感损失函数 ε 为 0.000 1,迭代过程如图 5 所示。

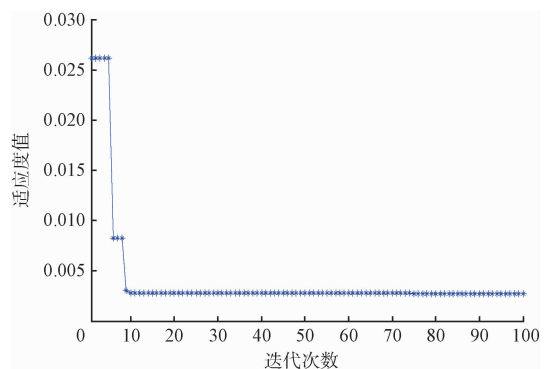


图 5 PSO 迭代过程

Fig. 5 PSO iterative process

然后使用 libsvm 工具箱,并在 MATLAB R2014a 平台上进行了实验,为了作对比分析,文中采用了 SVR 默认参数和优化后的 SVR 参数分别进行了实验,SVR 的默认

参数为 $\gamma = \frac{1}{k}$, $C = 1$, $\varepsilon = 0.1$, 实验结果如表 2 所示。

表 2 测试后的输出结果
Table 2 Output result after test

编号	预期输出	SVR 实际输出值	绝对误差	PSO-SVR 实际输出值	绝对误差
1	0	-0.025 7	0.025 7	-0.014 9	0.014 9
2	0	-0.047 4	0.047 4	0.002 2	0.002 2
3	0	-0.031 7	0.031 7	-0.003 0	0.003 0
4	0	-0.020 8	0.020 8	-0.034 6	0.034 6
5	1	0.838 5	0.161 5	0.950 0	0.050 0
6	0	-0.077 3	0.077 3	-0.044 7	0.044 7

注:表中第 2 列“预期输出”中的 0、1 代表窃电嫌疑系数,即一个值,并非代表窃电和非窃电的两种结果

从表 2 可知,当用户出现窃电后,实际的输出值很高,而没有出现窃电时的输出值都比较的低,虽然样本数较少,但从以上结果可以看出模型得出窃电和非窃电用户窃电嫌疑系数相差明显,窃电检测准确率很高,说明 SVR 算法在小样本应用泛化能力好的优势,即 SVR 算法与 PSO-SVR 算法是可行的,通过计算可知,未优化的 SVR 算法得到的均方误差为 0.006 1,平均绝对误差为 0.060 7,而优化的 SVR 算法得到的均方误差为 0.001 0,平均绝对误差为 0.024 9,所以相比较传统的 SVR 算法,经过 PSO 优化后的 SVR 算法的窃电检测准确性更高。

5 结 论

首先通过聚类分析完成了对用户负荷曲线的分类,得到不同用户的负荷特征曲线,为疑似窃电用户的判断提供了标准,其次将用户的典型负荷曲线与其负荷特征曲线进行匹配,筛选出疑似窃电用户,最后使用 PSO-SVR 模型对疑似窃电用户进行窃电检测,实验证明本文所使用方法比 SVR 模型的预测结果的均方误差和平均绝对误差分别提高了 0.005 和 0.036,证明了本文方法的有效性,同时为窃电检测提供了一种新方法,在实际应用中,需要对 PSO-SVR 模型输出值很高的用户进行重点监测和检查。

参考文献

- [1] YAP K S, TIONG S K, NAGI J, et al. Comparison of supervised learning techniques for non-technical loss detection in power utility [J]. International Review on Computers and Software, 2012, 7(2): 626-636.
- [2] 陈文瑛, 陈雁, 邱林, 等. 应用大数据技术的反窃电分析[J]. 电子测量与仪器学报, 2016, 30(10): 1558-1567.
- CHEN W, CHEN Y, QIU L, et al. Analysis of anti-stealing electric power based on big data technology[J].

Journal of Electronic Measurement & Instrumentation, 2016, 30(10): 1558-1567.

- [3] 杨佳. 基于电网电气参量特性分析的窃电辨识方法[D]. 保定: 华北电力大学, 2013.
- YANG J. Power stealing identification method based on electric parameter analysis of power grid [D]. Baoding: North China Electric Power University, 2013
- [4] NIKOVSKI D N, WANG Z H, ESENTER A, et al. Smart meter data analysis for power theft detection [C]. International Conference on Machine Learning & Data Mining in Pattern Recognition, 2013: 379-389.
- [5] SAHOO S, NIKOVSKI D, MUSO T, et al. Electricity theft detection using smart meter data [C]. Innovative Smart Grid Technologies Conference, IEEE, 2015: 1-5.
- [6] 方菁, 方众. 基于数据挖掘技术的窃电侦查方法[J]. 上海电力, 2013(2): 85-90.
- FANG J, FANG ZH. Power stealing detection method based on Data Mining Technology [J]. Shanghai Electric Power, 2013(2): 85-90.
- [7] MONEDERO I, BISCARRI F, LEÓN C, et al. Detection of frauds and other non-technical losses in a power utility using Pearson coefficient, Bayesian networks and decision trees [J]. International Journal of Electrical Power & Energy Systems, 2012, 34(1): 90-98.
- [8] 程超, 张汉敬, 景志敏, 等. 基于离群点算法和用电信息采集系统的反窃电研究[J]. 电力系统保护与控制, 2015(17): 69-74.
- CHAO CH, ZHANG H J, JING ZH M, et al. Study on the anti-electricity stealing based on outlier algorithm and the electricity information acquisition system [J]. Power System Protection & Control, 2015, 43(17): 69-74.
- [9] 张学工. 关于统计学习理论与支持向量机[J]. 自动化学报, 2000, 26(1): 32-42.
- ZHANG X G. Introduction to statistical learning theory and support vector machines [J]. Acta Automatica Sinica, 2000, 26(1): 32-42.
- [10] 杨浩, 张磊, 何潜, 等. 基于自适应模糊 C 均值算法的电力负荷分类研究[J]. 电力系统保护与控制, 2010, 38(16): 111-115.
- YANG H, ZHANG L, HE Q, et al. Study of power load classification based on adaptive fuzzy C means [J]. Power System Protection & Control, 2010, 38(16): 111-115.
- [11] 赵国生, 牛贞贞, 刘永光, 等. 基于自适应模糊 C 均值聚类算法的电力负荷特性分类[J]. 郑州大学学报(工学版), 2015, 36(6): 56-60.
- ZHAO G SH, NIU ZH ZH, LIU Y G, et al. Classification of power load characteristics based on

- adaptive fuzzy C-means clustering algorithm[J]. Journal of Zhengzhou University (Engineering Science), 2015, 36(6):56-60.
- [12] 刘燕驰, 高学东, 国宏伟, 等. 聚类有效性的组合评价方法[J]. 计算机工程与应用, 2011, 47(19):15-17.
LIU Y CH, GAO X D, GUO H W, et al. Ensembling clustering validation indices[J]. Computer Engineering & Applications, 2011, 47(19):15-17.
- [13] 周开乐, 杨善林, 丁帅, 等. 聚类有效性研究综述[J]. 系统工程理论与实践, 2014, 34(9):2417-2431.
ZHOU K L, YANG SH L, DING SH, et al. On cluster validation[J]. Systems Engineering-Theory & Practice, 2014, 34(9):2417-2431.
- [14] 蒋国栋. 基于 K-means 聚类算法的负荷模型研究[D]. 哈尔滨: 哈尔滨工业大学, 2009.
JIANG G. Research on load model based on K - means clustering algorithm [D]. Harbin: Harbin Institute of Technology, 2009.
- [15] 黄为勇, 王义, 田秀玲, 等. 一种基于 SVR 的发动机多模式故障诊断方法[J]. 仪器仪表学报, 2015, 36(9):2112-2119.
HUANG W Y, WANG Y, TIAN X L, et al. Multi-mode engine fault diagnosis method based on SVR[J]. Journal of Instrument & Instrumentation, 2015, 36(9):2112-2119.
- [16] 闫国华, 朱永生. 支持向量机回归的参数选择方法[J]. 计算机工程, 2009, 35(14):218-220.
YAN G H, ZHU Y SH. Parameters selection method for support vector machine regression [J]. Computer Engineering, 2009, 35(14):218-220.
- [17] 翟旭平, 杨兵兵, 孟田. 基于 PCA 和混合核函数 QPSO_SVM 频谱感知算法[J]. 电子测量技术, 2016, 39(9):87-90.
ZHAI X P, YANG B B, MENG T. Based on PCA and mixedkernel function QPSO_SVM spectral sensingalgorithm[J]. Electronic Measurement Technology, 2016, 39(9):87-90.
- [18] 董方. 粒子群算法研究及其在动态优化中的应用[D]. 杭州: 浙江大学, 2014.
DONG F. Particle swarm algorithm and its application in dynamic optimization [D]. Hangzhou: Zhejiang University, 2014.
- [19] 梁旭, 刘才慧. 基于混合粒子群算法的在线检测路径规划[J]. 国外电子测量技术, 2015, 34(12):30-34.
LIANG X, LIU C H. On - line detection path planning based on hybrid particle swarm optimization[J]. Foreign Electronic Measurement Technology, 2015, 34(12):30-34.

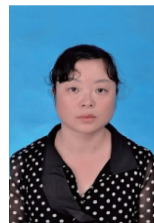
作者简介



康宁宁, 2015 年于安阳师范学院获得学士学位, 现为昆明理工大学硕士研究生, 研究方向为信息物理融合研究。

Email: hnwyknn@163.com

Kang Ningning received B. Sc. from Anyang Normal University in 2015. Now she is a M. Sc. candidate in Kunming University of Science and Technology. Her main research direction is research on information physics fusion.



李英娜(通信作者), 1996 年于云南大学获得学士学位。2009 年获得于昆明理工大学获得硕士学位。现为昆明理工大学信息工程与自动化学院副教授, 主要从事传感网组建与信息集成和智能分析研究。

E-mail: 1348415800@qq.com

Li Yingna (Corresponding author) received B. Sc. from Yunnan University in 1996, and M. Sc. from Kunming University of Science and Technology in 2009, respectively. Now she is an associate professor in School of Information Engineering and Automation, Kunming University of Science and Technology. Her main research interest includes sensor network construction and information integration and intelligent analysis.