

一种基于组合算法的异常用电模式辨识方法

袁翔宇¹, 张蓬鹤¹, 熊素琴¹, 赵波², 李求洋¹

(1. 中国电力科学研究院有限公司 北京 100192; 2. 北京信息科技大学 北京 100192)

摘要:针对电力用户异常用电的检测问题,提出了一种基于无监督组合算法的异常用电模式辨识方法。所提辨识方法由数据处理、特征提取、离群检测三部分组成。文中先获取用户的用电量及相关数据,进行数据清洗和缺失数值补全;再对数据进行特征提取,得到相应的异常用电识别特征量;通过 k 均值聚类将用户聚为两组,并分别对每组进行主成分分析优化特征空间,计算离群邻近度,通过 2 sigma 原则实现异常用电用户辨识。该方法通过聚类、优化特征空间、离群检测组合算法,提高了辨识效率。文中采用真实用电数据进行了异常用电用户辨识仿真实验,辨识结果验证了该方法的有效性。

关键词:异常用电; k 均值聚类; 主成分分析; 离群邻近度; 欧几里得距离; 2 sigma 原则

DOI:10.19753/j. issn1001-1390. 2023. 06. 023

中图分类号:TM933

文献标识码:B

文章编号:1001-1390(2023)06-0160-07

Identification of abnormal power consumption mode based on combination algorithm

Yuan Xiangyu¹, Zhang Penghe¹, Xiong Suqin¹, Zhao Bo², Li Qiuyang¹

(1. China Electric Power Research Institute Co., Ltd., Beijing 100192, China.

2. Beijing Information Science and Technology University, Beijing 100192, China)

Abstract: In a bid to detect abnormal electricity consumption of power users, a method for identifying abnormal electricity consumption mode based on unsupervised combination algorithm is proposed in this paper. The proposed identification method consists of three parts of data processing, feature extraction and outlier detection. The power consumption and related data of the users are obtained, and the data is cleaned and the missing value is supplemented, feature extraction is carried out on the data to obtain the corresponding features for abnormal electricity consumption recognition. Afterwards, k -means is used to cluster the users into two groups, and principal component analysis is performed on each group to optimize the feature space, the outlier proximity is calculated, and abnormal power consumption users are identified by 2 sigma principle. This method improves the identification efficiency by combining clustering, optimization of feature space and outlier detection. The simulation experiment of abnormal power consumption user identification is carried out with real power consumption data, and the identification results verify the effectiveness of the proposed method.

Keywords: abnormal power consumption, k -means clustering, principal component analysis, outlier proximity, Euclidean distance, 2 sigma principle

0 引言

近几年,智能电网成为当今世界的一个关键问题,并吸引了能源用户、政策制定者和工程师的日益关注。而智能电网的运行过程中会出现一定量的能量损失,其中可以用线路损失解释的部分称为技术性损失;不能被技术损失所解释的部分称为非技术性损失。非技术性损失指剔除技术损失后无法用技术解释的电网输

配电损失部分。电力用户的异常用电行为,比如窃电,则是造成非技术性损失的主要原因^[1]。

异常用电行为日益增加,这不仅增加了电力公司的经济损失,也严重影响了经济建设和社会稳定。据不完全统计,中国福建省每年因异常用电行为造成的电费损失可达 1 亿人民币^[2]。所以应当对异常用电行为进行检测打击,对异常用电用户进行严惩。目前的反异常用电技术不够完善,对于较为常见的异常用电

基金项目:国家电网有限公司总部科技项目(5442JL170007)

手段,无法做出准确的判断。

异常用电监测包括硬件监测和数据处理分析两方面的技术。硬件监测技术主要有智能表检测异常用电^[3]和电量计量采集系统^[4]。数据分析主要分为有监督学习^[5-8]和无监督学习^[9-15]两类。有监督算法条件较为苛刻,无法在没有数据和数据缺失的条件下使用。无监督学习对数据量的要求并不是很高,适用面更加广泛。文献[11]提出了一种融合聚类和异常点检测的算法;为了提高算法的查准率,文献[12]提出了一种基于方形邻域和裁剪因子的离群点检测算法;文献[13]采用基于密度的聚类算法检测电力用户异常用电;文献[14]提出一种基于高斯核的离群点检测方法对电力用户进行分类。文献[15]提出了一种基于距离的离群点检测法判断窃电。

为了进一步提高识别准确率,文中提出了一种新型无监督的异常用电检测算法,主要包括特征提取、聚类分析、主成分分析,优化特征空间和离群点检测。首先提取部分地区用电量及线路相关数据,进行数据处理,然后从数据集中提取出电量下降趋势指标、线损指标和告警类指标作为异常用电评判体系,再进行聚类分析、主成分分析和离群邻近度的计算,最后通过离群点检测辨识异常用电用户。

1 异常用电辨识流程

文中所提异常用电辨识方法主要有以下几个部分:数据处理、特征提取和离群检测。

异常用电用户辨识流程,主要包括以下步骤:

(1) 数据处理:考虑到部分用户是不存在发生窃电行为和部分数据存在缺失,所以对这些数据进行清洗和缺失数值处理;

(2) 特征提取:提取电量下降趋势指标、线损指标和告警类指标三个指标;

(3) 离群检测:先使用改进的 k 均值聚类算法将数据聚为两组。然后分别进行主成分分析,优化特征空间。然后对每组数据进行离群邻近度求解,采用 2 sigma 原则对离群用户进行筛选。

2 基于组合算法的异常用电辨识

2.1 数据处理

从电力计量装置中提取用户的用电负荷和客户端电能表终端报警次数等原始信息数据。异常用电用户在用电用户中只占有很小的一部分,即大部分用户是不可能发生窃电,比如说:银行、学校、税务、工商等非居民类别的用户数据。所以需要对这些用户数据进行清洗,即剔除这些用户的数据。

由于智能电能表故障、测量数据传输不可靠、系统维护不定时、存储问题等多种原因^[16],导致存在部分缺失数值,直接舍去会减少数据的真实性,所以采用拉格朗日插值方法进行补全。公式如下:

$$L_n(x) = \sum_{i=0}^n l_i(x) y_i \quad (1)$$

$$l_i(x) = \prod_{j=0, j \neq i}^n \frac{x - x_j}{x_i - x_j} \quad (2)$$

其中 x 为缺失数值对应的下标序号; $L_n(x)$ 为缺失数值的插值后的结果; x_i 为非缺失数值 y_i 的下标序号, n 为 10, 即采用缺失数据前后各 5 个数据, 共 10 个数据, 进行插值补全。对于所有含有缺失数据的用户依次进行插补, 直到不存在缺失为止。

2.2 特征提取

经过插值处理后, 所获取的用电量数据已不存在缺失数据。但作为用电特征, 这组数据并不充分, 所以文中对数据进行了相应地分析。

对用户的用电量信息进行分析后, 发现大部分正常用电用户的用电量变化较为平稳, 而窃电用户的用电量往往短期内呈下降趋势, 后期趋于稳定; 同时对用户的供电线路损失电量进行分析, 发现窃电用户线路损失明显高于正常用户; 最后统计用户的电能表故障情况, 同样发现窃电用户的电能表故障次数偏高。基于以上分析, 文中采用相关的异常用电评判体系对补全后的数据进行特征提取, 得到电量下降趋势指标、线损指标和告警类指标^[17]。

2.2.1 电量下降趋势指标

如果用户的用电量正常, 则用电量数据具有一定规律性。通常用户开始发生窃电行为, 用电量就会降低或波动。为了描述这个过程, 文中采用电量下降趋势指标作为衡量标准。对于某天的电量下降趋势指标, 设置该天前 5 天、后 5 天以及该天作为统计窗口期。先计算 11 天中每天的用电量趋势, 其中第 i 天的用电量趋势 k_i 的计算公式如下:

$$k_i = \frac{\sum_{l=i-5}^{i+5} (f_l - \bar{f})(l - i)}{\sum_{l=i-5}^{i+5} (l - i)^2} \quad (3)$$

$$\bar{f} = \frac{1}{11} \sum_{l=i-5}^{i+5} f_l \quad (4)$$

其中 k_i 为第 i 天的用电量趋势; f_l 为第 l 天的用电量。

如果电量趋势不断下降, 则认为该用户有可能窃

电。然后统计用户 11 天内,当天比前一天用电递减的天数,即令:

$$D(i) = \begin{cases} 1, & k_i < k_{i-1} \\ 0, & k_i \geq k_{i-1} \end{cases} \quad (5)$$

则 11 天内的电量趋势下降指标为^[17]:

$$T = \sum_{n=i-4}^{i+5} D(n) \quad (6)$$

2.2.2 线损指标

线路损失同样也是衡量异常用电的重要指标。文中采用线路供电量和线路用户的总用电量之间的差值百分比,作为线路损失,公式如下^[17]:

$$V_l = \frac{s_l - \sum_m f_{l,m}}{s_l} \times 100\% \quad (7)$$

式中 s_l 为第 l 天线路供电量; $\sum_m f_{l,m}$ 为线路上各个用户的总用电量; m 为该线路用户总数。

考虑到线损均有波动性,同电量下降趋势指标类似,文中统计当日的前 5 天和后 5 天为窗口期。记当天和前五天的线损率平均值为 V_i^1 和当天与后 5 天之间的线损率平均值 V_i^2 。若 V_i^1 比 V_i^2 增长率多 1%, 则认为该用户具有一定的嫌疑,故定义线损指标如下^[17]:

$$E(i) = \begin{cases} 0, & \frac{V_i^1 - V_i^2}{V_i^2} > 1\% \\ 1, & \frac{V_i^1 - V_i^2}{V_i^2} \leq 1\% \end{cases} \quad (8)$$

2.2.3 告警类指标

以用电用户为单位,统计某段时间内电流异常、电压异常、功率因素异常和有功异常等发生次数。由于这类情况发生时会发生终端报警,所以选取终端报警次数作为告警类指标。

2.3 基于组合算法的离群检测

该算法为无监督算法,由聚类、主成分分析优化特征空间和离群检测三个部分组成。先用聚类算法将用户数据聚为两组,再进行主成分分析优化特征空间,然后计算离群邻近度,最后用 2 sigma 原则进行离群检测。

2.3.1 k 均值聚类

k 均值算法,是一种聚类算法。首先确定聚类数目 k ,再从样本集中随机选择 k 个样本作为“簇中心”,并计算所有样本与这 k 个“簇中心”的欧几里得距离。然后对于剩余的每一个样本,都将被划分到与其最近“簇中心”的那一簇中。然后对于每个簇,文中重新计算“簇中心”^[18-20]。如果计算的“簇中心”与初始“簇中

心”相等,则聚类完成。反之,则重新随机生成 k 个“簇中心”,重复步骤,直到相等。

改进的 k 均值算法,则对 k 均值生成“簇中心”的方式做了一定的改进。为了减少“簇中心”的生成次数,当已经生成了 n 个初始“簇中心”后,则第 $n+1$ 个“簇中心”应该要尽可能远离前 n 个“簇中心”。所以改进后的 k 均值算法的设计步骤为:

步骤 1: 从数据集中随机选取一个样本(样本具有三个特征,电量下降趋势指标、线损指标和告警类指标)作为初始化“簇中心”;

步骤 2: 计算每一个样本与已有的“簇中心”的最短距离,并用 $D(x)$ 表示;

步骤 3: 通过式(9)计算每个样本作为下一个“簇中心”的概率,并选择概率最大的样本作为下一个“簇中心”;

$$p(x) = \frac{D^2(x)}{\sum_x D^2(x)} \quad (9)$$

步骤 4: 重复步骤 2 和步骤 3,直到生成 k 个“簇中心”;

步骤 5: 对于每个簇,重新计算新的“簇中心”。如果计算的“簇中心”与初始“簇中心”相等,则聚类完成。反之,则返回步骤 1。

文中选择两个“簇中心”,聚类运算后,将用电用户分为两簇,然后进行后续处理。

2.3.2 特征空间优化

为了更好地对数据进行处理,文中采用主成分分析,来优化特征空间。

主成分分析(Principal Component Analysis,简称 PCA)是一种分析、简化数据集的方式,常用来降低数据的维度,突出协方差贡献最大的特征量。PCA 的思想是将 n 维空间特征映射到 k 维($k < n$)空间中,得到样本集的新的 k 维特征,从而达到优化特征空间的目的。

算法步骤如下:

步骤 1: 将样本矩阵中心化;

步骤 2: 采用中心化后的样本矩阵,构建协方差矩阵;

步骤 3: 计算协方差矩阵的特征值、特征向量以及特征量的贡献率;

步骤 4: 据特征值贡献率,绘制贡献率条形图;

步骤 5: 取累计贡献率达到 85% 以上最少数目的特征量作为主成分,利用其特征向量对样本矩阵进行变换,得到新的样本数据。

通过上述主成分分析,得到新特征空间,再在新的特征空间里对数据进行后续处理。文中通过对电量下降趋势指标、线损指标和告警类指标进行主成分处理后,得到两个新特征量作为新的特征空间。

2.3.3 离群邻近度分析

经过主成分分析后,得到了两组数据的新特征量。将得到的新的两组数据分别进行离群检测。离群邻近度是一个衡量离群程度的参量,多用来辨识一组用户中的部分异常用户。

2 sigma 原则:数值分布在满足($\mu - 2\sigma, \mu + 2\sigma$)的区间内的概率为 0.9544。小概率事件通常指发生概率小于 5% 的事件,认为是不可能发生事件。即此处离群程度不满足 2 sigma 原则的用户为用电异常用户,将该类用户判断为异常用电用户。

算法步骤如下:

步骤 1:计算每组数据中各个样本之间的距离,记作 $D(i,j)$,即第 i 个样本与第 j 个样本之间的距离;

步骤 2:计算每组用户之间的所有距离的平均值作为邻域半径 r_d ;

$$r_d = \frac{1}{(n-1)!} \sum_{i,j \in \text{用户}, i \neq j} D(i,j) \quad (10)$$

步骤 3:统计每一个样本的邻域半径内样本的数量,作为离群邻近度 N_{dist} ,计算如下:

$$D(j) = \begin{cases} 1, & D(i,j) < r_d \\ 0, & D(i,j) \geq r_d \end{cases} \quad (11)$$

$$N_{dist}(i) = \sum_{j=0, i \neq j}^{j=n} D(j) \quad (12)$$

式中 $N_{dist}(i)$ 为第 i 个用户的离群邻近度;

步骤 4: N_{dist} 为衡量用户离群程度参量,文中采用 2 sigma 原则对用户进行离群划分,判断($\mu - 2\sigma, \mu + 2\sigma$)区间外的用户为离群用户,也称为窃电用户。

2.4 算法评判方法

异常用电用户检测过程实际是一个二分类问题,将异常用电用户和正常用电用户区分开来。混淆矩阵、ROC(Receiver Operating Characteristic)曲线和 AUC(Area Under Curve)三个参数常用来评判二分类的好坏。

2.4.1 混淆矩阵

混淆矩阵,也称为误差矩阵,一个 n 行 n 列的矩阵,表示识别精度。

表 1 为一个简易 $2 * 2$ 的混淆矩阵,其中行(Positive/Negative)为实际所属的类别,列(Ture/False)表示预测的类别。

表 1 混淆矩阵

Tab. 1 Confusion matrix

用户	检测为正常用户	检测为异常用户
实际正常用户	True Positive (TP)	False Positive (FP)
实际异常用户	False Negative (FN)	True Negative (TN)

其中 FP 是第一类错误,即将实际是异常用电用户被判断为正常用户;FN 是第二类错误,即将正常用户判定为异常用电用户。总而言之,TP 和 TN 的比重越高,判断的准确率就越高。

考虑到异常用电检测为不平衡分类,文中选用 $f_measure$ 标准作为分类方法的评价指标。 $f_measure$ 标准包括查全率(recall),即正类正确率和查准率(precision)。该标准均同时考虑到了正类和负类,且适合于评价不平衡数据分类效果,该值越高分类器的整体性能越好。其计算公式如下所示:

$$\left\{ \begin{array}{l} recall = TPR = \frac{TP}{TP + FN} \\ precision = \frac{TP}{TP + FP} \\ f_measure = \frac{(1 + \beta) * TPR * precision}{\beta^2 (TPR + precision)} \end{array} \right. \quad (13)$$

式中 β 通常取 1。

2.4.2 ROC 曲线和 AUC 值

ROC 曲线描述混淆矩阵中 TPR(真正类率)和 FPR(反正类率)两个指标增长速率的相对关系。就二分类来看,二分法可以分为正类和负类,大于阈值为负类,小于阈值为正类。小的阈值虽然能识别出更多的负类,使 TPR 提高;但同时也可能会导致将更多的正类错误地划分为负类,使 FPR 提高。ROC 曲线能形象反映这一关系,即曲线上每一个点对应分类器取某个阈值时分类结果的混淆矩阵。

在 ROC 空间坐标中,ROC 曲线与 $y=0$,和 $x=1$ 曲线围成的面积为 AUC 值。AUC 值越大,越接近于 1,表示分类效果越好。

3 算例及分析

为了检验模型的准确性以及可行性,文中提取了某市电网台区电能表及部分用户电能表用电量信息,进行数据处理,提取相应的指标(电量下降趋势指标、线损指标和告警类指标)后,随机抽取 280 组数据进行建模和检验。该检验在 Window 10 系统 MATLAB 2016a 环境下进行仿真验证。

3.1 实验与结果

文中对 280 组用户数据进行中心化处理,使用户

特征更加集中。通过改进的 k 均值聚类分析将用户分为两组(A 组和 B 组),并采用主成分分析,得到两组用户的特征贡献率条形图,如图 1(a)所示(特征 1 是电量下降趋势,特征 2 为告警类指标,特征 3 为线损指标)。可知,第三个特征可以进行降维。处理后,得到两组(A'组和 B'组)用户新特征量,并用特征 1 和特征 2 表示,如图 1(b)所示。

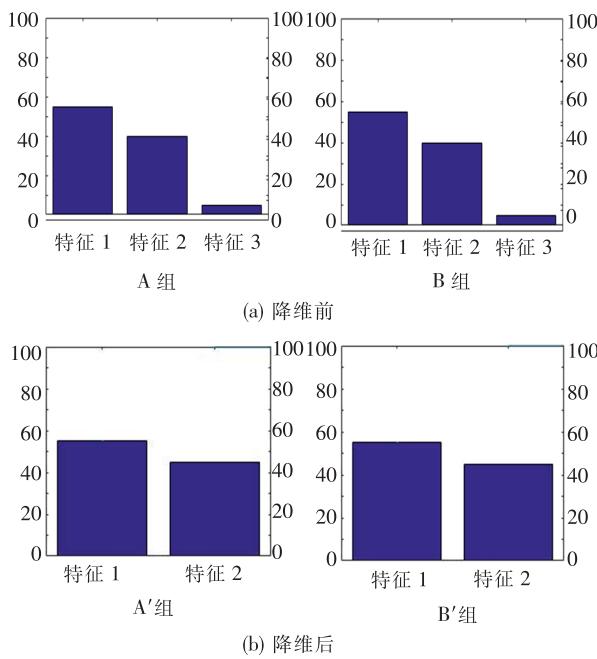


图 1 降维前和降维后的特征量贡献率条形图

Fig. 1 Bar chart of feature contribution rate before and after dimensionality reduction

然后,计算 A'组和 B'组中各个特征向量之间的欧几里得距离,取各组欧几里得距离的平均值作为各组的邻域半径 r_d ,计算各个特征向量的离群邻近度 N_dist 。

分别对两组用户的离群邻近度进行正态函数的标准化处理,筛选出 2σ 原则以外的用户作为异常用电用户。

离群点检测前,绘制 A'组和 B'组的散点图,见图 2。经过对比发现,A'组用户中异常用电远多于 B'组异常用电用户(A'组异常用电 28 个,B'组异常用电 4 个)。根据异常用电用户比重不同,设置 A'组和 B'组的离群个数分别为 28 和 4 个,即计算离群邻近度后,判断 A'组用户中 $\mu \pm 2\sigma$ 外和 B'组用户中的 $\mu \pm 2\sigma$ 外的若干组数据为异常用电数据。

检测完成后绘制混淆矩阵,如图 3 所示,即总的异常用电识别率到了 90.7%,实际为正常用电的用户识别正常的概率为 94.4%。ROC 曲线如图 4 所示,其 AUC 值接近 0.8。文中计算其 f -measure 指标 F_1 为 0.9474。

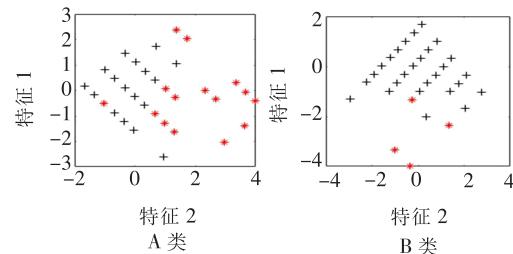


图 2 用户散点图

Fig. 2 Scatter plot of users

Confusion Matrix		
		Output Class
Target Class	1	2
	234 83.6%	14 5.0%
1	12 4.3%	20 7.1%
	95.1% 4.9%	58.8% 41.2%
		94.4% 5.6%
		62.5% 37.5%
		90.7% 9.3%

图 3 检测结果的混淆矩阵

Fig. 3 Confusion matrix of test results

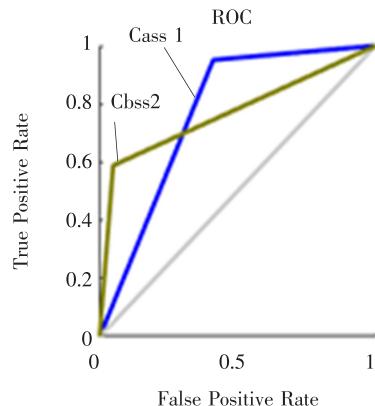


图 4 检测结果的 ROC 曲线

Fig. 4 ROC curve of test results

3.2 对比分析

文中将改进的 k 均值聚类算法和离群算法相结合,先改进的 k 均值聚类聚为 2 组,然后主成分降维得到新的两组数据,进行离群点检测。为了更好地体现聚类和离群组合算法的优越性,按照文献[15]直接采用基于距离的离群算法对数据进行了再次分析。

首先将数据中心化,随后将中心化的数据进行主成分分析。经过成分分析之后,三个特征贡献率接近 1:1:1(见图 5),只能在该特征空间进行数据的离群检测(其中特征 1 是电量下降趋势,特征 2 为线损指标,

特征 3 为告警类指标);随即进行欧几里得距离、邻域半径、离群邻近度计算和离群检验,检验结果如图 6~图 7 所示。图 6 显示的混淆矩阵中识别准确率为 86.8%,特别是异常用电中判断为异常用电的数量仅有 50%。就 ROC 曲线来看,AUC 值为 0.6。同时也计算其 $f_{measure}$ 指标 F_2 ,得: $F_2 = 0.9246$ 。

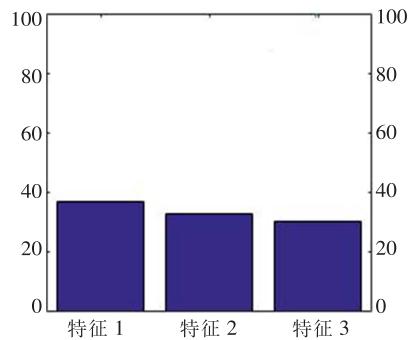


图 5 主成分占比图

Fig. 5 Principal component ratio map

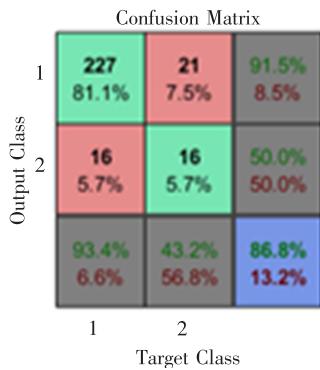


图 6 检测结果的混淆矩阵

Fig. 6 Confusion matrix of test results

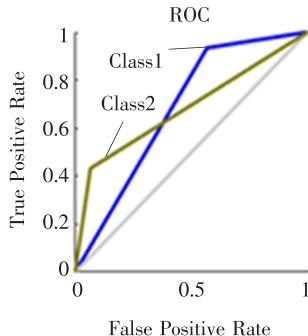


图 7 检测结果的 ROC 曲线

Fig. 7 ROC curve of test results

而采用文中所提组合算法,先通过改进的聚类 k 均值聚成两组后,进行 PCA 分析得到新的两组特征量,形成了新的特征空间。经过离群的 2σ 原则进行筛选,

实现了更好的检测效果。识别准确率也达到了 90.7%, ROC 曲线的 AUC 值也近乎 0.8。同时, $F_1 > F_2$, 组合算法的分类效果更好。

分析对比表明,在进行异常用电模式检测时,基于无监督组合算法的检测方法比直接离群检测的方法识别效果更好。

4 结束语

针对异常用电的检测问题,提出了一种基于无监督组合算法的异常用电模式辨识方法。该方法辨识异常用电模式的步骤包括数据处理、特征提取、离群检测。先获取用户的用电量及相关数据,进行数据清洗和缺失数值补全;再进行特征提取,得到相应的异常用电识别特征量;然后通过聚类将用户聚为两组,并分别对每组进行主成分分析优化特征空间;最后通过离群检测实现异常用电用户辨识。对真实用电数据进行了异常用电用户辨识仿真实验,辨识结果验证了该方法的有效性。

相较于直接进行离群点检测,所提方法通过聚类、优化特征空间、离群检测组合算法,提高了辨识效率。考虑到不同的算法之间的组合效果不一,后续研究不同算法之间的组合,希望进一步提高异常用电模式辨识方法的识别率以及应用范围。

参 考 文 献

- [1] Jiang R, LuR, Wang Y, et al. Energy-theft detection issues for advanced metering infrastructure in smart grid[J]. Tsinghua Science and Technology, 2014, 19(2): 105-120.
- [2] 陈启鑫, 郑可迪, 康重庆, 等. 异常用电的检测方法: 评述与展望 [J]. 电力系统自动化, 2018, 42(17): 189-199.
Chen Qixin, Zheng Kedi, Kang Chongqing, et al. Detection methods of abnormal electricity consumption behaviors: review and prospect [J]. Automation of Electric Power Systems, 2018, 42(17): 189-199.
- [3] 熊德智, 陈向群, 陈奕蕾, 等. 智能用电监控装置的研制 [J]. 电测与仪表, 2019, 56(5): 143-147, 152.
Xiong Dezhi, Chen Xiangqun, Chen Yilei, et al. Development of intelligent electricity monitoring device [J]. Electrical Measurement & Instrumentation, 2019, 56(5): 143-147, 152.
- [4] 胡圣尧, 关静, 杨子立, 等. 基于嵌入式的电量计量采集系统研究 [J]. 现代电子技术, 2016, (22): 163-166, 170.
Hu Shengyao, Guan Jing, Yang Liping, et al. Research on the embedded power metering and collection system [J]. Modern Electronics Technique, 2016, (22): 163-166, 170.
- [5] 周明, 宋旭帆, 涂京, 等. 基于非侵入式负荷监测的居民用电行为分析 [J]. 电网技术, 2018, 42(10): 3268-3276.
Zhou Ming, Song Xufan, Tu Jing, et al. Residential Electricity Consumption Behavior Analysis Based on Non-Intrusive Load Monitoring [J]. Power System Technology, 2018, 42(10): 3268-3276.

- [6] 于希娟, 孙宏伟. 基于图像处理和半监督学习的变电设备故障诊断[J]. 电网与清洁能源, 2022, 38(8): 60-68.
Yu Xijuan, Sun Hongwei. Fault Diagnosis of Substation Equipment Based on Image Processing and Semi-Supervised Learning [J]. Power System and Clean Energy, 2022, 38(8): 60-68.
- [7] 白浩, 潘姝慧, 邵向潮, 等. 基于小波去噪与随机森林的配电网高阻接地故障半监督识别方法[J]. 电力系统保护与控制, 2022, 50(20): 79-87.
Bai Hao, Pan Zhuhui, Shao Xiangchao, et al. A high impedance grounding fault semi-supervised identification method based on wavelet denoising and random forest [J]. Power System Protection and Control, 2022, 50(20): 79-87.
- [8] 卢克斌, 殷守林. 一种端到端弱监督学习网络模型的中国画情感识别[J]. 哈尔滨理工大学学报, 2022, 27(1): 69-78.
Lu Kebin, Yin Shoulin. Chinese Paintings Emotion Recognition Based on End-to-end Weakly Supervised Learning Network Model [J]. Journal of Harbin University of Science and Technology, 2022, 27(1): 69-78.
- [9] 王毅, 丁力, 侯兴哲, 等. 基于层次分析法的加权线异常用电检测方法[J]. 科学技术与工程, 2017, 17(33): 96-103.
Wang Yi, Ding Li, Hou Xingzhe, et al. Weighted LOF stealing detection method based on analytic hierarchy process [J]. Science Technology and Engineering, 2017, 17(33): 96-103.
- [10] 庄池杰, 张斌, 胡军, 等. 基于无监督学习的电力用户异常用电模式检测[J]. 中国电机工程学报, 2016, 36(2): 379-387.
Zhuang Chijie, Zhang Bin, Hu Jun, et al. Anomaly detection for power consumption patterns based on unsupervised learning [J]. Proceedings of the CSEE, 2016, 36(2): 379-387.
- [11] 张彼德, 洪锡文, 刘俊, 等. 基于无监督学习的MMC子模块开路故障诊断方法[J]. 电力系统保护与控制, 2021, 49(12): 98-105.
Zhang Bide, Hong Xiwen, Liu Jun, et al. Diagnosis method for submodule open-circuit fault in modular multilevel converter based on unsupervised learning [J]. Power System Protection and Control, 2021, 49(12): 98-105.
- [12] 涂晓敏, 石鸿雁. 基于方形邻域和裁剪因子的离群点检测方法[J]. 小型微型计算机系统, 2019, 40(1): 186-189.
Tu Xiaomin, Shi Hongyan. Square neighborhood and pruning factor based outlier detection algorithm [J]. Journal of Chinese Computer Systems, 2019, 40(1): 186-189.
- [13] 田力, 向敏. 基于密度聚类技术的电力系统用电量异常分析算法[J]. 电力系统自动化, 2017, 41(5): 64-70.
Tian Li, Xiang Min. Abnormal power consumption analysis based on density-based spatial clustering of applications with noise in power systems [J]. Automation of Electric Power Systems, 2017, 41(5): 64-70.
- [14] 孙毅, 李世豪, 崔灿, 等. 基于高斯核函数改进的电力用户用电数据离群点检测方法[J]. 电网技术, 2018, 42(5): 1595-1606.
Sun Yi, Li Shihao, Cui Can, et al. Improved outlier detection method of power consumer data based on gaussian kernel function [J]. Power System Technology, 2018, 42(5): 1595-1606.
- [15] 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, 43(17): 69-74.
Cheng Chao, Zhang Hanjing, Jing Zhimin, et al. Study on the anti-electricity stealing based on outlier algorithm and the electricity information acquisition system [J]. Power System Protection and Control, 2015, 43(17): 69-74.
- [16] ZHENG Z B, YANG Y T, NIU X D, et al. Wide and Deep Convolutional Neural Networks for Electricity-Theft Detection to Secure Smart Grids [J]. IEEE Transactions on Industrial Informatics, 2018, 14(4): 1606-15.
- [17] 张良均, 杨坦, 肖刚, 等. MATLAB 数据分析与数据挖掘实战 [M]. 北京: 机械工业出版社, 2015: 140-159.
- [18] 孙毅, 冯云, 崔灿, 等. 基于动态自适应K均值聚类的电力用户负荷编码与行为分析[J]. 电力科学与技术学报, 2017, 32(3): 3-8.
Sun Yi, Feng Yun, Cui Can, et al. Power user load code and behavior analysis based on dynamic adaptive k-means clustering [J]. Journal of Electric Power Science and Technology, 2017, 32(3): 3-8.
- [19] 姚黄金, 雷霞, 付鑫权, 等. 基于改进自适应密度峰值算法的日负荷曲线聚类分析[J]. 电力系统保护与控制, 2022, 50(3): 121-130.
Yao Huangjin, Lei Xia, Fu Xinquan, et al. Cluster analysis of daily load curves based on an improved self-adaptive density peak clustering algorithm [J]. Power System Protection and Control, 2022, 50(3): 121-130.
- [20] 张朝龙, 赵筛筛, 章博. 基于因子分析与K-means聚类的退役动力电池快速分选方法[J]. 电力系统保护与控制, 2021, 49(12): 41-47.
Zhang Chaolong, Zhao Shaishi, Zhang Bo. A fast classification method based on factor analysis and K-means clustering for retired electric vehicle batteries [J]. Power System Protection and Control, 2021, 49(12): 41-47.

作者简介:



袁翔宇(1988—),男,工程师,硕士,主要从事电测量技术、智能传感和检测设备自动化研究。
Email: yuanxiangyu@epri.sgcc.com.cn

张蓬鹤(1978—),女,教授级高级工程师,博士,主要从事电测量技术、电能表及元器件性能分析研究。

熊素琴(1979—),女,高级工程师,硕士,主要从事电测量技术和电能质量研究。

赵波(1977—),男,教授,博士,主要从事电能质量与电力电子技术研究。

李求洋(1988—),女,工程师,博士,主要从事智能传感技术研究。

收稿日期:2020-05-14;修回日期:2022-12-27

(田春雨 编发)