

# 基于 Mean-shift 聚类 and 孪生网络的非侵入式 负荷辨识方法

石少青<sup>1</sup>, 张乐平<sup>2</sup>, 张本松<sup>2</sup>, 吴昊文<sup>2</sup>, 连新凯<sup>3</sup>, 杜书帅<sup>3</sup>

(1. 中国南方电网有限责任公司, 广东 广州 510080; 2. 南方电网数字电网研究院有限公司, 广东 广州 510663; 3. 河南许继仪表有限公司, 河南 许昌 461000)

**摘要:** 为了面向智能电表负荷辨识技术的推广应用, 提出一种基于 Mean-shift 聚类和孪生网络的非侵入式负荷聚类辨识方法, 该方法针对所检测到的负荷事件, 提取有功-无功等电量特征, 并将其按运行时刻和运行时长进行细分, 获得负荷事件在时间特征上的分布, 并作为辅助特征。其次, 对时间特征采用 Mean-shift 进行聚类, 得到具有相同时间特征的负荷类别, 并采用孪生网络与负荷特征数据库匹配识别, 最终实现负荷辨识。最后, 通过实际家庭用户的负荷数据测试, 实验证明了所提方法能够有效辨识用户家庭内部的负荷类别, 从而为智能电表负荷辨识推广奠定基础。

**关键词:** 非侵入式; 时间特征; 智能电表; Mean-shift 聚类; 孪生网络

**中图分类号:** TM933 **文献标识码:** A **DOI:** 10.19457/j.1001-2095.dqcd23589

## Non-intrusive Load Identification Method Using Mean-shift Clustering Algorithm with Siamese Network

SHI Shaoqing<sup>1</sup>, ZHANG Leping<sup>2</sup>, ZHANG Bensong<sup>2</sup>, WU Haowen<sup>2</sup>, LIAN Xinkai<sup>3</sup>, DU Shushuai<sup>3</sup>

(1. China Southern Power Grid Company Limited, Guangzhou 510080, Guangdong, China; 2. Digital Grid Research Institute of China Southern Power Grid Research Institute, Guangzhou 510663, Guangdong, China; 3. Henan XJ Metering Co., Ltd., Xuchang 461000, Henan, China)

**Abstract:** In order to apply the smart meters with load identification technology into real application, a non-intrusive load identification method using Mean-shift clustering algorithm with siamese network framework was proposed. In this method, the load event was carried out for extracting the features, including active power and reactive power. Meanwhile, working time and working duration of the load was obtained to assist the load identification. Secondly, Mean-shift clustering was used to classify the load events into the load type, and the load type was then identified by comparing the siamese network with the load signature database. Thus, the load type which the load event belongs to can be obtained. Finally, through experiments on the load test data of actual household users, the result proved that the method proposed could obtain the load kinds within the user's home, and distinguish the load with a large coupling degree of power signatures which could provide the basis for application in smart meter.

**Key words:** non-intrusive; time signature; smart meter; Mean-shift clustering; siamese network

随着新一代智能电表的推广和应用, 非侵入式负荷辨识技术作为一种新兴的智能化负荷监测手段, 越来越受到国家电力公司等单位的青睐。电力公司通过智能电表非侵入式负荷辨识技术收集用户侧用电比例构成<sup>[1]</sup>, 了解各个负荷

的用电特点及潜在用电规律<sup>[2]</sup>, 可以充分挖掘用户的互动潜力, 进而为实现电网透明化<sup>[3]</sup>奠定基础。

智能电表非侵入式负荷感知系统主要由数据采集、特征选取和负荷辨识 3 个模块组成<sup>[4]</sup>。其中, 特征选取和负荷辨识是负荷识别技术的关

基金项目: 南方电网数字电网研究院有限公司创新项目 (670000KK52200001)

作者简介: 石少青 (1965—), 男, 本科, 高级工程师, Email: maxwell201904@163.com

通讯作者: 连新凯 (1986—), 男, 本科, 高级工程师, Email: 329101854@qq.com

键。负荷特征是反映负荷运行状态的关键信息,通常可将其分为电量特征和非电量特征。常见的电量特征包括:有功功率<sup>[5]</sup>、无功功率、电流谐波<sup>[6]</sup>和 $V-I$ 轨迹<sup>[7]</sup>等稳态特征<sup>[8]</sup>,以及尖峰特征、电压噪声等暂态特征<sup>[9]</sup>;而非电量特征则包括时间特征<sup>[10]</sup>、天气、温度等表达用户用电行为的外界特征<sup>[11]</sup>。基于这些特征,文献[12]中首先使用聚类算法将负荷按有功功率和无功功率进行分类实现粗辨识,再将二维 $V-I$ 轨迹图作为卷积神经网络模型的输入,完成精细化识别。在文献[13]中提出了一种融合排列熵算法和Yamamoto算法的暂态特征采样方法,并结合特征权重,采用基于模糊 $c$ 均值聚类实现非侵入式负荷辨识。然而仅使用电量特征会因为负荷种类的增多而需要不断加入新的特征提升辨识能力,难免会造成计算量过大的问题。文献[14]通过考虑电器运行过程与时间的关联关系,结合负荷运行模式的时间尺度信息和信号幅值构建序列翻译模型,将待分解的能量翻译为状态码,实现负荷能耗分解。

考虑到不同的负荷特征获取方式,负荷辨识方法通常可分为基于负荷事件型和基于负荷运行状态型<sup>[15]</sup>两种。基于负荷事件的辨识方法提取随负荷状态变化产生的负荷特征变化,并将其作为辨识依据,而基于负荷运行状态的辨识方法将负荷工作时间内的特征状态作为辨识依据。文献[16]在负荷事件的基础上提取负荷特征,然后通过融合Fisher得分和主成分分析算法的Fisher主元分析法剔除掉区分性较差的无效特征,同时降低有效特征之间的关联性,然后采用遗传算法(genetic algorithm, GA)对模型参数进行寻优,从而完成负荷识别。而文献[17]首先采用基于B样条曲线拟合算法提取负荷运行状态的互补特征,然后用支持向量机(support vector machine, SVM)算法构建分类器对负荷进行辨识。一般来说,基于负荷事件的辨识方法无需了解负荷整个运行状态的情况,且负荷特征数据易于获取、方便计算和存储,具有准确率较高的特点。

基于此,本文通过负荷事件检测提取负荷事件的有功-无功功率、时间等特征,并对时间特征进行细化建模。考虑到用户用电行为的规律性,采用Mean-shift聚类对具有相同时间特征的负荷进行聚类,进而引入孪生神经网络进行负荷细分,最终实现负荷辨识。最后,通过实际数据测试验证文中方法的有效性。

## 1 负荷特征表达

负荷特征是用于区分不同类型负荷的内在特性,合理的选取负荷特征对后续的负荷辨识具有重要意义。本文选取的电量特征包括有功功率和无功功率特征;非电量特征则包括负荷运行时长特征、运行时刻等特征。

### 1.1 电量特征

#### 1.1.1 有功功率特征

有功功率 $P$ 作为一种易于获取的稳态特征,能直观地反映负荷的能耗且也是负荷投切时最为明显的特征。在现有智能电表,负荷辨识模块可直接与计量芯通信读取实时的有功功率值。

#### 1.1.2 无功功率特征

无功功率 $Q$ 与有功功率 $P$ 相对应,也是一种表现负荷内在性质的稳态特征。同样地,无功功率可直接从计量芯实时获取。

通常,负荷可根据无功和有功功率将负荷划分为阻性、容性以及感性等类别。

### 1.2 非电量特征

非电量特征,例如时间、温度、节假日等,是反映用户用电行为的特征,可作为负荷不易区分时潜在的辅助特征。时间特征应用较为广泛<sup>[18-19]</sup>,主要包括负荷运行时长和启动时刻。

#### 1.2.1 负荷运行时长特征

负荷运行时长是对负荷单次运行时间长短的描述,不同的负荷通常运行时间长短不一。图1为某负荷运行时长特征示意图,图中 $t_1$ 为负荷开启事件发生时刻点, $t_2$ 为该负荷关闭事件发生时刻点,两时刻点的差值即为负荷运行时长 $L$ 的值。

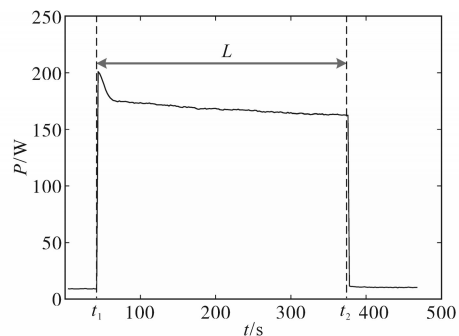


图1 负荷运行时长特征示意图

Fig.1 Feature of load running time

对于常见的家用电器,大致可将其分为两类,其中建模方式以高斯和均匀分布为主,如表1所示。因此对负荷的运行时长 $L$ ,其概率密度分布可描述为

$$L = \begin{cases} N(\mu, \sigma^2) \\ l \end{cases} \quad (1)$$

式中: $N(\mu, \sigma^2)$ 为高斯分布; $l$ 为常数,代表均匀分布。

表1 负荷运行时间分布

Tab.1 Distribution of time length

类别	特点	建模方式	典型负荷
类别1	运行时长相对较为固定的负荷	高斯模型	电水壶、电饭煲等
类别2	运行时长不确定的负荷,由几十min到几h不等,具有很强的随机性	均匀分布	电风扇、电视机等

一般而言,负荷运行时长的统计依赖于负荷自身的特性以及用户的用电行为。本文在负荷事件检测的基础上,对某家庭用户各负荷运行时长进行了统计,如图2所示,其中各图形上方的数字表示各负荷的平均运行时长。

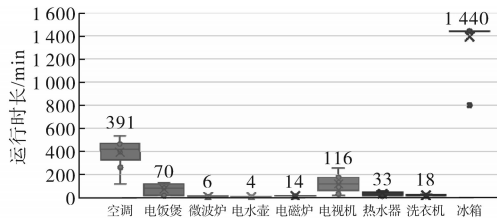


图2 负荷运行时长特征统计结果

Fig.2 The length of statistical time of load running

从图2中可以看出,各负荷运行时长之间具有一定的区分性,例如空调通常是长时间运行,电视机平均运行时长在2h左右,而微波炉、电水壶、电磁炉以及洗衣机的运行时长较短,仅为几min或十几min,且分布相对较为集中。

### 1.2.2 负荷运行时刻特征

负荷运行时刻特征通常与用户作息、行为习惯有较大关联。图3为统计某家庭用户各负荷运行时刻特征统计热力图,图右侧刻度条颜色表示对应的运行次数。图中色块的位置反映各负荷处于运行状态的时间段,颜色越深代表负荷在该时间段内使用频率越高。例如,冰箱几乎处于全天运行状态,而热水器通常在晚上19:00—20:00之间运行。此外,从图中也可以看出空调的运行时段也较为固定,均为夜里22:00以后开启到次日早晨关闭,这与用户睡眠时间有关,而电视机则通常在中午和晚上运行。这些也从侧面反映了用户对负荷的使用习惯。因此,可将一天24h分为若干个时间段,并用二值函数表示在某时间段( $t_1, t_2$ )内是否有负荷事件发生,即

$$l_i = \begin{cases} 1 & t \in (t_1, t_2) \\ 0 & \text{其他} \end{cases} \quad (2)$$

$l_i=1$ 表示发生负荷事件; $l_i=0$ 表示未发生负荷事件。

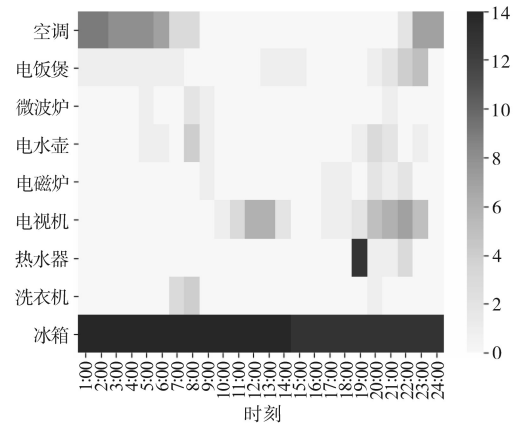


图3 负荷运行时刻特征统计结果

Fig.3 Statistical running time of load

### 1.2.3 间歇性变化特征

除了上述一些特征之外,还存在一些工作波形特征。图4所示为洗衣机工作时所展现的波形曲线特性。不难看出,该特征具有较好的区分性,因此在非电量特征中可直接进行负荷判别,即

$$T = \begin{cases} 1 & \text{间歇性变化} \\ 0 & \text{非间歇性变化} \end{cases} \quad (3)$$

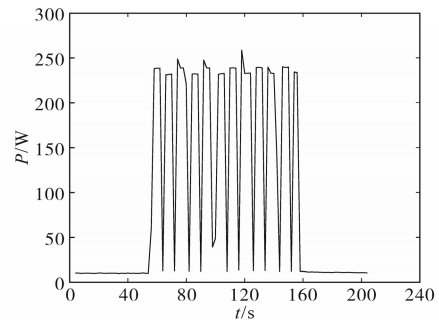


图4 负荷运行间歇性变化特征示意图

Fig.4 Variation feature of load during running

## 2 负荷时间特征统计及细分

考虑到用户用电行为通常具有统计规律,因此时间特征是除电量特征以外能够有效地提高负荷辨识准确率的特征。本文首先对负荷时间特征进行统计以及细化建模。

### 2.1 负荷运行时长特征分段

采用负荷运行时长特征进行分类,最为直接的方式是按运行时长的长短进行划分。为了确定具体的运行时长的细分类型数和分类阈值,本文对所有负荷的运行时长采用核密度估计方法,得到负荷时长特征概率密度分布函数 $f(x)$ 。对 $f(x)$ 进行积分,并令积分结果满足:

$$\int f(x) dx = \varepsilon \quad (4)$$

根据  $\varepsilon$  值的不同,可将负荷按运行时长分为若干个类型,使得每个类型包含的负荷种类数大致相等,同时得到时长阈值  $x$  的值。

### 2.2 负荷运行时刻特征分段

考虑到负荷在一天中的运行时刻具有不确定性,可能存在多负荷同时运行的高峰期(晚上下班后),也存在几乎无负荷运行的低谷期(在白天的工作时间),因此负荷运行时间特征的分段往往不是均匀等距的,需要根据具体的场景进行划分。

对于不同用户的不同用电行为规律,为进一步确定具体划分时间段的时间点,本文采用一种最优化的方法,规定所有负荷所占的时间段数总和最少即为最优的时间区间划分结果,这能最大程度地将不同运行时刻的负荷区分开来。

令  $N_{T_i}$  表示第  $i$  个负荷占用的时间段数,  $i \in 1 \sim s$ ,  $s$  为总的负荷类型数,则目标函数的表达式如下式所示:

$$g(N_{T_i}) = \min \sum_{i=1}^s N_{T_i} \quad (5)$$

对于具有  $s$  类负荷的环境而言,可以用稀疏二值矩阵  $LT_{s \times 24}$  表示各负荷的运行时段(下标 24 代表一天 24 h),矩阵中 1 代表存在负荷运行,0 代表不存在负荷运行,结合式(2)可将某一天的各负荷运行时刻分布情况表示为

$$LT = \begin{bmatrix} lt_{11} & \cdots & lt_{124} \\ \vdots & \ddots & \vdots \\ lt_{s1} & \cdots & lt_{s24} \end{bmatrix} \quad (6)$$

其中,矩阵的每一行代表各负荷存在运行状态的时间段,每一列代表该小时内运行的负荷类型。若将时刻  $t_{m1}, t_{m2}, \dots, t_{mt}$  ( $1 < m1 < m2 < \dots < mt < 24$ ) 作为分割 24 h 的时刻点,将负荷运行时刻分为若干个部分,则对于负荷  $i$ ,其运行的时间段数  $N_{T_i}$  为

$$N_{T_i} = \text{sgn}(\sum_{j=1}^{m1-1} lt_j) + \text{sgn}(\sum_{j=m1}^{m2-1} lt_j) + \dots + \text{sgn}(\sum_{j=mt}^{24} lt_j) \quad (7)$$

式中: $j$  为时刻点,  $j \in [1, 24]$ 。

定义符号函数  $\text{sgn}(\cdot)$  为

$$\text{sgn}(x) = \begin{cases} 1 & x > 0 \\ 0 & \text{其他} \end{cases} \quad (8)$$

将式(7)、式(8)计算结果代入式(5)即可求得  $m1, \dots, mt$  的值。

## 3 Mean-shift 负荷事件聚类和孪生网络判断

在智能电表非侵入式负荷辨识研究中,时间特征和电量特征具有一定的差异性。本文对时间特征采用 Mean-shift 聚类,得到用户在相同时间特征下所使用的负荷,同时利用孪生神经网络,通过与数据库负荷匹配辨识出所使用的负荷类型。

### 3.1 Mean-shift 聚类

Mean-shift 算法作为一种非参数化聚类算法,其基本思想是假设不同簇类的数据集符合不同的概率密度分布,通过找到任一样本点密度增大的最快方向,那么样本密度高的区域对应于该分布的最大值,这些样本点最终会在局部密度最大值处收敛,且收敛到同一局部最大值的样本点就被认为是同一簇类。相比于其他聚类算法如 K-means 算法,Mean-shift 算法无需预先设置聚类中心数,其可以借助数据分布概率密度自适应地选取聚类中心数<sup>[20]</sup>,这对一天内使用的负荷类别的快速确定具有重要意义。

假定给定  $d$  维空间的  $n$  个样本点  $x_i$  ( $i=1, \dots, n$ ),空间中任一点  $x$  的 Mean-shift 漂移向量的基本形式为

$$M_h = \frac{1}{K} \sum_{x_i \in S_h} (x_i - x) \quad (9)$$

其中

$$S_h(x) = \{y: \|y - x\|^2 < h^2\}$$

式中: $y$  为在空间尺度  $h$  内的变量; $K$  为  $n$  个样本点中分布在区域  $S_h$  内的个数; $S_h$  为以  $x$  为中心、 $h$  为半径的高维球域区域。

在聚类的过程中,按照下式不断地计算漂移向量进行迭代,最终获得聚类中心。

$$\hat{x} = x + M_h(x) = \frac{1}{h} \sum_{x_i \in S_h} x_i \quad (10)$$

为了优化聚类效果,引入核函数  $K(x_i - x)$ <sup>[21]</sup>,则均值偏移向量更新为

$$M_h(x) = \frac{\sum_{x_i \in S_h} [K(x_i - x)\omega(x_i)x_i]}{\sum_{x_i \in S_h} [K(x_i - x)\omega(x_i)]} - x \quad (11)$$

式中: $\omega(x_i)$  为样本  $x_i$  的权重,由其概率分布决定。

### 3.2 孪生网络判断负荷类型

在聚类之后,本文采用如图 5 所示的孪生网络的方法对负荷进行细分类。

孪生网络主要用来衡量两个输入的相似程度,其有两个输入( $X_1$ 和 $X_2$ ),将两个输入分别输入两个神经网络(Network1和Network2),通常两个神经网络的结构和参数一般是相同的,它们分别将输入映射到新的空间形成新的特征,并最终通过计算距离来评价两个输入的相似度。

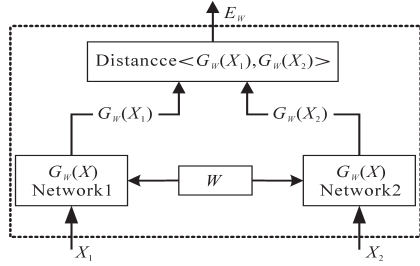


图5 孪生网络结构图

Fig.5 Siamese network framework

图5中, $X_1, X_2$ 为负荷特征和数据库中负荷特征; $G_w$ 表示神经网络模型,下角 $W$ 表示权重, $G_w$ 的作用就是将电量和非电量特征数据 $X$ 转换为一组特征向量; $E_w$ 则用于衡量特征向量之间的距离,距离越小,则代表越相似。

本文采用弹性BP神经网络<sup>[22]</sup>(resilient back propagation, RPROP)作为孪生神经网络中的 $G_w$ 。假设分类对象为 $\Psi_i(i=1, 2, \dots, p)$ ,每个对象 $\Psi_i$ 对应有 $q$ 个样本 $\Psi_{ij}(j=1, 2, \dots, q)$ ,其中样本特征维数为 $r$ ,即

$$\mathbf{x}_{ij} = [x_{ij}^{(1)} x_{ij}^{(2)} \dots x_{ij}^{(r)}]^T \quad (12)$$

则特征集的个数为 $p \times q$ ,此时将这些样本输入神经网络,得到 $p \times q$ 个输出矢量:

$$\mathbf{y}_{ij} = [y_{ij}^{(1)} y_{ij}^{(2)} \dots y_{ij}^{(u)}]^T \quad (13)$$

式中: $u$ 为输出层数据维度。

则生成的特征矩阵为

$$\mathbf{Y} = [\mathbf{y}_{11} \mathbf{y}_{12} \dots \mathbf{y}_{1q} \mathbf{y}_{21} \mathbf{y}_{22} \dots \mathbf{y}_{2q} \dots \mathbf{y}_{p1} \mathbf{y}_{p2} \dots \mathbf{y}_{pq}] \quad (14)$$

为了降低计算量,对矩阵 $\mathbf{Y}$ 进行主成分分析。令 $\mathbf{Y}$ 的协方差矩阵的最大特征值对应的特征矢量为 $\mathbf{t}$ ,则在第一主元方向的投影所得数据为

$$\begin{aligned} \boldsymbol{\alpha} &= \mathbf{t}^T \mathbf{Y} \\ &= [\boldsymbol{\alpha}_{11} \boldsymbol{\alpha}_{12} \dots \boldsymbol{\alpha}_{1q} \boldsymbol{\alpha}_{21} \boldsymbol{\alpha}_{22} \dots \boldsymbol{\alpha}_{2q} \dots \boldsymbol{\alpha}_{p1} \boldsymbol{\alpha}_{p2} \dots \boldsymbol{\alpha}_{pq}] \end{aligned} \quad (15)$$

即第 $i$ 类对象的第 $j$ 个样本对应于第一主元方向上的投影数据 $\boldsymbol{\alpha}_{ij}$ 。假定当第 $n$ 个主元成分上投影的方差和占总方差超过某个阈值(文中设置为0.90),则确定了对应的 $n$ 个特征向量,从而生成特征集。

为了判别神经网络输入属于具体某个负荷类别,将数据库中负荷类别特征作为孪生神经网络

输入,其中孪生神经网络输出的距离度量采用K-means分类方法。

### 3.3 实现流程

图6给出了本文方法的具体流程框架。在该框架中,首先对负荷事件通过时间特征进行聚类,获得相应的负荷类别,然后通过孪生神经网络,将与数据库中与之具有相同时间特征的负荷进行匹配,最终得到负荷辨识结果。

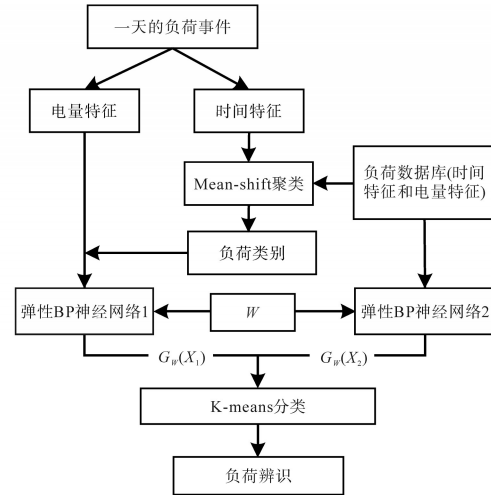


图6 本文方法模型结构图

Fig.6 Frame chart of proposed method

## 4 实验结果与分析

为了验证文中方法的有效性,本文以某一家庭用户作为具体的研究对象,连续对该家庭用户监测30 d的负荷数据进行离线分析,并从中分析提取出负荷运行的特征数据,构建负荷特征数据库。用户使用的负荷包括空调、电饭煲、微波炉、电水壶、电磁炉、电视机、热水器、洗衣机和电冰箱。为此,本文将该9种负荷作为研究对象进行辨识,表2给出了该用户负荷有功-无功功率特征和间歇性特性数据库,运行时长和投切时刻统计图见图2、图3。文中设置Mean-shift算法中时长尺度为10 min,投切时刻尺度为相差30 min。

### 4.1 参数设置

#### 4.1.1 负荷运行时长分段阈值

按照1.2.1中的方法,为确定具体的运行时长分类阈值,图7给出了负荷时长特征概率密度分布结果。为了获得时长分类的界限,本文对概率密度函数进行积分,取积分结果中值处作为负荷运行时长分界线。为了方便表示,令 $t_0$ 表示时长分类阈值点,设置式(4)中 $\varepsilon=1/2$ ,计算得到分类阈值

表2 负荷特征数据库

Tab.2 Database of load signature

负荷	有功功率/W	无功功率/var	平均有功功率/W	平均无功功率/var	时长中值/min	间歇性跳变特征 $T$
空调	451.8~635.9	-104.8~85.7	487.7	-33.7	27	0
电饭煲	511.0~639.0	-29.0~36.0	592.4	-12.0	34	0
微波炉	852.9~930.8	-351.6~-203	884.9	-253.0	3	0
(2档)	1 425.0~1527.2	-433~-239	1 469.0	-348.1	3	0
电水壶	1 200.0~1 275.0	-19.1~28.0	1 260.0	-15.6	7	0
电磁炉	1 452.0~1 506.5	-72.0~48.2	1 489.0	-44.1	23	0
电视机	101.0~131.7	-33.8~34.6	117.0	18.8	129	0
热水器	1 441.6~1 503.0	-32.9~26.8	1 477.0	-14.5	18	0
洗衣机	193.8~272.1	-89.0~95.0	233.7	-40.6	63	1
冰箱	60~105.3	-114.1~-42.8	81.4	-74.3	37	0

$t_0 \approx 40$  min, 由此可将  $t_0$  作为负荷事件分类的运行时长阈值, 即短时长为  $0 < L \leq 40$ , 长时长为  $L > 40$ 。

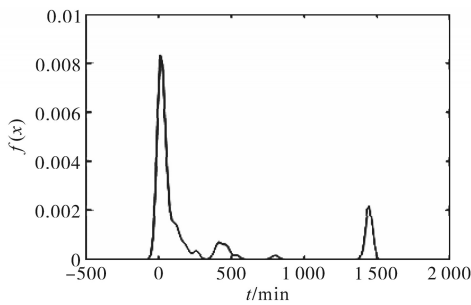


图7 负荷运行时长特征统计示意图

Fig.7 Statistical length of running time

#### 4.1.2 负荷运行时刻特征分段阈值

由图3可以看出, 用户用电行为较为活跃的时间段为早晨和夜晚, 因此可依据用户的用电规律, 将24 h分为3段, 即式(7)中  $N_{T_i} \in [1, 3]$ 。由于在该实验对象家庭中负荷类型数  $s=9$ , 为使每个时段时长分布较为均匀, 令  $m_1 \in [1, 12], m_2 \in [14, 20]$ , 此时可通过计算式(7)得到满足式(5)的  $m_1$  和  $m_2$ , 遍历后得到  $m_1$  取值为10,  $m_2$  取值为18, 此时  $\sum_{i=1}^s N_{T_i}$  的值最小为19。因此, 可将该家庭用户负荷事件按负荷时间特征进行分段处理, 分别为  $t_1 \in 0:00-10:00, t_2 \in 10:00-18:00, t_3 \in 18:00-24:00$ 。从而将该家庭用户负荷事件按负荷运行时刻进行分段处理。

#### 4.2 测试案例分析

在统计获得负荷特征数据库的基础上, 对用户某一天的用电数据作为测试数据进行实验。

首先, 对该用户的日用电功率曲线进行负荷事件检测, 并提取负荷事件的有功-无功功率特征和时间特征信息, 测试数据如图8所示, 负荷事件特征信息如表3所示。由时间特征分类可知, 该日负荷事件分为5类, 与之匹配

的在数据库中得到的负荷类别如表4所示。然后, 与数据库中的时间特征采用 Mean-shift 聚类后, 得到聚类匹配结果如表5所示。显然, 从数据库中聚类得到的结果可以发现, 时间特征上相似的负荷设备可能存在多个, 因此需要进一步细分。

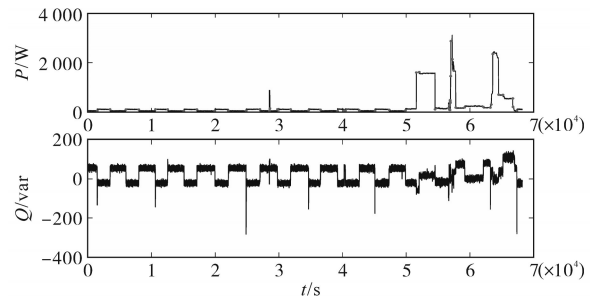


图8 测试数据日负荷事件检测结果

Fig.8 Detect result of test data in one day of home

为了最终判断负荷类型, 将电量特征和时长特征作为孪生网络中弹性BP神经网络1的输入, 同时, 将数据库中训练的负荷类别的特征信息输入到BP神经网络2, 通过生成的特征进行K-means聚类, 最终得到的匹配结果如表6所示, 与实际负荷投切运行结果一致。

为了进一步验证本文方法的有效性, 本文与文献[22]中的神经网络负荷辨识方法进行比较, 其中输入特征为有功、无功, 辨识结果如表7所示。从表中可以发现, 在没有时间特征细分情况下, 单一地采用神经网络进行负荷辨识, 因特征存在混叠导致其错误率较高; 而本文方法首先通过对时间特征的细分, 并采用 Mean-shift 算法与数据库中的负荷进行聚类, 选择具有相似的时间特征, 然后通过孪生神经网络进行细分, 实现了较高的辨识准确率。在该用户场景下, 其平均辨识准确率为0.9763。

表3 负荷事件特征信息

负荷事件标号	投入时刻 /min	负荷时长 /min	有功功率 /W	无功功率 /var
1	24	34	83.92	-72.61
2	100	34	84.37	-74.92
3	177	35	82.67	-85.88
4	258	34	86.1	-81.08
5	339	36	81.23	-74.95
6	423	36	74.82	-82.16
7	508	36	76.48	-67.53
8	593	38	80.13	-72.1
9	937	39	85.96	-82.43
10	1 023	38	85.73	-85.85
11	1 140	51	1 430.9	4.56
12	1 235	14	1 436	5.32
13	1 235	6	1 205	10.73
14	1 277	37	74.9	-63.96
15	1 352	61	468.53	28.81
16	1 356	16	1 460.7	6.7

表4 负荷时间特征匹配信息

类别	时间特征	负荷类别
K1	$t \in [0:00-10:00] \& L \in [0, 40]$	冰箱, 空调, 微波炉, 电水壶, 电饭煲, 电磁炉
K2	$t \in [10:00-18:00] \& L \in [0, 40]$	电冰箱, 电饭煲, 电磁炉
K3	$t \in [18:00-24:00] \& L \in [0, 40]$	空调, 电视机, 微波炉, 电饭煲, 电水壶, 热水器, 电磁炉
K4	$t \in [0:00-10:00] \& L > 40$	空调, 洗衣机
K5	$t \in [10:00-18:00] \& L > 40$	电视机
K6	$t \in [18:00-24:00] \& L > 40$	电视机, 热水器, 洗衣机

表5 时间特征 Mean-shift 聚类结果

负荷事件标号	投入时刻 /min	负荷时长 /min	负荷类别
1	24	34	冰箱, 电饭煲, 空调
2	100	34	冰箱, 电饭煲, 空调
3	177	35	冰箱, 电饭煲, 空调
4	258	34	冰箱, 电饭煲, 空调
5	339	36	冰箱, 电饭煲, 空调
6	423	36	冰箱, 电饭煲, 空调
7	508	36	冰箱, 电饭煲, 空调
8	593	38	冰箱, 电饭煲, 空调
9	937	39	冰箱, 电饭煲, 空调
10	1 023	38	冰箱, 电饭煲, 空调
11	1 140	51	热水器, 空调
12	1 235	14	空调, 热水器, 电磁炉
13	1 235	6	电水壶, 微波炉
14	1 277	37	冰箱, 电饭煲, 空调
15	1 352	61	空调, 洗衣机
16	1 356	16	空调, 热水器, 电磁炉

表6 孪生神经网络匹配聚类结果

负荷事件标号	负荷类别	负荷事件标号	负荷类别
1	冰箱	9	冰箱
2	冰箱	10	冰箱
3	冰箱	11	热水器
4	冰箱	12	热水器
5	冰箱	13	电水壶
6	冰箱	14	冰箱
7	冰箱	15	空调
8	冰箱	16	热水器

表7 对比分析结果

负荷	测试样本数	辨识准确率	
		RPROP神经网络	本文方法
空调	39	0.717 9	0.923 1
电饭煲	32	0.812 5	0.937 5
微波炉	5	1.000 0	1.000 0
电水壶	24	1.000 0	1.000 0
电磁炉	18	0.555 6	0.944 4
电视机	38	0.789 5	1.000 0
热水器	56	0.785 7	0.946 4
洗衣机	34	1.000 0	1.000 0
冰箱	134	0.820 8	1.000 0

## 5 结论

针对实际家庭用电场景中用户使用的负荷类型辨识问题,文中提出了基于Mean-shift聚类和孪生网络的负荷辨识方法,该方法按家庭负荷时间特征分布情况统计结果,将负荷事件按运行时长和运行时刻特征通过Mean-shift聚类方法,获得相似时间特征的负荷,然后采用孪生网络进行负荷匹配辨识,提高辨识准确率。实验结果证明,该方法可有效弥补因有功-无功特征混叠而引起的负荷辨识准确率低的问题。在下一步工作中,我们也将通过采集其他不同家庭用户负荷特征,将文中算法进行应用推广。

### 参考文献

- [1] 徐涛,黄莉,李敏蕾,等. 基于多维细粒度行为数据的居民用户画像方法研究[J]. 电力需求侧管理, 2019, 21(3): 47-52, 58. Xu Tao, Huang Li, Li Minlei, et al. Research on portrait method of residential users based on multi-dimensional fine-grained behavior data[J]. Power Demand Side Management, 2019, 21(3): 47-52, 58.
- [2] 李俊雄,黎灿兵,曹一家,等. 面向智能电网的互动式节能调度初探[J]. 电力系统自动化, 2013, 37(8): 20-25. Li Junxiong, Li Canbing, Cao Yijia, et al. Preliminary investigation on interactive energy-saving dispatch oriented to smart grid

- [J]. Automation of Electric Power Systems, 2013, 37(8): 20-25.
- [3] 李钦豪,张勇军,陈佳琦,等. 泛在电力物联网发展形态与挑战[J]. 电力系统自动化, 2020, 44(1): 13-22.  
Li Qin hao, Zhang Yongjun, Chen Jiaqi, *et al.* Development patterns and challenges of ubiquitous power internet of things[J]. Automation of Electric Power Systems, 2020, 44(1): 13-22.
- [4] Abubakar I, Khalid S, Mustafa M, *et al.* An overview of non-intrusive load monitoring methodologies[C]//2015 IEEE Conference on Energy Conversion, CENCON, 2015: 54-59.
- [5] Lin Y, Hu Y. Electrical energy management based on a hybrid artificial neural network-particle swarm optimization-integrated two-stage non-intrusive load monitoring process in smart homes[J]. Processes, 2018, 6(12): 236.
- [6] 吕志宁,赵少东,饶竹一. 非侵入负荷辨识的谐波特征量提取改进方法研究[J]. 电子测量技术, 2019, 42(7): 29-34.  
Lü Zhining, Zhao Shaodong, Rao Zhu yi. Improved method research on extracting load harmonic feature for non-intrusive load monitoring[J]. Electronic Measurement Technology, 2019, 42(7): 29-34.
- [7] 杨云瑞,黄宇魁,陈文浩. 基于V-I轨迹特征的非侵入式负荷监测算法[J]. 计算机与数字工程, 2020, 48(10): 2353-2358.  
Yang Yunrui, Huang Yukui, Chen Wenhao. Non-intrusive load monitoring algorithm based on V-I trajectory feature[J]. Computer and Digital Engineering, 2020, 48(10): 2353-2358.
- [8] 师洪涛,蒋中南,张巍巍,等. 基于Popov理论的微电网逆变器控制策略研究[J]. 电气传动, 2021, 51(11): 48-54.  
Shi Hongtao, Jiang Zhongnan, Zhang Weiwei, *et al.* Study on power quality control strategy of microgrid inverter based on Popov theory[J]. Electric Drive, 2021, 51(11): 48-54.
- [9] Liang J, Ng S, Kendall G, *et al.* Load signature study—Part I: Basic concept, structure, and methodology[J]. IEEE Transactions on Power Delivery, 2010, 25(2): 551-560.
- [10] 李如意,张鹏,刘永光,等. 基于随机森林的非侵入式家庭负荷辨识方法[J]. 电测与仪表, 2021, 58(4): 9-16.  
Li Ruyi, Zhang Peng, Liu Yongguang, *et al.* Non-intrusive household load identification method based on random forest[J]. Electrical Measurement & Instrumentation, 2021, 58(4): 9-16.
- [11] 李博,门德月,严亚勤,等. 基于数值天气预报的母线负荷预测[J]. 电力系统自动化, 2015, 39(1): 137-140.  
Li Bo, Men Deyue, Yan Yaqin, *et al.* Bus load forecasting based on numerical weather prediction[J]. Automation of Electric Power Systems, 2015, 39(1): 137-140.
- [12] 汪颖,杨维,肖先勇,等. 基于U-I轨迹曲线精细化识别的非侵入式居民负荷监测方法[J]. 电网技术, 2021, 45(10): 4104-4113.  
Wang Ying, Yang Wei, Xiao Xianyong, *et al.* Non-intrusive residential load monitoring method based on refined identification of U-I trajectory curve[J]. Power System Technology, 2021, 45(10): 4104-4113.
- [13] 高浩瀚. 非侵入式负荷辨识的特征分析研究[D]. 济南: 山东大学, 2019.
- Gao Haohan. Signature analysis of non-intrusive load identification[D]. Jinan: Shandong University, 2019.
- [14] 任文龙,许刚. 基于深度序列翻译模型的非侵入式负荷分解方法[J]. 电网技术, 2020, 44(1): 27-37.  
Ren Wenlong, Xu Gang. Non-intrusive load decomposition method based on deep sequence translation model[J]. Power System Technology, 2020, 44(1): 27-37.
- [15] 刘卫涛,刘博,栾文鹏,等. 考虑稀疏性的多特征电力负荷功率实时分解[J]. 电力系统及其自动化学报, 2018, 30(2): 79-84.  
Liu Weitao, Liu Bo, Luan Wenpeng, *et al.* Multi-characteristic power load power real-time decomposition considering sparsity[J]. Journal of Electric Power System and Automation, 2018, 30(2): 79-84.
- [16] 全瑞宁,李鹏,郎恂,等. 基于Fisher主元分析和核极限学习机的非侵入式电力负荷辨识模型[J]. 电力建设, 2021, 42(2): 85-92.  
Tong Ruining, Li Peng, Lang Xun, *et al.* Non-intrusive power load identification model based on FPCA and KELM[J]. Electric Power Construction, 2021, 42(2): 85-92.
- [17] Su D, Shi Q, Xu H, *et al.* Nonintrusive load monitoring based on complementary features of spurious emissions[J]. Electronics, 2019, 8(9): 1002.
- [18] Lin Y H, Tsai M S. Development of an improved time-frequency analysis-based nonintrusive load monitor for load demand identification[J]. Instrumentation & Measurement, 2014, 63(6): 1470-1483.
- [19] Liang J, Ng S, Kendall G, *et al.* Load signature study—Part II: Disaggregation framework, simulation, and applications[J]. IEEE Transactions on Power Delivery, 2010, 25(2): 561-569.
- [20] 文亚凤,崔亮节,孙毅,等. 考虑状态概率因子和状态修正的非侵入式负荷分解方法[J]. 电网技术, 2019, 43(11): 349-355.  
Wen Yafeng, Cui Liangjie, Sun Yi, *et al.* Non-intrusive load decomposition method considering state probability factors and state corrections[J]. Power System Technology, 2019, 43(11): 349-355.
- [21] 陈建华,王磊,巨云涛. 基于Mean-shift聚类的大规模新能源对电网线损影响评估方法[J]. 供用电, 2021, 38(2): 59-64, 76.  
Chen Jianhua, Wang Lei, Ju Yuntao. Evaluation method of large scale new energy impact on power grid line loss based on Mean-shift clustering[J]. Distribution & Utilization, 2021, 38(2): 59-64, 76.
- [22] 李如意,王晓换,胡美璇,等. RPROP神经网络在非侵入式负荷分解中的应用[J]. 电力系统保护与控制, 2016, 44(7): 55-61.  
Li Ruyi, Wang Xiaohuan, Hu Meixuan, *et al.* Application of RPROP neural network in nonintrusive load decomposition[J]. Power System Protection and Control, 2016, 44(7): 55-61.

收稿日期:2021-05-31

修改稿日期:2021-06-22