

• 电气工程 •

DOI:10.15961/j.jsuese.202200034



本刊网刊

基于CFSFDP图拉普拉斯算法的非侵入式负荷监测方法

林平川¹, 路磊¹, 谷超¹, 冯俊国¹, 张仕文¹, 杨顺尧¹, 于丹¹, 郑迪文^{2*}, 汪颖²

(1.国网河北省电力有限公司 石家庄供电分公司, 河北 石家庄 050051; 2.四川大学 电气工程学院, 四川 成都 610065)

摘要: 非侵入式负荷监测(NILM)是中国未来电网建设的重要发展方向之一。为克服传统非侵入式负荷监测方法的计算数据量大、辨识准确率较低等问题,提出了一种基于快速密度峰值搜索算法(CFSFDP)图拉普拉斯算法的非侵入式负荷监测方法。首先,该方法利用输入的设备有功功率数据采取快速密度峰值搜索聚类算法构建家用电器的功率阈值向量和先验图结构;然后,结合图信号的平滑度特征和总功率信号构建图拉普拉斯二次型最优函数,利用Tikhonov正则化方法以迭代的方式求得最优解,从而实现用电负荷图信号的重构;最后,根据功率阈值向量将图信号转换为功率信号,即可实现用户的非侵入式负荷监测。对某一家庭2 d的实测用电数据进行仿真分析,包括2 d内的负荷监测结果和采样频率对算法性能的影响,结果如下:1)该方法能够识别出第1天内工作的所有设备,各用电设备消耗用电量比例与实际耗电量比例接近。2)该方法对第2天的负荷识别准确率达到90.1%,优于4种对比算法。单个用电设备的分解精度达到91%以上,绝大多数设备的用电量误差都低于对比算法。3)当数据采样间隔增大为2 min,所提算法的准确率、辨识精度和单设备分解精度都有所降低,但数值上优于对比算法,并且有更优的时间复杂度。研究结果验证了所提非侵入式负荷监测方法的有效性及其优越性,对于解决实际低频NILM问题有很大的优势。

关键词: 非侵入式负荷监测; CFSFDP聚类算法; 图拉普拉斯二次型; Tikhonov正则化

中图分类号: TM714

文献标志码: A

文章编号: 2096-3246(2023)04-0216-08

Non-intrusive Load Monitoring Method Based on CFSFDP Graph Laplace Algorithm

LIN Pingchuan¹, LU Lei¹, GU Chao¹, FENG Junguo¹, ZHANG Shiwen¹,
YANG Shunyao¹, YU Dan¹, ZHENG Diwen^{2*}, WANG Ying²

(1.Shijiazhuang Power Supply Branch, State Grid Hebei Electric Power Co., Ltd., Shijiazhuang 050051, China;
2.School of Electrical Eng., Sichuan Univ., Chengdu 610065, China)

Abstract: Non-intrusive load monitoring (NILM) is one of the important development directions of power grid construction in China in the future. In order to overcome the problems of large amount of calculation data and low identification accuracy of traditional NILM methods, a non-intrusive load monitoring method based on CFSFDP (clustering by fast search and find of density peaks) graph Laplace algorithm was proposed in this paper. Firstly, the power threshold vector and the prior graph structure were constructed using the active power data adopting the CFSFDP algorithm. Then, the graph Laplacian quadratic optimal function was constructed by combining the total power signal and graph signal smoothness, and the optimal solution was obtained iteratively by Tikhonov regularization method, so as to realize the reconstruction of graph signal of appliance. Finally, the graph signals were converted into power signals according to the power threshold vector, which enabled non-intrusive load monitoring. The following results were obtained from the simulation analysis of two days of measured electricity consumption data of a real household, including the load monitoring results within two days and the impact of sampling frequency on the algorithm performance. 1) The pro-

收稿日期:2022-01-11

基金项目:国家自然科学基金青年基金项目(52107117)

作者简介:林平川(1994—),男,助理工程师。研究方向:智能用电。E-mail: linpc13931170195@163.com

*通信作者:郑迪文, E-mail: 864647962@qq.com

网络出版时间:2022-08-12 14:39:53

网络出版地址:https://kns.cnki.net/kcms/detail/51.1773.TB.20220812.0854.001.html

posed method can identify all the equipments running within the first day, and the calculated proportion of electricity consumed by each electric equipment is close to the actual situation. 2) The load identification accuracy of the proposed method for the next day is 90.1%, which is superior to four comparison methods. The decomposition accuracy of a single appliance is more than 91%, and the vast majority of devices have lower power consumption error than the comparison methods. 3) When the data sampling interval is increased to 2 min, although the precision, identification accuracy and single appliance decomposition accuracy of the proposed method are all reduced, its calculation results gains superior performance and time complexity than the comparison methods. The simulation results verify the effectiveness of the proposed non-intrusive load monitoring method and its superiority for solving practical low-frequency NILM problems.

Key words: NILM(non-intrusive load monitoring); CFSFDP clustering algorithm; graph Laplace quadratic; Tikhonov regularization

随着高效低碳经济和环境友好型经济的可持续发展,非侵入式电力负荷监测(non-intrusive load monitoring, NILM)技术逐渐成为国内外研究者的热点。麻省理工学院的Hart教授^[1]在1982年首次提出非侵入式负荷监测的概念,随后建立非侵入式负荷监测系统正式用于监测家用负荷的用电情况^[2]。该技术仅通过在用户入口处安装监测设备采集的总用电信息辨识用户家用设备的工作状态和能耗。

近年来,国内外研究者从不同角度、方向上提出了许多不同的非侵入式负荷分解方法。Held^[3]、Le^[4]、Ghosh^[5]等着重研究了单个设备的有功功率、无功功率、电流波形、多次谐波等特征为负荷特征库。Le等^[6]通过希尔伯特变换生成一种新的瞬态特征APF(振幅-相位-频率)并应用序列对序列长-短期记忆算法来识别设备。Chen等^[7]从监测的能量数据中构造出保留足够信息的时间和光谱负荷特征,并提出了双流卷积神经网络来从这两类负载特征中提取特征并进行分类。汪颖^[8]和杜刃刃^[9]等分别利用U-I(电压-电流)轨迹曲线特征构建卷积神经网络和小波变换提取的能量系数进行支持向量机分类来完成非侵入式负荷监测。以上研究都是利用高频采样获得了功率、电流、电压等高频数据,由此变换计算得到各种新颖实用的负荷特征库,利用不同的负荷特征提取方法和分解算法可以实现精度较高的负荷辨识。但是高频采样对采样设备要求较高,数据存储、处理和计算难度极大,不利于非侵入式技术的推广^[10]。目前,国内外大量使用的智能电表也因制造成本、存储成本、数据传输容量的限制,一般适用于间隔比较大的采样率^[11]。因此不少学者针对低频采样负荷分解问题进行研究。

汪繁荣等^[12]利用聚类数据建立特征集并提出果蝇优化广义回归神经网络模型进行负荷监测,但需要提前训练大量参数获得设备状态模型。Wittmann^[13]和Liu^[14]等分别提出了两种基于混合线性规划的负荷分解算法。前者加入状态开关和持续时间限制来有效建立模型,并且使用时间窗将序列分解为多个子序列来提高计算效率;后者采用2-SAX(符号集合

近似法)算法提取各设备的状态转移行为特征和运行概率特征,以克服传统线性规划方法的缺点。但是此类方法针对多状态设备有较大的局限性。Liu等^[15]提出了一种具有深度编码器-解码器架构的能量分解网络(EDNet)实现无延迟NILM,但在多个设备同时开启时算法的分解效果不理想。徐青山^[16]和He^[17]等均利用聚类算法构建用电设备特征集并结合遗传算法实现负荷监测,算法效率较高,但聚类过度容易降低算法准确率。

为克服高频数据计算量大和低频算法准确率不高等问题,本文提出了一种基于CFSFDP图拉普拉斯算法的非侵入式负荷监测方法。该方法首先利用CFSFDP聚类算法构建家用电器的先验图结构,根据图平滑度建立图拉普拉斯二次型最优函数。然后通过Tikhonov正则化方法求解图拉普拉斯二次型优化问题的最优解,将求解得到的图信号转换为功率信号,即可实现用户的非侵入式负荷监测。该方法仅利用电表采集的低频用电功率数据,降低了对硬件设备的要求和数据处理的时间需求。同时,CFSFDP聚类算法可以自动检测簇的数量而不会聚类过度。图拉普拉斯矩阵存储了用电信号的拓扑结构和图总变化量等信息,独特的负荷特征提高了负荷辨识度,并且不需要大量的训练。经过实验验证表明,本文算法能够准确监测家用电器的用电情况,相比于已有算法有更高的准确率和辨识精度。

1 基础方法论

1.1 CFSFDP算法的基本原理

快速密度峰值搜索算法^[18](clustering by fast search and find of density peaks, CFSFDP)是一种快速聚类的算法。该算法假设聚类中心的密度大于周围邻点的密度,且聚类中心与其他密度较大的点之间的距离相对更大。其基本思路为:首先,计算每个数据点的局部密度和相邻距离;然后,构建决策图辅助选取聚类中心,将具有极大局部密度且与其他密度更高点有较远距离的数据视为聚类中心;最后,对非聚类中点进行归类。该算法既可以处理非球面的簇,又可

以自动检测簇的数量,并且不需要反复迭代运算,效率较高。

1.1.1 局部密度 ρ 计算

待聚类的数据集 $X = \{x_i | x_i \in R, i = 1, 2, \dots, N\}$,其中, x_i 为数据样本, N 为数据集的长度。通常用欧式距离计算样本点 x_i 与样本点 $x_j (j=1, 2, \dots, N)$ 之间的距离,如式(1)所示:

$$d(x_i, x_j) = \sqrt{\sum_{i=1}^m (x_i^k - x_j^k)^2} \quad (1)$$

式中, $(x_i^k - x_j^k)^2$ 为样本点 x_i 和 x_j 在 k 维空间上的欧式距离, k, m 为空间维数。

根据高斯核函数计算局部密度公式如下:

$$\rho_i = \sum_{i, j \in I_X} e^{-\left(\frac{d(x_i, x_j)}{d_c}\right)^2} \quad (2)$$

式中, ρ_i 为样本 x_i 的局部密度, d_c 为截断距离, $I_X = \{1, 2, \dots, N\}$ 为数据集 X 对应的指标集,而样本点 x_i 的局部密度由截断距离 d_c 的取值确定。实验证明^[19]: d_c 在数据集 X 中所有样本点之间的欧式距离按照依次升序排列的1%~2%的样本间距离之间选取。

1.1.2 高局部密度点距离 δ 确定

聚类中心与周围样本峰值点的距离越大表示聚类效果越好。这个距离表示在所有局部密度大于 ρ_i 的数据点中,与数据点 x_i 最近的数据点与该数据点之间的距离。

那么,高局部密度点距离 δ 可用如下公式表示:

$$\delta_{q_i} = \begin{cases} \min\{d(q_i, q_j)\}, & i \geq 2, j < i; \\ \max\{\delta_{q_j}\}, & i = 1, j \geq 2 \end{cases} \quad (3)$$

式中, $\{q_i\}_{i=1}^N$ 为 $\{\rho_i\}_{i=1}^N$ 的一个呈降序排列的序下标组,即满足 $\rho_{q_1} \geq \rho_{q_2} \geq \dots \geq \rho_{q_N}$ 。

1.1.3 聚类中心的选取

因为聚类中心是局部密度 ρ 和高局部密度点距离 δ 均较大的样本数据点,则聚类中心选择如式(4)所示:

$$\gamma_i = \rho_i \delta_i \quad (4)$$

从式(4)可以看出, γ 的值越高,该样本点被识别为聚类中心的概率越高。最终,算法输出的不同组的聚类都有相应的最大值和最小值,取适当的数值构建阈值向量 \mathbf{M} ,其表达式为:

$$\mathbf{M} = (m_1, m_2, \dots, m_\tau) \quad (5)$$

式中, τ 为算法输出的聚类簇的数量。

1.2 图信号基本理论

1.2.1 图论基本概念

图作为图谱论的基本研究对象,可以表示真实

网络的拓扑结构,也可抽象表示数据集内部的结构^[20]。一个节点的图可记为 $\mathbf{G}(\mathbf{V}, \mathbf{E})$,其中: $\mathbf{V} = \{1, 2, \dots, N\}$, \mathbf{V} 为图的节点集合; \mathbf{E} 为各节点之间形成的边的集合。图 \mathbf{G} 也可以用 N 阶的权重矩阵邻接矩阵 \mathbf{A} 来表示,其元素 $A_{i,j}$ 反映了节点 i 到节点 j 这条边的权重,利用高斯核函数表示如下:

$$A_{i,j} = \begin{cases} \exp\left(-\frac{[\text{dist}(i, j)]^2}{2\theta^2}\right), & \text{dist}(i, j) \leq \omega; \\ 0, & \text{dist}(i, j) > \omega \end{cases} \quad (6)$$

式中, $\text{dist}(i, j)$ 为节点 i 与节点 j 之间的物理距离或节点信号值的差值, ω 为阈值, θ 为尺度参数。

为了确保图信号在计算过程中的数值稳定性,将邻接矩阵 \mathbf{A} 进行归一化处理得到:

$$\mathbf{A}_{\text{norm}} = \frac{1}{|\lambda|_{\max}} \mathbf{A} \quad (7)$$

式中, λ_{\max} 为邻接矩阵特征值绝对值的最大值。 N 节点无向图 \mathbf{G} 的拉普拉斯矩阵为:

$$\mathbf{L} = \mathbf{D} - \mathbf{W} \quad (8)$$

式中: \mathbf{W} 为图 \mathbf{G} 的权重矩阵; $\mathbf{D} = \text{diag}\{d_1, d_2, \dots, d_N\}$, \mathbf{D} 为节点度的对角阵,其对角元素为 $d_i = \sum_{j=1}^N A_{i,j}$, $i = 1, 2, \dots, N$ 。

1.2.2 图信号处理

在经典的信号处理中,噪声是不可避免的,图信号处理亦如此。带有噪声的图信号可以表示为:

$$\mathbf{Y} = \mathbf{S} + \boldsymbol{\eta} \quad (9)$$

式中, \mathbf{S} 为待重构信号的图信号, \mathbf{Y} 为带有噪声的节点图信号向量, $\boldsymbol{\eta}$ 为噪声向量。对图信号处理的基本思路是根据抽取的一定量已知节点数据的拓扑关系和相关的先验信息来对整个图信号进行重构。同时,在信号恢复过程中,噪声也是不可忽略的。对图信号的采样公式表示为:

$$\mathbf{Y}^* = \mathbf{J} \circ \mathbf{S}^* + \mathbf{V} \quad (10)$$

式中: \mathbf{Y}^* 为采样所得数据矩阵, $\mathbf{Y}^* \in \mathbf{R}^{N \times m}$; \mathbf{S}^* 为原始初值数据向量; \mathbf{J} 为采样算子矩阵, $\mathbf{J} \in \mathbf{R}^{N \times m}$,其表示为当节点 i 在第 t_m 时刻被采样时算子值赋值为1,否则赋值为0; \mathbf{V} 为噪声矩阵。

图信号处理中常用图拉普拉斯二次型^[21-22]来研究平滑单一的图信号 \mathbf{S} ,可表示为:

$$S_{2L}(\mathbf{S}) = \mathbf{S}^T \mathbf{L} \mathbf{S} \quad (11)$$

式中, $S_{2L}(\mathbf{S})$ 越小表示图信号越发的平滑。此时可以利用Tikhonov正则化方法^[23],以迭代的方式求解最优解,从而实现图信号的重构。其优化表示如下:

$$\min_{\mathbf{S}} \|\mathbf{J} \circ \mathbf{S} - \mathbf{Y}\|_F^2 + \lambda S_{2L}(\mathbf{S}) \quad (12)$$

式中: λ 为归一化参数; 式(12)的第1部分表示计算误差; 第2部分表示图信号重构的增益, 作为最优化问题的正则项。

2 非侵入式负荷识别的应用

本文首先基于用电设备的先验用电数据利用CFSFDP聚类算法建立设备的初始图信号; 再利用非侵入式负荷监测装置采样获得的用户总用电数据建立总用电图结构以及相关的图拉普拉斯矩阵, 将非侵入式负荷识别问题转化为寻找合适的使图拉普拉斯二次型函数最优的图信号 \mathbf{S} 的重构问题。

2.1 设备用电信号的图结构

设家庭中共有 L 个用电设备且每个设备均存在稳定的工作状态。设备 l 的用电信号包含 N 个时间点, 前 q 点时序信号值已知 (即 $x_1^l, x_2^l, \dots, x_q^l$ 为设备 l 的先验用电数据), 后 $N-q$ 点的信号未知。利用以上数据作为CFSFDP聚类算法的输入, 可以求得式(5)的阈值向量, 那么, 根据设备先验曲线聚类结果, 用电设备已知的先验信息的图信号可表示为:

$$s_i^l = \begin{cases} \tau, & x_i^l > m_\tau; \\ \vdots & \\ 1, & m_1 < x_i^l \leq m_2; \\ 0, & 0 < x_i^l \leq m_1 \end{cases} \quad (13)$$

式中: s_i^l 为用电设备 l 在时刻 i 的图信号值, $i=1, 2, \dots, q$; x_i^l 为用电设备在时刻 i 的先验用电数据。对于未知的 $N-q$ 点的图信号值, 则设其数值为0, 那么, 设备 l 的先验图信号可表示为 $\mathbf{S}^l = [s_1^l, s_2^l, \dots, s_q^l, 0, 0, \dots, 0]^T$ 。

2.2 非侵入式负荷图信号重构

由于家庭中用电设备数量和测试功率序列的长度往往比较大, 针对式(12)的最优化问题求解比较困难。因此, 本文将问题的求解分为两步: 第1步, 通过最小化正则项得到满足全局变化量最小的解; 第2步, 在第1步所得解的基础上对式(12)进行迭代优化, 寻找全局最优解。具体如下:

首先, 进行正则项最小化, 按照设备平均功率递减的顺序依次从总功率信号中分离负荷功率信号, 每分离一个设备, 就从总功率信号中除去, 再分离下一个直到所有设备都分离完成。这一步骤减少了大功率电器对小功率电器的覆盖作用, 识别的准确率明显提高。

其次, 为了能更快速正确地求解式(12)所代表的优化问题, 令其等于函数 $F(\mathbf{S})$, 通过梯度下降的方法迭代寻找最优解。即有:

$$F(\mathbf{S}) = \min_{\mathbf{S}} \|\mathbf{J} \circ \mathbf{S} - \mathbf{Y}\|_F^2 + \lambda S_{2L}(\mathbf{S}) \quad (14)$$

迭代过程中, 设置最大迭代次数为 K , 迭代门限

为 δ , 设置合适的迭代步长 σ 和归一化参数 λ 。则对函数 $F(\mathbf{S})$ 求梯度有:

$$\nabla F(\mathbf{S}) = \mathbf{J} \circ \mathbf{S} - \mathbf{Y} + \lambda \nabla S_{2L}(\mathbf{S}) \quad (15)$$

式中, 梯度下降算法的搜索方向 $\mathbf{d}^k = -\nabla F(\mathbf{S}^k)$, 那么, 迭代公式可表示如下:

$$\mathbf{S}^{k+1} = \mathbf{S}^k + \sigma \mathbf{d}^k \quad (16)$$

式中, $\sigma = \frac{\|\nabla F(\mathbf{S}^k)\|_F^2}{\|\nabla F(\mathbf{S}^{k-1})\|_F^2}$, 经过逐次迭代寻优, 直到达到最大迭代次数 K 或者满足条件 $\|\nabla F(\mathbf{S}^k)\|_F^2 \leq \delta$, 即可找到最优解, 实现对设备图信号的重构。

2.3 用电设备功率信号重构

第2.2节求解得到的图信号各数值蕴含了用电设备的工作情况信息包括启停时刻、持续时间、运行功率。数值的变化时刻即代表工作状态的变化。某工作状态的持续时间可通过信号正值与另一非零值之间的采样点数来确认。根据第1.1.3节的CFSFDP聚类算法求得的阈值向量 \mathbf{M} 来将图信号数值转换为功率信号值。若图信号值小于1, 则对应的功率值为0; 若图信号值介于1~2之间, 则对应的功率值为该设备阈值向量中的 m_1 ; 以此类推, 若图信号值大于 τ , 则取功率信号为 m_τ 。当所有的图信号值转换成功率信号值, 则完成了该用电设备的功率信号重构。

由于总功率信号是多个用电设备叠加而成, 因此每个子功率信号的干扰亦会叠加。为了尽可能地降低单个设备的干扰, 当重构出一个用电设备的功率信号之后, 将此信号从总功率信号中除去, 再进行下一个设备的图信号与功率信号重构。

3 算例与仿真

根据成都某一小区家庭的真实测量数据为基础进行所提方法的仿真与验证。利用基本计量仪表分别在用户的总进线端和各识别负荷的进线端进行数据采集, 采样频率为1/60 Hz。

对负荷分解结果的度量可以从多个角度评判, 本文主要从负荷辨识的辨识精度、准确率、分解精度、用电量误差4个方面对算法结果的合理性进行度量。辨识精度表示分解得到负荷正确的个数和辨识负荷集合的个数的比值, 用于判断算法是否完整正确分解出所有负荷, 表示为:

$$A_p = L'/L \quad (17)$$

式中, L' 为识别出的设备状态个数, L 为家庭中实际有的设备状态总数。

准确率表示每个采样点的分解结果与实际的完全相同的比例表示如下:

$$A_{ac} = R/(R+Q) \quad (18)$$

式中, R 为负荷完全识别正确的时间点数量, Q 为存在负荷错误识别的时间点数量。

分解精度 A_d 用于度量单个负荷分解的效果, 它表征计算得到的设备工作曲线与实际工作情况的接近程度, 具体公式如下:

$$A_d = \frac{1}{T} \sum_{j=1}^T I(P_j = Q_j) \quad (19)$$

式中: P_j 和 Q_j 分别为设备功率曲线的第 j 个功率点设备的开启或关闭状态; 若计算与实际状态相同则 $I(P_j=Q_j)=1$, 反之, $I(P_j=Q_j)=0$ 。

为提高用户节能意识, 需要了解各个负荷的用电量情况。根据负荷辨识的结果, 用电量和用电量误差的计算公式分别如式(20)、(21)所示:

$$\hat{W} = \sum_{t=1}^T \hat{P}_t \Delta t \quad (20)$$

$$\varepsilon = \frac{\hat{W} - W}{W} \times 100\% \quad (21)$$

式中, \hat{W} 为计算所得负荷用电量, \hat{P}_t 为设备计算功率值, Δt 为采样间隔, W 为实际用电量。

3.1 算例1: 典型日负荷分解结果

本算例根据家庭实际测量数据的某1 d的用电数据进行算法分析(对单负荷数据和总功率数据以1 min间隔为采样间隔共1 440个时间点作为实验样本。数据1 min采样作为实际工作曲线用于结果校验)。在此之前, 需要对单设备用电数据进行连续2 d的监测, 利用CFSFDP聚类算法以得到家庭中各电器的所有阈值向量, 为了减少噪声功率带来的影响, 若功率的阈值小于20 W则置为0。该家庭包含有微波炉、洗碗机、电视机、LED灯等9个设备, 共18个稳定工作状态, 设备的状态功率阈值如表1所示。

表1 算例1中的负荷状态功率阈值

Tab. 1 Power thresholds of load states in scenario 1

负荷设备名称	m_1	m_2	m_3
LED灯	40	80	—
洗碗机	120	—	—
洗衣机	200	500	—
冰箱	220	—	—
电视机	250	—	—
电吹风	300	700	1 100
微波炉	600	900	1 500
空调	780	1 280	1 780
火炉	960	1 950	—

基于本文所提的聚类图信号分解算法对家用负

荷进行监测, 监测结果如图1所示, 由图1可知, 本文提出的算法成功识别出了9个家用电器, 共包含12个状态, 辨识精度达到了100%。

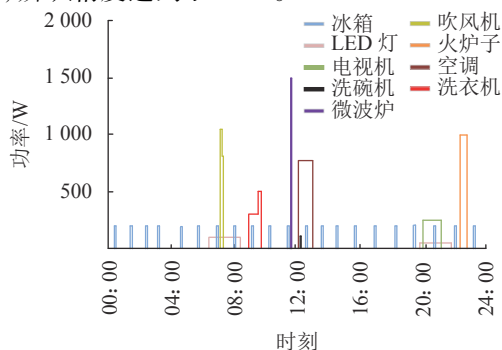


图1 家庭负荷监测结果

Fig. 1 Results of household load monitoring

同时, 图2为家用负荷分解后各个设备的用电量占比的计算值与实际值。由图2可知该场景下家庭中用电量分布比较分散。对比图2(a)、(b)可以看出, 各负荷消耗的用电量比例与实际家用用电量比例的误差低于3%, 表明了所提算法对能耗监测的可行性和较高的准确性。

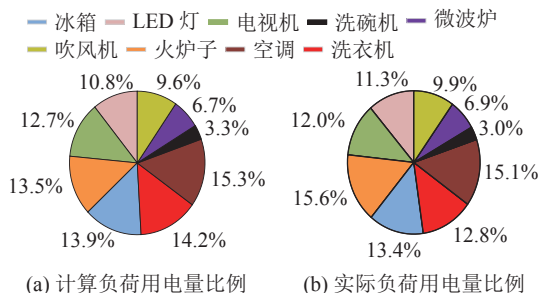


图2 用电设备的用电量占比

Fig. 2 Proportion of power consumption of appliances

3.2 算例2: 典型日负荷分解结果对比

为比较本文所提出的负荷监测算法与其他典型算法的辨识结果, 现仍选取算例1家庭用电数据进行实验验证, 与算例1不同的是, 本算例实验日参与工作的用电设备仅有5个, 这5个设备共存在有9个工作状态(并不代表有9个状态在该日都有开启状态)。对该日的用电数据进行设备数据和总数据进行1 min采样共得到1 440×6个采样点作为实验样本数据。与本文算法比较的其他典型辨识算法分别取自文献[13,24-25], 不同算法识别结果的准确率 A_{ac} 、辨识精度 A_p 的对比如表2所示。

由表2可以看出, 度量4种算法结果的指标基本呈上升趋势, 本文算法的性能相比之下更好。从辨识精度来看, 前两种算法均会发生状态的漏判, 仔细分析发现, LED灯和微波炉在某几个采样间隔内完成了启停动作, 采样得到的负荷总有功率数据中不

包含该电器的运行数据,因而文献[13]的IP算法无法从采样数据中分析出微波炉启停信息。而文献[24]的EOI算法在重构总功率曲线时有较大误差,导致LED灯的其中一个状态没有被监测出来。本文算法与文献[25]的多特征融合算法能完整识别出所有负荷状态。根据重构的设备功率曲线求得各个算法的不同电器的分解精度如表3所示,用电量误差如表4所示,清晰直观地表征了负荷分解的效果。

表2 准确率与辨识精度对比

Tab. 2 Comparison of accuracy and recognition precision

算法	准确率 $A_{ac}/\%$	辨识精度 $A_p/\%$
文献[13]算法	62.7	71.4
文献[24]算法	87.6	85.7
文献[25]算法	89.3	100
本文算法	90.1	100

表3 设备分解精度对比

Tab. 3 Comparison of disaggregation accuracy of appliances

设备	单设备分解精度 $A_d/\%$			
	文献[13]算法	文献[24]算法	文献[25]算法	本文算法
总功率	79.6	82.5	89.6	90.7
LED灯	82.1	89.5	92.2	95.7
电视机	80.2	86.3	93.8	94.5
洗衣机	82.7	90.6	94.6	93.7
微波炉	93.5	95.7	97.3	98.6
冰箱	74.8	87.4	92.7	91.6

表4 用电量误差对比

Tab. 4 Comparison of power consumption error

设备	用电量误差 $\varepsilon/\%$			
	文献[13]算法	文献[24]算法	文献[25]算法	本文算法
LED灯	4.1	-5.9	3.7	-2.7
电视机	8.2	-2.9	-5.6	2.8
洗衣机	-9.3	-4.1	-3.9	-2.4
微波炉	-13.3	6.8	7.3	3.6
冰箱	-28.5	-9.9	5.5	4.2

对比于表2的准确率,表3中设备分解精度的数值明显更大,这是因为大多数电器处于关闭状态的时间更长,导致与实际工作情况重合的时间更多。微波炉的工作时间很短,它的高分解精度源于长时间的关闭状态。由表3可见:本文提出的算法与文献[25]算法的设备分解精度都比较高,表明得到的分解负荷曲线与实际负荷曲线的重合度相对较高。除了不可避免的功率波动造成的误差,设备种类漏判是导致误差增大的最主要因素。文献[13]算法因为未识别出微波炉和LED灯的几个工作周期导致分解精度低。

文献[24]算法在针对小功率电器LED灯时会影响重构功率曲线的效果从而得到较低的分解精度。

从表4可以看出:文献[13]算法的用电量误差数值和波动最大,而本文算法的误差变化最稳定,对比算法的负误差较多是因为设备的部分工作状态未被正确识别出来。而文献[25]算法在辨识精度、设备分解精度都与本文算法较接近的情况下有更高的用电量误差,这表明本文算法在设备功率的重构上更加准确。微波炉的用电量误差较大是因为其工作时间较短的情况下只要启停时间不准确就会对结果有较大影响。本文利用1 min低频用电数据作为已知条件来合理估计工作中的设备,对多解的处理不会影响用电量误差。准确求取家中电器的能耗情况,对于解决实际NILM问题有很大的优势。

3.3 算例3:采样间隔设置的影响

为了测试数据采样间隔对所提算法的影响,本算例在算例2的基础上将采样数据降采样到周期为2 min,对所提算法与对比算法进行进一步的测试,得到准确率和辨识精度如表5所示。

表5 降采样后的准确率与辨识精度对比

Tab. 5 Comparison of accuracy and recognition precision after downsampling

算法	准确率 $A_{ac}/\%$	辨识精度 $A_p/\%$
文献[13]算法	58.4	57.1
文献[24]算法	79.3	71.4
文献[25]算法	80.6	85.7
本文算法	82.7	85.7

对比表5和2可以看出,由于采样间隔的增加,4个算法的辨识精度有所下降,表明算法成功识别的设备状态减少,这是由于设备的工作周期在两次采样间隔之间而无法识别到。相应地,算法的准确率也有较大幅度的降低。因此,采样间隔的增加对算法的准确率和辨识精度有负面影响。为了测试采样间隔对单设备分解性能的影响,根据重构的负荷功率曲线计算设备的分解精度和用电量误差如表6和7所示。

表6 降采样后的设备分解精度对比

Tab. 6 Comparison of disaggregation accuracy of appliances after downsampling

设备	单设备分解精度 $A_d/\%$			
	文献[13]算法	文献[24]算法	文献[25]算法	本文算法
总功率	73.9	79.6	85.3	87.5
LED灯	76.3	85.1	90.7	91.8
电视机	78.6	84.7	82.5	92.7
洗衣机	80.1	80.2	88.1	90.2
微波炉	82.7	84.6	85.9	87.3
冰箱	70.6	77.4	83.9	86.2

表 7 降采样后的用电量误差对比

Tab. 7 Comparison of power consumption error after downsampling

设备	用电量误差 $e/\%$			
	文献[13]算法	文献[24]算法	文献[25]算法	本文算法
LED灯	5.6	-6.8	2.8	-4.6
电视机	8.0	-3.2	-6.1	2.1
洗衣机	-10.8	-7.9	-7.5	-6.5
微波炉	-20.6	1.2	2.8	-2.6
冰箱	-30.9	-12.7	1.7	-3.4

对比表6和3可以看出,各个设备在降采样后的分解精度都有不同程度的下降,其中微波炉的下降幅度最大。这与微波炉的工作状态有关,它的工作周期比较短,当在采样间隔之间运行时就会被算法漏识别。而对比表7和4的结果可以发现,用电量误差在数值上都有一定的下降,而从误差分析角度来讲,部分设备的计算用电量更接近实际情况。这是由于在降采样被漏识别的部分用电量与算法误判的部分用电量抵消了。因此,采样间隔的变化对用电量误差的影响并不明显。一般来讲,数据的采样频率越高包含的有效信息越多,信号重构的效果就越好,算法的准确率也会随之提高。但是,高采样频率对采样设备硬件要求更高,算法的时间复杂度和空间复杂度都呈指数增长,即使算法精度有所提高,所消耗的时间成本和硬件成本都导致算法的整体性能大幅度降低。因此,本文算法在处理采样周期小于2 min的负荷用电数据时效果更好。

对所提算法和对比算法的时间复杂度进行分析。家庭中用电设备数量 L 固定,测试功率序列长度为 N 。测试功率序列的维度和最大迭代次数都可被认为是常数。在负荷监测过程中,利用图拉普拉斯算法进行图信号迭代重构的复杂度为 $O(LT)$,功率信号转换的时间复杂度为 $O(LT)$,则本文所提算法的时间复杂度为 $O(LT)$ 。文献[13,24]算法的整体时间复杂度为 $O(L^2T)$,文献[25]算法的整体时间复杂度为 $O(L^3T)$,证明上述算法在时间复杂度上均可行,并且本文算法相比之下有更优的时间复杂度。总的来说,文献[13,24]算法在识别设备开关状态上的性能较文献[25]算法和本文算法低,而本文算法有更好的重构负荷功率的效果。降采样后各个算法的性能有一定的下降,但是,相比之下本文算法依然有较好的负荷监测效果,针对NILM问题有更大的优势。

4 结 论

本文提出了一种基于CFSFDP图拉普拉斯算法

的非侵入式负荷监测方法。该方法利用一种CFSFDP聚类算法构建家用电器的先验图结构,然后通过Tikhonov正则化方法求解图拉普拉斯二次型优化问题的最优解,将求解得到的图信号转换为功率信号,即可实现用户的非侵入式负荷监测。经过实验验证表明,本文算法能够准确监测家用电器的用电情况,相比于已有算法有更高的准确率和辨识精度。

本文算法仍有一些不足之处和尚未解决的问题。算法无法较好地监测某些在启停时刻有极高暂态功率的用电设备,在算法中没有考虑到用电设备的用电规律和约束。今后的工作将进一步考虑用电设备的运行特征,对运行时功率变化趋势不同对电器进行分类,做更细致的研究。再者,考虑对商业用户用电行为进行监测,制订合理的用电计划,实现资源的有效利用。

参考文献:

- [1] Hart G W. Nonintrusive appliance load monitoring[J]. *Proceedings of the IEEE*, 1992, 80(12): 1870–1891.
- [2] Bouloutas A, Hart G W, Schwartz M. Two extensions of the Viterbi algorithm[J]. *IEEE Transactions on Information Theory*, 1991, 37(2): 430–436.
- [3] Held P, Mauch S, Saleh A, et al. Frequency invariant transformation of periodic signals (FIT-PS) for classification in NILM[J]. *IEEE Transactions on Smart Grid*, 2019, 10(5): 5556–5563.
- [4] Le T T H, Kang H, Kim H. Household appliance classification using lower odd-numbered harmonics and the bagging decision tree[J]. *IEEE Access*, 2020, 8: 55937–55952.
- [5] Ghosh S, Chatterjee A, Chatterjee D. An improved load feature extraction technique for smart homes using fuzzy-based NILM[J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 1–9.
- [6] Le T T H, Heo S, Kim H. Toward load identification based on the Hilbert transform and sequence to sequence long short-term memory[J]. *IEEE Transactions on Smart Grid*, 2021, 12(4): 3252–3264.
- [7] Chen Junfeng, Wang Xue, Zhang Xiaotian, et al. Temporal and spectral feature learning with two-stream convolutional neural networks for appliance recognition in NILM[J]. *IEEE Transactions on Smart Grid*, 2022, 13(1): 762–772.
- [8] Wang Ying, Yang Wei, Xiao Xianrong, 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. [汪颖, 杨伟, 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.]
- [9] Du Renren, Yang Chao, Wang Huayong. Non-intrusive load identification based on wavelet features and support vector

- machine[J].*Journal of Guizhou University(Natural Sciences)*, 2020,37(6):67–72.[杜刃刃,杨超,王华勇.基于小波特征和支持向量机的非侵入式负荷识别[J].*贵州大学学报(自然科学版)*,2020,37(6):67–72.]
- [10] Xu Daqing,Luan Wenpeng,Wang Peng,et al.Analysis and applications of smart meter data[J].*Distribution & Utilization*,2015,32(8):25–30.[徐大青,栾文鹏,王鹏,等.智能电表数据分析方法及应用[J].*供用电*,2015,32(8):25–30.]
- [11] Chen Liang,Lv Fangcheng,Xie Qing,et al.Research of real-time energy measurement and bidirectional communication smart meter[J].*Proceedings of the CSEE*,2011,31(Supp1):94–99.[陈亮,律方成,谢庆,等.实时测量双向通信智能电表的研发[J].*中国电机工程学报*,2011,31(增刊1):94–99.]
- [12] Wang Fanrong,Xiang Kun,Liu Hui.Research on non-intrusive load decomposition based on improved AP clustering and optimized GRNN[J].*Advanced Engineering Sciences*, 2020,52(4):56–65.[汪繁荣,向堃,刘辉.基于改进AP聚类与优化GRNN的非侵入式负荷分解研究[J].*工程科学与技术*, 2020,52(4):56–65.]
- [13] Wittmann F M,López J C,Rider M J.Nonintrusive load monitoring algorithm using mixed-integer linear programming[J].*IEEE Transactions on Consumer Electronics*,2018, 64(2):180–187.
- [14] Liu Yaoxian,Sun Yi,Li Bin.A modified IP-based NILM approach using appliance characteristics extracted by 2-SAX[J].*IEEE Access*,2019,7:48119–48128.
- [15] Liu Yinyan,Qiu Jing,Lu Junda,et al.A single-to-multi network for latency-free non-intrusive load monitoring[J].*IEEE Transactions on Network Science and Engineering*,2022,9 (2):755–768.
- [16] Xu Qingshan,Lou Oudie,Zheng Aixia,et al.A non-intrusive load decomposition method based on affinity propagation and genetic algorithm optimization[J].*Transactions of China Electrotechnical Society*,2018,33(16):3868–3878.[徐青山,娄藕蝶,郑爱霞,等.基于近邻传播聚类 and 遗传优化的非侵入式负荷分解方法[J].*电工技术学报*,2018,33(16):3868–3878.]
- [17] He Jialing,Zhang Zijian,Zhu Liehuang,et al.An efficient and accurate nonintrusive load monitoring scheme for power consumption[J].*IEEE Internet of Things Journal*, 2019,6(5):9054–9063.
- [18] Rodriguez A,Laio A.Clustering by fast search and find of density peaks[J].*Science*,2014,344(6191):1492–1496.
- [19] Xie Juanying,Gao Hongchao,Xie Weixin.K-nearest neighbors optimized clustering algorithm by fast search and finding the density peaks of a dataset[J].*Scientia Sinica (Informationis)*,2016,46(2):258–280.[谢娟英,高红超,谢维信.K近邻优化的密度峰值快速搜索聚类算法[J].*中国科学(信息科学)*,2016,46(2):258–280.]
- [20] Sandryhaila A,Moura J M F.Discrete signal processing on graphs[J].*IEEE Transactions on Signal Processing*,2013,61 (7):1644–1656.
- [21] Shuman D I,Narang S K,Frossard P,et al.The emerging field of signal processing on graphs:Extending high-dimensional data analysis to networks and other irregular domains[J].*IEEE Signal Processing Magazine*,2013,30(3):83–98.
- [22] Batabyal T,Weller D,Kapur J,et al.Efficient learning of transform-domain LMS filter using graph Laplacian[J/OL].*IEEE Transactions on Neural Networks and Learning Systems*[2022-01-01].<https://dx.doi.org/10.1109/TNNLS.2022.314463>.
- [23] Loukas A,Perraudin N.Stationary time-vertex signal processing[J].*EURASIP Journal on Advances in Signal Processing*,2019,2019:36.
- [24] Wu Shuomin,Zhang Kaifeng,Wang Ying,et al.Operation identification method of electrical appliance considering signals of energy consumption based on low frequency sampling[J].*Automation of Electric Power Systems*,2018,42 (17):87–95.[吴烁民,张凯锋,王颖,等.计及低频采样用电量信号的电器运行情况识别方法[J].*电力系统自动化*,2018, 42(17):87–95.]
- [25] Yang Dongsheng,Kong Liang,Hu Bo,et al.Load identification method based on multi-feature sequence fusion[J].*Automation of Electric Power Systems*,2017,41(22):66–73.[杨东升,孔亮,胡博,等.基于多特征序列融合的负荷辨识方法[J].*电力系统自动化*,2017,41(22):66–73.]

(编辑 赵婧)

引用格式: Lin Pingchuan,Lu Lei,Gu Chao,et al.Non-intrusive load monitoring method based on CFSFDP graph laplace algorithm[J].*Advanced Engineering Sciences*,2023,55(4):216–223.[林平川,路磊,谷超,等.基于CFSFDP图拉普拉斯算法的非侵入式负荷监测方法[J].*工程科学与技术*,2023,55(4):216–223.]