

文章编号: 1005-0523(2021)05-0056-09



基于MK变点检测和加权HMM算法的NILM研究

傅军栋, 耿余, 赵颖, 刘珺

(华东交通大学电气与自动化工程学院, 江西 南昌 330013)

摘要:为了更可靠的提升居民用户非侵入式负荷的分解能力, 将整个识别过程分为负荷事件识别和状态识别两个过程。分别提出一种基于双滑动窗的MK变点检测算法以及加权双参量隐马尔可夫模型进行负荷事件识别和状态识别。并在状态识别中将每个设备各内部状态在每个时间段的出现概率作为权值带入到算法。通过数据验证以及对比实验, 表明NILM算法能够较为有效地对居民负荷进行识别。

关键词:非侵入式负荷监测; MK变点检测; 双滑动窗; 隐马尔可夫模型

中图分类号: TM933

文献标志码: A

本文引用格式:傅军栋, 耿余, 赵颖, 等. 基于MK变点检测和加权HMM算法的NILM研究[J]. 华东交通大学学报, 2021, 38(5): 56-64.

DOI: 10.16749/j.cnki.jecjtu.20211026.006

NILM Research Based on Mann-Kendall and Weighted HMM Algorithm

Fu Jundong, Geng Yu, Zhao Ying, Liu Jun

(School of Electrical & Automation Engineering, East China Jiaotong University, Nanchang 330013, China)

Abstract: In order to more reliably improve the non-intrusive load decomposition ability of residents, the whole identification process is divided into two processes: load event identification and state identification. A MK change point detection algorithm based on dual sliding windows and a weighted dual-parameter hidden Markov model are respectively proposed for load event recognition and state recognition. In the state recognition, the appearance probability of each internal state of each device in each time period is brought into the algorithm as a weight. Through data verification and comparative experiments, it is shown that the NILM algorithm mentioned can more effectively identify the residents' load.

Key words: NILM; MK change point detection; double sliding window; HMM

Citation format: FU J D, GENG Y, ZHAO Y, et al. NILM research based on mann-kendall and weighted HMM algorithm[J]. Journal of East China Jiaotong University, 2021, 38(5): 56-64.

非侵入式负荷监测(non-intrusive load monitoring, NILM)早在20世纪80年代就已经提出。近年来, 如何科学用电、引导用户侧提高用电效率引起了诸多关注^[1]。所谓的非侵入式就是在用户入口处采集用户总负荷的能耗等用电信息, 通过一定的方

法分解识别出末端的各个设备^[2]。此方式较侵入式监测投入小, 实用性强。成熟的非侵入式负荷监测模型将会对包括智能电表等用户侧设备的革新带来新的方向, 同时也可以优化故障检测及定位、节能等研究内容, 具有很大的实用价值。

收稿日期: 2021-01-11

基金项目: 江西省科技厅重点研发计划(20192BBE50017)

国内外对于NILM的研究一直持续着,孙毅^[3]将总谐波失真系数作为负荷新特征引入目标函数,采用基于动态粒子群自适应算法的非侵入式负荷分解方法对实测用电数据进行负荷分解,但是两个方法都采用仿真所得数据,过于理想化,实用价值不高。刘兴杰^[4]提出一种基于改进鸡群算法的负荷监测方法,综合考虑稳态谐波电流和功率特征的正态分布度量函数。对于变功率负荷,两种方法的识别能力较差。Racines^[5]采用两个独立的神经网络模型,分别对事件和负载状态进行识别,但是模型之间的联系不够紧密,导致识别率不高。牟魁翌^[6]提出了一种基于分段线性近似及高斯动态弯曲核支持向量机的非侵入式负荷监测方法。Wang^[7]提出一种基于V-I轨迹特征支持向量机的多分类算法进行负荷识别。两个模型的识别精度都有所提高,但基于支持向量机模型的算法对于相似负荷的识别无法达到预期效果。张丽^[8]基于智能负荷控制器(SRLC)的用电参数检测功能和非侵入式负荷监测原理,提出一种基于多参量隐马尔可夫模型的负荷辨识方法,但是由于实验在特定环境下进行,缺乏多样性,导致在个别环境下,识别误差增大。王晓换^[9]提出了一种决策融合方法。由于模型的复杂化,导致加权值无法准确选取,降低了识别精度。

针对以上研究中存在的不足,通过对几种具有代表性的家用电器的分析,将用于环境变化检测以及降水趋势变化中的Mann-Kendall(MK)^[10]变点检测算法应用于NILM中的事件检测中,并加以改进,提出基于双滑动窗的MK变点检测算法,并以每个设备各内部状态在每个时间段的出现概率作为权值带入到算法中,以负荷有功功率和电流电压相位差为参数,提出加权双参量隐马尔可夫模型实现负荷状态识别。最后以ECO数据集^[11]提供的的数据验证算法的先进性。

1 双滑动窗MK变点检测原理

负荷状态变化时,会伴随着从一个稳态到另一个的稳态的突变过程,可以通过变点识别的方法检测到状态改变事件。史树彬^[12]提出了一种基于复合滑动窗的CUSUM暂态事件检测算法,但是该方法需要负荷参数在投切时的爬坡过程来累计偏移量,适用于高频的负荷参数,且出现大量的漏检现象。

在低频情况下,CUSUM变点检测算法无法满足要求。为了满足低频信号下能够准确检测到突变点,将MK变点检测应用到负荷事件检测中,该方法具有很高的抗干扰能力,同时也避免了漏检和误检的发生。

1.1 非参数化的MK变点检测算法原理

负荷状态的改变,将检测到负荷信息从一个稳态到另一个稳态的转变,这样的过程可以通过变点检测的方法检测到。变点检测即利用一定的方法,对时间序列的趋势状态进行检测^[12]。

MK是一种非参数统计检验方法,该方法的样本不遵从特定的分布方式,检测结果也不会因为个别异常值而出现很大的误差,且能客观地展现序列变化趋势^[13]。本文对有功功率的时间序列 P 作MK变点检测,对 P (含有 n 个样本)构造一个秩序列

$$s_k = \sum_{i=1}^k r_i (k=2,3,\dots,n) \quad (1)$$

式中: $r_i = \begin{cases} 1, & p_i > p_j \\ 0, & \text{其他} \end{cases}$ ($1 \leq j \leq i$),秩序列 s_k 是第 i 个时刻数值大于第 j 个时刻时,数值个数的累加。已知序列 P 是关于有功功率的时间序列,是随机序列。可定义统计量

$$UF_k = \frac{[s_k - E(s_k)]}{\sqrt{D(s_k)}} (k=1,2,\dots,n) \quad (2)$$

式中: $UF_k=0$, $E(s_k)$ 和 $D(s_k)$ 分别是 s_k 的均值和方差,且 p_1, p_2, \dots, p_n 互相独立时,它们具有相同连续分布,可以推算出

$$\begin{cases} E(s_k) = \frac{n(n+1)}{4} \\ D(s_k) = \frac{n(n-1)(2n+5)}{72} \end{cases} \quad (3)$$

UF_k 为标准正态分布,是在时间序列 P 的正顺序情况下计算出的统计量,并形成相应的序列。按时间序列 P 的逆序列,重复上述公式计算 UF_k ($k'=n+1-k$),并且令 $UB_{k'} = -UF_{k'}$,当 $k'=1$ 时, $UB_1=0$ 。

给定显著性水平 α ,将 UF_k 和 UB_k 序列曲线与 U_α 的两条临界直线一起绘制。若 UF_k 的值大于0,则表明有功功率序列呈上升趋势,小于0则表明呈下降趋势。如果 UF_k 和 UB_k 两条曲线出现交点,且交点在显著性水平曲线之间,则交点对应时刻或者前一时刻即为事件发生的时刻。

MK 变点检测在处理变点检测问题上有很大的优势。然而,在处理多次突变的时间序列时,由于 UF_k 和 UB_k 的在图像中的完整性,会出现无法完全检测突变点的情况,导致突变点的漏检和误检,本文基于这一问题提出复合滑动窗的方法将 MK 变点检测序列细分成更小的窗口进行逐一检测,这一手段也大大减少了 MK 变点检测的工作量。

1.2 双滑动窗变点检测

在实际的用电情况中,无事件情况下功率等参数波动小。当出现设备状态改变时,伴随着负荷参数的突变,窗口的均值和方差也会突变。根据这一特性,提出基于双滑动窗的 MK 变点检测算法,其检测过程如图 1 所示。

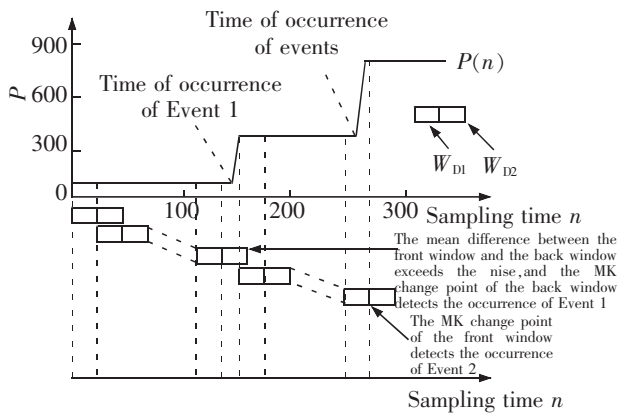


图 1 基于双滑动窗的 MK 变点检测过程图
Fig.1 MK change point detection process based on double sliding windows

在图 1 中,前窗口 W_{D1} 的宽度为 L_1 ,后窗口 W_{D2} 的宽度为 L_2 。 W_{D1} 和 W_{D2} 互为均值窗口,同时也互为检测窗口。 W_{D1} 和 W_{D2} 均值分别为 $\mu_{W_{D1}}$ 和 $\mu_{W_{D2}}$,并比较均值差是否在噪声 h 范围内。若 $|\mu_{W_{D1}} - \mu_{W_{D2}}| \leq h$,则判断窗口内无事件发生;若 $|\mu_{W_{D1}} - \mu_{W_{D2}}| > h$,则判断窗口内有事件发生,再分别对 W_{D1} 和 W_{D2} 进行 MK 变点检测,检测出事件发生的时刻,如果 W_{D1} 和 W_{D2} 均未检测到事件,则对 $W_{D1} + W_{D2}$ 的大窗口进行检测,判断是否在 W_{D1} 末端和 W_{D2} 前端。当检测到事件时,窗口滑动到检测到变点的时刻;如果未检测到事件,直接滑动 L_1 步长继续检测。两次事件的间隔只要超过 L_1 ,就能被检测出来。 L_1 、 L_2 以及 h 取值对检测结果有着一定的影响,本文结合负荷曲线的特点以及各负荷的波动特性经过不断的参数优化,

取 $L_1=L_2=20, h=15, \alpha=0.05, U_\alpha=\pm 1.96$ 。

考虑到事件发生在 W_{D1} 前端或者 W_{D2} 尾端时,均值 $\mu_{W_{D1}}$ 和 $\mu_{W_{D2}}$ 的差值可能不明显,增加了漏检的概率。可以适当增加大窗口两端采样点的参考价值。为此对大窗口内每个采样点值进行加权处理,采用类似于正态窗的方式获取 W_{D1} 和 W_{D2} 的均值。

通过对正态分布函数进行变换用于确定权值

$$f(a_i) = \theta + \frac{1}{\sigma\sqrt{2\pi}} \left(1 - e^{-\frac{(a_i - a_s)^2}{2\sigma^2}} \right) \quad (4)$$

式中: a_i 表示大窗口内每个采样点的位置; a_s 表示大窗口中间采样点位置,即 W_{D1} 尾部和 W_{D2} 头部重叠采样点位置; θ 为 0~1 的一个系数; σ 为大于 0 的常数,决定权值曲线的平滑程度,其值越大,权值曲线越平滑,每个采样点作用越均匀。通过调节 σ , 可以改变大窗口两端采样值在计算均值时的权重。

由式(4)可知,中间采样点 a_s 处权值 $f(a_s) = \theta$, 为最小权值。对 $f(a_i)$ 进行归一化处理,得到权值 $\varphi(a_i)$, 最终得到窗口 W_{D1} 和 W_{D2} 的均值 $\mu_{W_{D1}}$ 和 $\mu_{W_{D2}}$ 为

$$\begin{cases} \mu_{W_{D1}} = \frac{\sum_{a_i=k_{start}}^{k_{start}+L_1-1} \varphi(a_i)W(a_i)}{L_1-1} \\ \mu_{W_{D2}} = \frac{\sum_{a_i=k_{start}+L_1+L_2-2}^{k_{start}+L_1+L_2-1} \varphi(a_i)W(a_i)}{L_2-1} \end{cases} \quad (5)$$

式中: k_{start} 表示窗口开始位置时刻; L_1 和 L_2 为 W_{D1} 和 W_{D2} 的窗口宽度; $W(a_i)$ 为窗口中采样点 a_i 的采样值。

2 加权双参量隐马尔可夫状态识别

获取事件点及对应负荷参数后,即可进行该突变点的设备状态识别,即识别出每次突变后各设备的运行状态。文献[8, 14-15]均采用隐马尔可夫模型(hidden Markov model, HMM)进行负荷识别,在 HMM 的基础上做了相应的改进,分别提出基于免疫优化和增量学习改进的 HMM^[14]、可加因子 HMM^[15] 以及多参量 HMM 算法^[8]。然而 3 种方法均未进行前期的事件识别操作,每个采样点均需要进行状态识别,算法运行负担加大。对此本文在识别负荷事件的基础上,采用有功功率和电流电压相位差两个参量,并以每个设备各内部状态在每个时间段的出现概率作为权值带入到算法中,提出基于加权双参量

HMM 的负荷状态识别方法。将事件点负荷参数带入 HMM 中,获取各状态组合的概率并通过负荷各状态出现概率作加权处理,找出最大概率对应状态组合即为该事件发生后各设备的状态,最终实现负荷识别的目的。

2.1 隐马尔可夫模型原理

HMM 作为典型的机器学习模型在模式识别等领域得到了广泛的应用。在 HMM 模型中,主要有以下几个概念:隐藏状态集合 Q ,观测集合 V ,状态序列 I ,观测序列 O ,状态转移矩阵 A ,观测概率矩阵 B 以及初始概率分布向量 C 。

在 HMM 中,采用三元组 $\lambda=(A, B, C)$ 表示一个 HMM 模型。在预测模型中,通常采用 Viterbi 算法对隐藏状态序列 I 进行解码。在给定了 $\lambda=(A, B, C)$ 及观测序列 O 后,即可解码出观测序列 O 对应的状态序列 I 。其过程如下:

- 1) 需要进行局部状态的初始化

$$\left. \begin{aligned} \delta_1(i) &= c_i b_i(o_1), i=1, 2, \dots, N \\ \psi_1(i) &= 0, i=1, 2, \dots, N \end{aligned} \right\} \quad (6)$$

式中: $\delta_1(i)$ 为初始化中在观测值为 o_1 时隐藏状态为 q_i 的概率; c_i 为初始状态为 i 的初始概率; b_i 为观测值为 o_1 时隐藏状态为 i 的概率; $\psi_1(i)$ 为初始化中隐藏状态 i 的索引值。

- 2) 在初始化局部状态后,即可通过观测序列 O 以及 $\lambda=(A, B, C)$ 递推每个观测值对应的局部状态

$$\left. \begin{aligned} \delta_t(j) &= \max_{1 \leq i < N} [\delta_{t-1}(i) a_{ij}] b_j(o_t), i=1, 2, \dots, N \\ \psi_t(j) &= \arg \max_{1 \leq i < N} [\delta_{t-1}(i) a_{ij}], i=1, 2, \dots, N \end{aligned} \right\} \quad (7)$$

式中: $\delta_t(j)$ 表示在 t 时对应的最大概率隐藏状态为 j ; a_{ij} 为从状态 i 转移到状态 j 的状态转移概率; $b_j(o_t)$ 表示观测值为 o_t 时隐藏状态为 j 的概率; $\psi_t(j)$ 表示 t 时最大概率隐藏状态为 j 的索引值。

- 3) 当式(7)递推至观测序列末端 t_{\max} 时,终止递推

$$\left. \begin{aligned} \delta_{t_{\max}} &= \max_{1 \leq i < N} \delta_{t_{\max}}(i) \\ \psi_{t_{\max}} &= \arg \max_{1 \leq i < N} [\delta_{t_{\max}}(i)] \end{aligned} \right\} \quad (8)$$

- 4) 当递推至 $\delta_{t_{\max}}$ 和 $\psi_{t_{\max}}$ 后,通过 $\psi_t(i)$ 回溯每个 o_t 对应的状态,其中 $t=t_{\max}-1, t_{\max}-2, \dots, 2, 1$ 。最终得到状态序列 $I=(i_1, i_2, \dots, i_{t_{\max}})$ 。

2.2 设备状态组数的确定

实际运行中,很多设备比如冰箱存在关机、待

机、制冷等多个工作状态,需要确定每个设备的状态数。本文采用 PMF 方法^[16]进行设备状态数的确定。

在 PMF 中,通过统计各设备有功功率可能存在的取值的概率来确定设备状态数。公式如下:

$$\eta(l) = \begin{cases} \eta(P=l), l \in \{0, 1, 2, \dots, P_{\max}\} \\ 0 \end{cases} \quad (9)$$

式中: η 为设备有功功率为 l 的概率; P_{\max} 为设备最大的有功功率取值。

在获取设备状态数时,认为设备状态可由 PMF 的峰值来确定。当出现 $\eta(l)-\eta(l-1)>0$ 且 $\eta(l)-\eta(l+1)<0$ 时,即认为 $P=l$ 为设备的一个状态,且 l 为该状态下设备的有功功率取值。赋予设备每个状态一个索引值,其中包括设备的关闭状态,索引值为“1”即 $l=0$ 。

通过 PFM 方法分别获取各个设备的状态数并赋予各状态索引值。即可获取隐藏状态集合 Q 中的各状态组合

$$S=(k^{(1)}, k^{(2)}, \dots, k^{(N)}) \quad (10)$$

式中: S 为当前的设备隐藏状态组合; $k^{(n)}$ 为第 n 个设备的状态索引,对应设备的内部状态。

隐藏状态集合 Q 即由设备状态组合 S 组成,其组合 S 的总数表示为

$$\text{num}(Q) = \prod_{n=1}^N K^{(n)} \quad (11)$$

式中: N 为设备总数; $K^{(n)}$ 设备 n 的内部状态总数。例如,当存在 4 个设备,各设备内部状态数分别为 4, 5, 2 和 7。则组合 S 的总数 $\text{num}(Q)=4 \times 5 \times 2 \times 7=280$ 个。

2.3 加权双参量隐马尔可夫模型

关于状态转移概率矩阵 $A_{\text{num}(Q) \times \text{num}(Q)}$, a_{ij} 表示从状态 i 转移到状态 j 的概率。通过计算状态 i 转换为状态 j 的次数并通过从状态 i 转换所有状态的总数进行归一化,以最大似然准则估算 a_{ij}

$$a_{ij} = \frac{D_{ij}}{\text{num}(Q)}, i, j=1, 2, \dots, \text{num}(Q) \quad (12)$$

$$\sum_{j'=1} D_{ij'}$$

式中: D_{ij} 是从状态组合 i 到状态组合 j 的转变数; $\text{num}(Q)$ 为状态总数。由于该算法是在检测到变点后,认为不存在状态没有改变的情况,即矩阵 A 对角线上的值全部设为 0。

同理,也可以得出各种状态组合下对应各观测值组的概率矩阵 $B_{\text{num}(Q) \times \text{num}(V)}$ 。

$$b_{iv} = \frac{E_{iv}}{\sum_{i'=1}^{\text{num}(Q)} E_{i'v}}, \begin{cases} i=1, 2, \dots, \text{num}(Q) \\ v=1, 2, \dots, \text{num}(V) \end{cases} \quad (13)$$

式中: b_{iv} 为观测值 v 对应状态组合 i 的概率; $\text{num}(v)$ 为观测值总数; E_{iv} 为观测值 v 对应状态 i 的数量。

不同的设备每天的开启时间是不一样的,以至于设备每个内部状态出现的时间段也是不同的,本文将一天分为 24 个时间段,统计每个设备各内部状态在每个时间段内出现的概率

$$\phi_{mk} = \frac{F_{mk}}{\sum_{k'=1}^K F_{mk'}}, \begin{cases} m=1, 2, \dots, 24 \\ k=1, 2, \dots, K \end{cases} \quad (14)$$

式中: ϕ_{mk} 为设备内部状态 k 在第 m 个时间段内出现的概率; F_{mk} 是在第 m 个时间段内设备内部状态 k 出现的总数。

在进行状态识别时即采用 Viterbi 算法获取设备的状态组合。不同的是,在计算 δ 时,由于观测值序列对应为事件发生后的负荷参数,即认为当前状态组合只与前一状态组合有关,且所有观测值之间相互独立,只与生成它的状态组合有关。即

$$\left. \begin{aligned} P(i_t | i_{t-1}, o_{t-1}, i_{t-2}, o_{t-2}, \dots, i_1, o_1) &= P(i_t | i_{t-1}) \\ P(o_t | i_t, o_t, i_{t-1}, o_{t-1}, \dots, i_1, o_1) &= P(o_t | i_t) \end{aligned} \right\} \quad (15)$$

状态序列 I 及观测值序列 O 均只包含当前事件和前一事件相关参数。当检测到事件发生,即可得到该时刻的观测值组 o_t , 并列入观测序列 $O = o_{t-1}, o_t$ 中。HMM 通过观测值 o_t , 结合每个设备各内部状态在每个时间段内出现的概率?计算当前事件后处于不同状态组合的最大概率 δ_t :

$$\delta_t(j) = \max \left[\left[\prod_{n=1}^N \phi_{nk}^{(n)} \right] a_{ij} b_j(o_t) \right] \quad (16)$$

式中: $\phi_{nk}^{(n)}$ 表示第 n 个设备在当前状态组合为 j , 自身内部状态为 k , 本次事件发生时间段为 m 的情况下的出现概率; a_{ij} 为从状态组合 i 转移到状态组合 j 的状态转移概率; $b_j(o_t)$ 为观测值为 o_t 时隐藏状态组合为 j 的概率。

最大的 δ_t 对应的隐藏状态组合即为本次事件后各设备状态的组合。通过状态组合即可获取每个设备的状态信息,实现负荷识别的目的。

3 算法验证

本文采用 ECO (electricity consumption & occupancy, ECO) 数据集对算法进行验证。该数据集以 1HZ 的频率收集了瑞士 6 个家庭历时 8 个月的总负荷以及各用电设备参数。包含功率、电流、电压、相位差等数据。

3.1 事件识别算法验证

选取该低压台区其中一户家庭某一周的总负荷数据进行算法验证。对该周用户负荷事件进行统计并采用基于双滑动窗的 MK 变点检测方法进行负荷事件识别。计算事件识别方法的精确率 P 及准确率 E

$$\left. \begin{aligned} P &= \frac{T_P}{T_P + F_P} \\ E &= \frac{T_P + T_N}{T_P + F_N + F_P + T_N} \end{aligned} \right\} \quad (17)$$

式中: T_P 表示模型判断结果为正例,且实际上判断结果为正确,即真正例; T_N 表示模型判断结果为负例,且实际上判断结果为正确,即真负例; F_P 表示模型判断结果为正例,但实际上判断结果为错误,即假正例; F_N 表示模型判断结果为负例,但实际上判断结果为错误,即假负例。

经统计该周共发生了 417 次负荷波动,其中设备状态改变事件为 406 次,采用基于双滑动窗的 MK 变点检测算法识别的事件结果为:真正例 $T_P=399$;假负例 $F_N=7$;假正例 $F_P=4$;真负例 $T_N=7$ 。即精确率 $P = \frac{T_P}{T_P + F_P} = 0.990$, 准确率 $E = \frac{T_P + T_N}{T_P + F_N + F_P + T_N} = 0.966$ 。由此可见,将基于双滑动窗的 MK 变点检测算法应用在负荷事件识别中具有一定的先进性。

选取该低压台区其中一户家庭某一天的总负荷数据进行 NILM 算法验证。图 2 为家庭该天的 12 时至 14 时的总有功率曲线,通过统计可知该时间段内实际发生了 14 次事件。

对该时间段内的设备事件进行监测。在进行窗口检测的过程中,通过得到的 UF 与 UB 曲线及时的确定窗口内是否发生负荷事件以及其确切的时间,并在监测到事件后进入设备状态识别阶段。

通过基于双滑动窗的 MK 变点检测的算法,对图 2 整个时间段的负荷数据进行事件检测,检测出事件数为 14 个,分别发生在以下时刻:12:17:29, 12:

35:41, 12:48:18, 12:56:49, 13:09:06, 13:10:48, 13:15:31, 13:36:43, 13:37:03, 13:54:06, 13:57:00, 13:57:41, 13:58:14, 13:58:36。这与实际的设备启停事件发生时刻几乎一致。在图 3 中分别取第 46 098 s, 第 44 249 s 的事件对应窗口进行分析。

图3(a)的 UF 和 UB 曲线存在大于 0 的交点,且交点在临界线之间,即在交点前一刻出现了功率突增的情况;图 3(b)的两条曲线出现了小于 0 的交点,且交点在临界线之间,表明在交点前一刻出现了功率突减的情况,这与实际情况完全吻合。

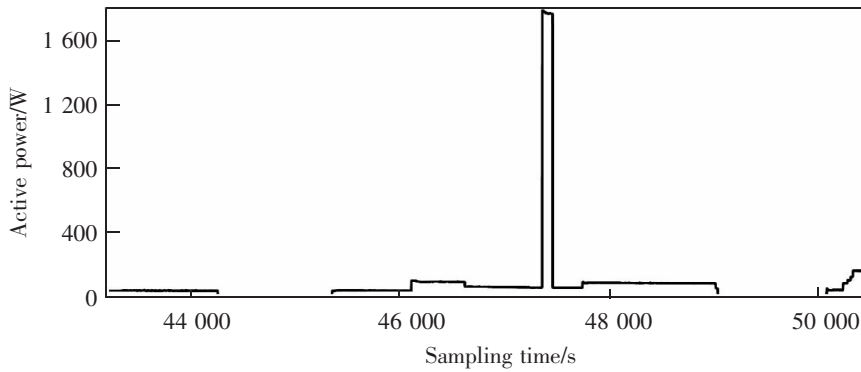


图 2 家庭 1 某时间段总负荷有功功率曲线图

Fig.2 Family 1 total load active power curve for a certain period of time

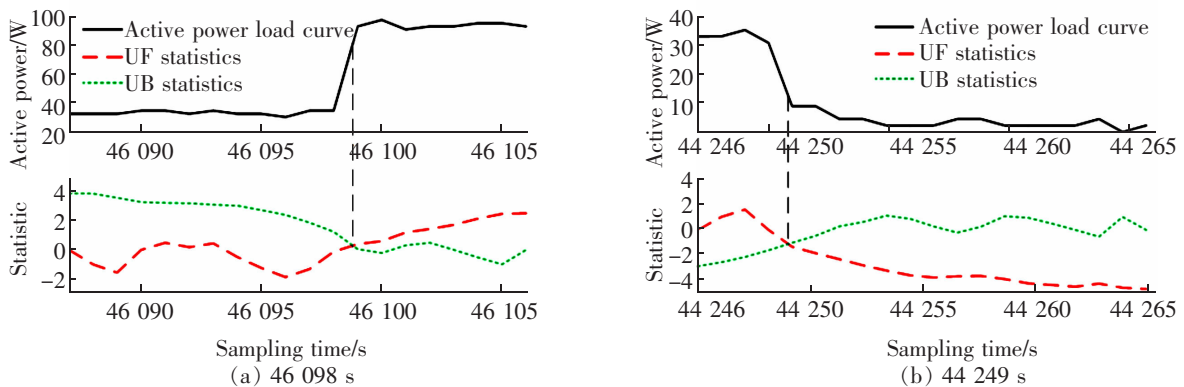


图 3 MK 变点检结果

Fig.3 MK change check result

3.2 状态识别算法验证

上文已验证基于双滑动窗的 MK 变点检测算法的先进性,在此基础上进行基于加权双参 HMM 的设备状态识别算法验证。以家庭 1 连续 180 d 的负荷数据作为统计数据,在该时间段里,待识别设备包括冰箱 1、水壶、洗衣机、冰箱 2、台式电脑和照明 6 个设备。在此基础上,以统计天以外的负荷数据作为验证数据进行算法验证。

将本文提出的 MK-WHMM 算法与孙毅^[14]中的 IA-IL-HMM 算法、Bomfifli^[15]提出的 AFHMM 算法以及基于 k-means 的 NILM 算法进行对比,分别从状态识别准确率和有功功率平方误差两个方面进行算法性能比较。两种性能指标如式(18)式(19)所示。

$$R_s = \frac{\text{num}(S_T)}{\text{num}(S_T) + \text{num}(S_F)} \times 100\% \quad (18)$$

式中: R_s 为状态识别准确率; $\text{num}(S_T)$ 为正确识别状态数; $\text{num}(S_F)$ 为错误识别状态数。

$$R_{PE} = \frac{\sum_{t=1}^T (P_t - \bar{P}_t)^2}{\sum_{t=1}^T P_t^2} \quad (19)$$

式中: R_{PE} 为有功功率平方误差; P_t 为 t 时刻设备实际有功功率; \bar{P}_t 为 t 时刻算法估计设备有功功率。

如图 4 和图 5 分别为 4 种 NILM 算法对家庭 1 的冰箱 1、水壶、洗衣机、冰箱 2、台式电脑以及照明设备的状态识别精度及有功功率平方误差对比。

由图4可知,在对该家庭的6个用电设备进行NILM算法识别时,本文提出的方法状态识别精度均高于其他3种方法,较其他3种算法在负荷状态识别上更具有优势,且对于冰箱1和冰箱2两个负荷特性较为相似的设备以及照明设备等负荷较小的设备同样可以进行有效的识别。

在图5中,除水壶和照明设备的功率平方误差较其他3种算法具有优势,其他4个设备的功率平方误差较文献[14-15]的方法更大。分析可知,这4个设备均为多状态或者负荷波动频繁的设备,本文算法旨在检测到负荷事件后进行负荷状态识别,这

样大大减小了负荷状态识别的负担,但设备在同一状态下负荷功率保持为一个定值,忽略了设备自身的负荷波动,从而增加了有功功率平方误差。而文献[14-15]跳过负荷事件检测这一步,直接通过对每一个采样点都进行一次HMM负荷状态识别,使得负荷功率更加接近实际,但同时也增加了HMM状态识别的负担。综合各方面因素,本文NILM算法较另外3种算法具有一定的先进性。

在此基础上,随机选择除家庭1外任意两户居民,以设备状态识别准确率进行两个场景的算法性能比较,结果如表1和表2所示。

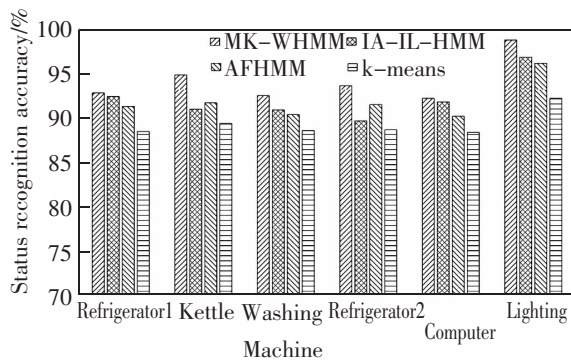


图4 家庭1的设备状态识别准确率对比
Fig.4 Comparison of device status recognition accuracy rate of Family 1

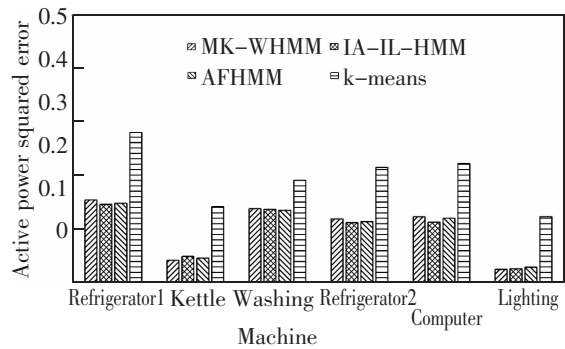


图5 家庭1的设备有功功率平方误差对比
Fig.5 Comparison of the square error of the active power of the equipment of Family 1

表1 场景1设备状态识别准确率对比

Tab.1 Comparison of device status recognition accuracy rate of Scenario 1

Algorithm	Washing machine	Refrigerator	Computer	TV	Air conditioner	Illumination
MK-WHMM	93.3	93.5	92.7	92.3	93.9	97.4
IA-IL-HMM	92.0	91.7	91.5	91.9	91.5	94.5
AFHMM	91.8	92.1	91.0	91.4	91.2	94.8
k-means	88.2	88.1	88.3	87.8	88.1	90.9

表2 场景2设备状态识别准确率对比

Tab.2 Comparison of device status recognition accuracy rate of Scenario 2

Algorithm	Kettle	Refrigerator	TV	Drying machine	Washing machine	Computer	Illumination
MK-WHMM	95.5	92.9	91.2	93.3	92.4	91.9	97.4
IA-IL-HMM	92.5	91.7	90.1	91.4	90.5	90.2	95.8
AFHMM	92.8	91.4	89.7	91.5	90.0	90.9	94.3
k-means	88.2	88.1	87.3	88.2	88.1	87.0	91.1

由表1和表2可见,本文提出的NILM算法能够适应不同的场景,且相比于其他3种方法效果明显。根据以上验证可知通过本文提出的NILM算法可以实现对居民用户内负荷进行分解。

4 结论

本文将NILM具体分为负荷事件识别和负荷状态识别两个部分。

1) 采用基于双滑动窗的MK变点检测算法进行负荷事件识别,同时在HMM的基础上以每个设备各内部状态在每个时间段的出现概率作为加权值,提出基于加权双参量HMM的负荷状态识别方法。

2) 通过验证,结果表明该NILM算法能够及时有效地识别出用户内各设备的运行状态。为后期与IEC61850结合的低压用户用电信息通信提供技术支撑,同时为今后基于NILM的相关研究提供参考。

参考文献:

- [1] 张巧来,马运东,王愈. 基于用户侧需求响应的互联微电网研究[J]. 电测与仪表,2019,56(10):34-39.
ZHANG Q L,MA Y D,WANG Y. Research on micro-grid interconnection based on demand side response[J]. Electrical Measurement & Instrumentation,2019,56(10):34-39.
- [2] 程祥,李林芝,吴浩,等. 非侵入式负荷监测与分解研究综述[J]. 电网技术,2016,40(10):3108-3117.
CHENG X,LI L Z,WU H,et al. A Survey of the research on non-intrusive load monitoring and disaggregation[J]. Power System Technology,2016,40(10):3108-3117.
- [3] 孙毅,张璐,赵洪磊,等. 基于动态自适应粒子群算法的非侵入式家居负荷分解方法[J]. 电网技术,2018,42(6):1819-1826.
SUN Y,ZHANG L,ZHAO H L,et al. A non-intrusive household load monitoring method based on dynamic adaptive particle swarm optimization algorithm[J]. Power System Technology,2018,42(6):1819-1826.
- [4] 刘兴杰,曹美晗,许月娟. 基于改进鸡群算法的非侵入式负荷监测[J]. 电力自动化设备,2018,38(5):235-240.
LIU X J,CAO M H,XU Y J. Non-intrusive load monitoring based on improved chicken swarm optimization algorithm[J]. Electric Power Automation Equipment,2018,38(5):235-240.
- [5] RACINES D,CANDELO J E,MONTANA J. Non-intrusive electrical load monitoring system applying neural networks with combined Steady-state electrical variables[J]. Tehnicki Vjesnik,2018,25(5):1321-1329.
- [6] 牟魁翌,杨洪耕. 基于PLA-GDTW支持向量机的非侵入式负荷监测方法[J]. 电网技术,2019,43(11):4185-4193.
MOU K Y,YANG H G. Non-intrusive load identification method based on PLA-GDTW support vector machine[J]. Power System Technology,2019,43(11):4185-4193.
- [7] WANG A L,CHEN B X,WANG C G,et al. Non-intrusive load monitoring algorithm based on features of V-I trajectory[J]. Electric Power Systems Research,2018,157:134-144.
- [8] 张丽,张涛,张宏伟,等. 一种基于多参量隐马尔可夫模型的负荷辨识方法[J]. 电力系统保护与控制,2019,47(20):81-90.
ZHANG L,ZHANG T,ZHANG H W,et al. Research on a method of load identification based on multi parameter hidden Markov model[J]. Power System Protection and Control,2019,47(20):81-90.
- [9] 王晓换,李如意,周东国,等. 基于决策融合的非侵入式电力负荷辨识方法及应用[J]. 电力系统保护与控制,2016,44(7):115-121.
WANG X H,LI R Y,ZHOU D G,et al. Non-intrusive power load disaggregation method based on decision fusion and its applications[J]. Power System Protection and Control,2016,44(7):115-121.
- [10] GIOIA A,BRUNO M F,TOTARO V,et al. Parametric assessment of trend test power in a changing environment[J]. Sustainability,2020,12(9):3889.
- [11] BECKEL C,KLEIMINGER W,CICCHETTI R,et al. The ECO data set and the performance of non-intrusive load monitoring algorithms[C]//Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings,2014:80-89.
- [12] 史帅彬,张恒,邓世聪,等. 基于复合滑动窗的CUSUM暂态事件检测算法[J]. 电测与仪表,2019,56(17):13-18.
SHI S B,ZHANG H,DENG S C,et al. CUSUM Transient event detection algorithm based on compound sliding window[J]. Electrical Measurement & Instrumentation,2019,56(17):13-18.
- [13] 田甜,黄强,郭爱军,等. 基于标准化降水蒸散指数的渭河流域干旱演变特征分析[J]. 水力发电学报,2016,35(2):16-27.
TIAN T,HUANG Q,GUO A J,et al. Drought evolution characteristics in wei river basin based on standardized precipitation evapotranspiration index[J]. Journal of Hydro-

electric Engineering, 2016, 35(2): 16-27.

- [14] 孙毅,李昊洋,刘耀先,等. 基于改进隐马尔科夫模型的非侵入式家居负荷识别[J]. 电力建设, 2020, 41(4): 73-80.

SUN Y, LI H Y, LIU Y X, et al. Non-intrusive home-load identification based on improved hidden Markov model[J]. Electric Power Construction, 2020, 41(4): 73-80.

- [15] BOMFILI R, PRINCIPI E, FAGIANI M, et al. Non-intrusive load monitoring by using active and reactive power in additive factorial hidden Markov models[J]. Applied Energy, 2017, 208: 1590-1607.

- [16] LIN Y H, TSAI M S. Non-intrusive load monitoring by novel neuro-fuzzy classification considering uncertainties[J]. IEEE Transactions on Smart Grid, 2014, 5(5): 2376-2384.



第一作者:傅军栋(1972—),男,副教授,硕士,硕士研究生导师,研究方向为电力系统、智能建筑及其自动化。Email: 8755915@qq.com。

(责任编辑:姜红贵)