

# 基于简化HMM和时间分段的非侵入式负荷分解算法

刘凯<sup>1</sup>,符玲<sup>2</sup>,杨金刚<sup>1</sup>,熊思宇<sup>2</sup>,蒿保龙<sup>2</sup>,刘丽娜<sup>3</sup>

(1. 西南交通大学 唐山研究院,河北 唐山 063000;

2. 西南交通大学 电气工程学院,四川 成都 611756;3. 国网四川省电力公司计量中心,四川 成都 610045)

**摘要:**针对现有非侵入式负荷分解算法需要以过去时刻的分解结果为依据,从而造成误差累积的问题,提出一种基于简化的隐马尔可夫模型和时间分段的非侵入式负荷分解算法,以实现居民家庭的负荷分解。对负荷的低频功率信号进行分层抽样和聚类分析,构建负荷功率模板并利用独热码对超状态进行编码表示。基于简化的隐马尔可夫模型和普遍生活规律对家庭用电时间段进行划分,在每个时间段内单独训练参数。结合总线数据和各时间段参数实现对各时刻负荷功率的独立求解。基于2种国外公开数据集的测试结果验证了所提算法的准确性和实时性。

**关键词:**负荷分解;隐马尔可夫模型;亲和力传播聚类;时间分段;超状态

**中图分类号:**TM715

**文献标志码:**A

**DOI:**10.16081/j.epae.202304010

## 0 引言

以高级量测体系(advanced metering infrastructure, AMI)等技术作为支撑,家庭用户与供电公司之间的关系由原来的家庭用户单方向用电发展为家庭用户与供电公司双向灵活互动。负荷监测是AMI的一种重要功能,可为供电公司提供详细的用户用电数据信息,同时也可使用户更清晰地了解自己的用电行为,从而帮助用户/供电公司选择更节能高效的用电/供电策略。

负荷监测按照各种设备级负荷数据的获取方式分为侵入式负荷监测(intrusive load monitoring, ILM)和非侵入式负荷监测(non-intrusive load monitoring, NILM)。相比于直接获取负荷数据的ILM, NILM更依赖于算法,通过分解在用户电力入口端测量的集总数据得到用户家庭内部各种设备的用电详情<sup>[1]</sup>。NILM由于具有更好的经济性和隐私保护性,受到了大量学者的关注<sup>[2]</sup>。NILM的求解需要以负荷特征作为支撑,根据数据的不同采样频率,可以将负荷特征分为高频(频率大于1 Hz)特征和低频(频率不大于1 Hz)特征<sup>[3]</sup>。高频特征主要包括负荷的电流谐波、电压-电流轨迹、高频暂态波型等,低频特征包括负荷的有功功率、无功功率、电流有效值等。虽然高频特征包含更多的负荷信息,但是现有AMI中的智能电表一般仅提供低频数据的采集支持<sup>[4]</sup>,要获得高频特征就需要额外添加采集设备,这会增加经济和人工成本,因此,利用智能电表提供的低频特征数

据来实现居民家庭的负荷分解,是一种具有实际应用价值的方案<sup>[5]</sup>。

负荷分解算法可以分为2类。一类是组合优化,如:文献[6]提出利用时间概率因子来改进目标函数,再利用遗传优化算法进行负荷分解,该算法的准确率高但求解时间长;文献[7]结合改进的交叉熵算法和惩罚函数,通过迭代更新概率求取最优解并将其作为分解结果。另一类是模式识别,如:文献[8]利用各种电器间的关联规则,基于信息熵和 $k$ 近邻算法实现状态辨识,但仅考虑两两负荷之间的关联;文献[9]提出利用模体挖掘的方式,结合事件检测与调和函数进行事件匹配和设备识别,但算法依赖于运行窗的选取。近年来,随着神经网络的发展,利用神经网络来实现负荷分解也受到了研究人员的关注,如:文献[10]提出利用深度神经网络,综合电器的时间和功率信息进行负荷分解,但是模型参数的训练需要大量的历史数据和时间。在负荷分解的求解算法中经常采用隐马尔可夫模型(hidden Markov model, HMM),如:文献[11]为每种负荷均建立一条HMM链,再结合二次整数规划进行求解,通过假设初始分布概率相等简化了部分运算,但是算法的计算复杂度较高;文献[12]为一天中的每个小时单独建立HMM,再结合行为影响因子来实现负荷辨识。总体而言,现有的低频负荷分解算法模型复杂,求解时间长,在部署监测设备方面压力较大。此外,大部分算法需要将前一个或多个时刻的分解结果作为当前时刻分解的依据,若前面的分解结果出错,则会造成误差累积。

基于以上分析,本文提出一种利用有功功率,并在每个时刻独立求解的负荷分解算法。首先,在训练阶段,利用改进的亲和力传播聚类算法自适应地获取负荷功率模板,并将数据按照时间分段;然后,

收稿日期:2022-08-02;修回日期:2023-01-19

在线出版日期:2023-04-19

基金项目:四川省科技计划项目(2021YFG0294)

Project supported by the Science and Technology Planning Project of Sichuan Province(2021YFG0294)

对HMM进行简化,构建负荷分解算法,并在不同的时间段内分别训练模型参数,在负荷分解阶段,结合时间信息实时求取用户负荷功率组合;最后,基于2个国外的低频公开数据集的测试结果表明,本文算法准确率高且所需的求解时间短。

## 1 负荷的超状态编码

### 1.1 负荷分解

假设居民家庭内部共包含 $M$ 个负荷,任意负荷 $i$ 的有功功率 $P_i \geq 0$ ,则在 $t$ 时刻用户总的有功功率可以表示为:

$$P_t^{\text{all}} = \sum_{i=1}^M P_t^i + P_t^{\text{noise}} \quad (1)$$

式中: $P_t^{\text{all}}$ 为 $t$ 时刻用户总的有功功率; $P_t^i$ 为 $t$ 时刻负荷 $i$ 的有功功率; $P_t^{\text{noise}}$ 为 $t$ 时刻的背景噪声功率。假设负荷 $i$ 存在 $N_i$ 种状态,则式(1)可以改写为:

$$P_t^{\text{all}} = \sum_{i=1}^M \sum_{n=1}^{N_i} s_t^{i,n} p^{i,n} + P_t^{\text{noise}} \quad (2)$$

式中: $s_t^{i,n}$ 为 $t$ 时刻负荷 $i$ 是否运行在状态 $n$ 的标志,若运行在状态 $n$ ,则其值为1,否则为0; $p^{i,n}$ 为负荷 $i$ 在状态 $n$ 下的功率模板。通常,在正常工作情况下,负荷功率模板不会发生变化,因此,在确定负荷功率模板后,负荷功率分解问题就可以表示为 $t$ 时刻所有负荷状态的求解问题。

### 1.2 负荷状态编码

可以将居民家庭中的负荷按照运行时可能出现的状态分为只存在开关状态的I型负荷、具有有限运行状态数的II型负荷、状态可以连续变化的III型负荷以及一直处于某一功率运行的IV型负荷<sup>[13]</sup>。当负荷为I型负荷时,以“0”表示关闭状态,以“1”表示开启状态;当负荷为II型负荷时,以有限位数的“0”“1”组成的向量表示负荷状态;当负荷为III型负荷时,负荷功率特征在有限的范围内连续变化,一般通过聚类的方式对负荷状态进行离散化,从而将其转化为有限种状态<sup>[8]</sup>;当负荷为IV型负荷时,可以将其等效为缺少“0”状态的前3种负荷。由于通过聚类可将所有负荷均等效为有限状态机,因此,采用独热码来表示各种负荷状态<sup>[14]</sup>。对于拥有 $N_i$ 种状态的负荷,其独热码位数为 $N_i$ ,且对于任意给定的状态,独热码向量的各元素中仅有1个元素置1。若将整个家庭所有负荷的独热码按顺序拼接,则可以将家庭中所有的负荷状态表示为1个超状态向量 $S^{[15]}$ ,如式(3)所示。

$$S = [s_1, s_2, \dots, s_K] \quad (3)$$

式中: $s_k (k=1, 2, \dots, K)$ 的取值为0或1, $K$ 为所有负荷的运行状态总数。例如,假设用户家庭中只有电灯和洗衣机,其中电灯为I型负荷,只有开关状态,洗衣机为II型负荷,有浸泡、洗涤、脱水3种开启状

态,则可以在超状态向量中用第1个元素表示电灯的状态,第2个元素表示洗衣机的浸泡状态,第3个元素表示洗衣机的洗涤状态,第4个元素表示洗衣机的脱水状态,若 $S=[0, 1, 0, 0]$ ,则表示电灯处于关闭状态,洗衣机处于浸泡状态。

## 2 负荷功率模板的获取

获取负荷功率模板是进行负荷分解的重要前提,考虑到每种负荷的运行状态数未知, $K$ -means和模糊 $C$ 均值聚类算法需提前确定聚类中心数<sup>[16]</sup>,而亲和力传播(affinity propagation, AP)算法不需要提前确定最终的聚类中心数,因此,本文利用AP算法来获取负荷功率模板。

### 2.1 聚类数据的预处理

AP算法需要人为调整的参数少,降低了人为因素对结果的干扰<sup>[17]</sup>。AP算法的具体计算流程参考文献<sup>[17]</sup>,该算法主要包括相似度矩阵 $E$ 、归属度矩阵 $A$ 和吸引度矩阵 $R$ 这3个参数矩阵。算法所需时间和计算结果受到样本点数的影响:当聚类的样本点过少时,聚类结果会不准确;当聚类的样本点过多时,算法的计算时间会增加。因此,本文提出对大量的样本进行分层抽样,在满足较长时间跨度的同时,减少样本点。

本文采用的分层抽样规则如下:

1)由于实际运行时负荷处于关闭状态的时间远长于处于开启状态的时间,同时考虑到噪声干扰,先剔除负荷功率小于10 W的部分;

2)将10 W与最大功率之间的部分等分为5层,在每层随机抽取 $\alpha$ 个样本点,对于样本点数不足的层,保留该层的全部样本点;

3)将样本点数小于5的层中所有数据剔除,消去异常样本点的干扰;

4)对剩余的样本点进行聚类。

通过分层抽样可将尽可能多地反映负荷运行状态的数据作为聚类依据,同时剔除了异常点,减少了用于聚类算法的样本点。

### 2.2 AP算法的改进

在传统AP算法中,当归属度矩阵 $A$ 和吸引度矩阵 $R$ 的计算结果在一定的迭代次数内保持不变或者迭代达到设定次数时,停止迭代,并且聚类中心 $c$ 由式(4)确定。

$$R(c, c) + A(c, c) > 0 \quad (4)$$

在确定各点所属的簇中心时,会出现由式(4)得到节点 $c_i$ 和节点 $c_j$ 均为聚类中心,但节点 $c_j$ 所属簇中心的计算结果却为节点 $c_i$ 的情况。传统AP算法直接将节点 $c_j$ 所属簇中心强制确定为节点 $c_i$ ,这会造成2个相近的点被划分至2个簇的情况。为了避免产生相近的簇中心,本文将节点 $c_j$ 和归属于节点 $c_j$ 的所

有点都重新划归为属于节点  $c_i$ 。

此外,最终的聚类中心数受到参考度取值的影响。在得到聚类中心后,可以计算各点到其聚类中心的距离之和  $D$ ,  $D$  的大小能够反映聚类中心数的合理性。聚类中心过多的极限情况是每个点均为聚类中心,此时  $D$  为 0; 聚类中心过少的极限情况是聚类中心只有 1 个,此时  $D$  的最大值为样本点到其余样本点距离之和的最大值,即相似度矩阵中绝对值最大的列元素之和。为了使不同分布状况的负荷均能获得适宜的簇中心,求取  $D$  与相似度矩阵中列元素之和最小值的比值  $\gamma$ ,再限定该比值的范围,并以该范围为依据,采用迭代更新的方式自适应选择每种负荷的参考度,具体方法如下。

1) 由于参考度绝对值越接近 0, 最终的聚类中心数越多, 因此, 初始参考度选取为所有样本点两两之间距离中值的 1/4, 通过聚类得到各点所属的簇, 并计算  $D$ 。

2) 将  $\gamma$  与阈值进行比较。通过实验, 本文选取阈值为 0.01 和 0.1, 即当  $\gamma$  小于 0.01 时, 减小参考度, 当  $\gamma$  大于 0.1 时, 增大参考度, 并重新聚类获取聚类结果, 直到满足范围要求。

最终, 负荷功率模板的获取过程包括聚类前的样本点分层抽样、聚类时相近簇中心的剔除以及参考度的自适应迭代更新, 具体如附录 A 图 A1 所示。在通过聚类得到各负荷功率模板后, 可以计算超状态向量  $S$  的元素数, 再将  $S$  与负荷功率模板  $P_{tem}$  位置一一对应。式(2)可以表示为:

$$P_t^{all} = S_t P_{tem} + P_t^{noise} \quad (5)$$

式中:  $S_t$  为  $t$  时刻的超状态向量。

### 3 HMM 的简化与分时段负荷的分解

#### 3.1 HMM 的简化

HMM 主要包括负荷的初始概率分布矩阵  $\pi$ 、负荷的状态转移概率矩阵  $Z$  以及表示负荷状态与功率关系的发射矩阵  $B$  这 3 个参数矩阵。基于 HMM 的负荷分解算法满足马尔可夫链的齐次假设, 即马尔可夫链的一步转移概率与马尔可夫链所处时间无关。但是, 考虑到用户的用电习惯, 各种负荷在不同时间段内运行的概率不同, 这导致  $\pi$ 、 $Z$ 、 $B$  也不同。此外, HMM 需要根据前一时刻的负荷状态分解结果, 结合状态转移概率矩阵  $Z$  来求解当前时刻的状态, 若前一时刻负荷的状态识别结果与真实情况不符, 则会影响后续的负荷分解结果。

本文将 HMM 进行简化: ① 不考虑状态转移概率矩阵, 只保留  $\pi$  和  $B$ , 从而在算法求解时减少运算量, 缩减计算时间; ② 在参数学习阶段, 通过用户家庭的历史用电数据减少超状态种类数, 并通过时间段进一步限制在不同时间段内可能出现的超状态

种类数, 同时, 分别在不同时间段通过训练获取  $\pi$  和  $B$ , 从而在这 2 个矩阵中保留各负荷在不同时间段的运行差异信息。

#### 3.2 时间段的划分

按照一般用户家庭生活习惯, 可以将一天划分为睡觉时间段、与饮食相关的时间段、晚饭后的休闲娱乐时间段以及上述各时间段之间的过渡时间段。本文将一天划分为 11 个时间段, 其中 01:00—05:00 为第 1 个时间段, 将其余时间段按照 2 h 为 1 个时间段进行划分, 如附录 A 图 A2 所示。

#### 3.3 超状态参数的获取

根据负荷功率模板, 将训练集中各负荷功率序列标记为负荷状态, 将训练集数据按照时间段划分为各子数据集, 在每个时间段, 统计超状态的种类数, 计为  $J_\lambda$ 。初始概率分布矩阵  $\pi_\lambda$  可以表示为:

$$\pi_\lambda = \left[ \frac{V_{s_1}}{V_\lambda}, \frac{V_{s_2}}{V_\lambda}, \dots, \frac{V_{s_{J_\lambda}}}{V_\lambda} \right] \quad (6)$$

式中:  $V_\lambda (\lambda=1, 2, \dots, 11)$  为第  $\lambda$  个时间段内的样本总数;  $V_{s_j} (j=1, 2, \dots, J_\lambda)$  为超状态向量  $S_j$  下的样本点数。

在每个时间段计算超状态的发射矩阵  $B_\lambda \in \mathbf{R}^{J_\lambda \times O_\lambda}$ ,  $O_\lambda = P_{\max}^\lambda - P_{\min}^\lambda + 1$  为在第  $\lambda$  个时间段总功率的取值区间,  $P_{\max}^\lambda$ 、 $P_{\min}^\lambda$  分别为第  $\lambda$  个时间段总功率的最大值和最小值。为了减小发射矩阵的维数, 将功率值取整。发射矩阵的元素  $b_{jo}$  表示当用户家庭整体状态构成超状态向量  $S_j$  时功率观测值为  $o$  的概率, 计算公式为:

$$b_{jo} = H_{jo} / H_j \quad (7)$$

式中:  $H_{jo}$  为超状态向量  $S_j$  下功率观测值为  $o$  的频数;  $H_j$  为超状态向量  $S_j$  下的总频数。

#### 3.4 负荷功率分解

算法的输入为总功率的观测值  $o_t$  及其对应的时间信息。基于分时段初始概率分布矩阵和发射矩阵的功率求解步骤如下。

1) 利用输入的时间信息选择不同时间段的初始概率分布矩阵  $\pi_\lambda$  和发射矩阵  $B_\lambda$ 。

2) 由  $o_t$  确定发射矩阵中的列,  $t$  时刻估计的超状态矩阵  $\hat{S}_t$  为:

$$\hat{S}_t = \underset{j \in \{1, 2, \dots, J_\lambda\}}{\operatorname{argmax}} \left[ \pi_\lambda B_\lambda(j, o_t - P_{\min}^\lambda + 1) \right] \quad (8)$$

式中:  $\operatorname{argmax}$  为求参函数。由于  $B_\lambda$  中每行对应一种超状态, 因此, 式(8)实际是求取得概率值最大的行数, 从而确定对应的超状态。

3) 由超状态和负荷功率模板矩阵解码求得各负荷在  $t$  时刻的功率值。

### 4 算例分析

#### 4.1 算法评估指标

负荷识别结果准确率指标主要包括负荷状态识

别准确率指标和负荷功率分解准确率指标。

为了便于与现有方法进行对比,本文选择文献[18]定义的有限状态识别准确率指标  $F_1$ , 计算公式为:

$$F_1 = 2 \frac{pr}{p+r} \quad (9)$$

$$p = \frac{T_p - T_{inacc}}{T_p + F_p} \quad (10)$$

$$r = \frac{T_p - T_{inacc}}{T_p + F_n} \quad (11)$$

式中: $p$ 为识别为开启状态样本的分解结果中正确结果的比例; $r$ 为真实开启状态样本的分解结果中正确结果的比例; $T_p$ 为正确识别负荷处于开启状态的次数; $F_p$ 为负荷实际处于关闭状态但识别结果为开启状态的次数; $F_n$ 为负荷实际处于开启状态但识别结果为关闭状态的次数; $T_{inacc}$ 为附加惩罚项,引入该变量是由于对于有多种开启状态的负荷,存在虽然识别结果和实际都是开启状态但识别的具体开启状态与实际不符的情况,其计算公式如式(12)所示。

$$T_{inacc} = \sum_{i=1}^T \frac{|\hat{s}_{t,i} - s_{t,i}|}{N_i} \quad (12)$$

式中: $T$ 为测试集的总时长包含的时刻数; $\hat{s}_{t,i}$ 、 $s_{t,i}$ 分别为 $t$ 时刻负荷 $i$ 的估计状态和真实状态,当估计状态和真实状态不同时,两者差值为1。

负荷功率分解准确率  $a_{power}$  也常用于评估负荷识别的效果<sup>[19]</sup>, 计算公式为:

$$a_{power} = 1 - \frac{\sum_{i=1}^M \sum_{t=1}^T |\hat{P}_{t,i} - P_{t,i}|}{\sum_{i=1}^M \sum_{t=1}^T P_{t,i}} \quad (13)$$

式中: $\hat{P}_{t,i}$ 、 $P_{t,i}$ 分别为 $t$ 时刻负荷 $i$ 的估计功率和真实功率。

## 4.2 分解结果

为了测试算法效果,本文选取2个由真实家庭采集的低频数据集 REDD (reference energy disaggregation dataset)<sup>[20]</sup> 和 AMPds (almanac of minutely power dataset)<sup>[21]</sup>, 其中 REDD 来自美国波士顿地区的多个家庭,约每3 s进行1次负荷功率数据的采集和储存。AMPds 记录了温哥华地区1户居民2 a的功率数据,每1 min进行1次数据记录。同时,这2个数据集均以时间戳为时标对每个数据进行标注,结合地理时区划分能够得到每个数据在当地的具体时间,方便进行时间段的划分。

### 1) 在 REDD 上的测试。

将文献[11]基于二次规划的因子 HMM 算法和文献[12]利用行为影响因子的 HMM 算法与本文算法进行对比,选取相同的家庭 1、2、3、6 进行测试。

为了获取负荷在各时间段内的参数以及功率模板,将前2周的数据作为训练集,将其余数据作为测试集,负荷功率分解准确率对比如表1所示。

表1 负荷功率分解准确率对比

Table 1 Comparison of load power decomposition accuracy rate

| 家庭 | $a_{power} / \%$ |          |      |
|----|------------------|----------|------|
|    | 文献[11]算法         | 文献[12]算法 | 本文算法 |
| 1  | 78.4             | 94.4     | 94.6 |
| 2  | 86.4             | 96.4     | 95.2 |
| 3  | 83.5             | 94.4     | 94.1 |
| 6  | 91.6             | 92.1     | 95.8 |

由表1可知,本文算法的负荷功率分解准确率均高于文献[11]算法,本文算法对家庭1和家庭6的负荷功率分解准确率高于文献[12]算法,对家庭2和家庭3的负荷功率分解准确率与文献[12]算法相当。

为了便于验证本文算法对单个负荷的识别准确率,选择与文献[7]中场景5相同的实验设置,并与该文献中的改进交叉熵(modified cross-entropy, MCE)算法和分段整数约束规划(segmented integer quadratic constraint programming, SIQCP)算法进行对比,分解结果如附录A表A1所示。由表可知:本文算法对大多负荷的识别效果均更具优势,且整体识别效果优于文献[7]算法;虽然本文算法对个别负荷(如烤箱)的识别效果与文献[7]算法差距很大,但是整体识别效果和  $F_1$  是将整个测试阶段的所有负荷分解结果作为整体进行计算的,烤箱等负荷使用较少,对整体结果的影响很小。

### 2) 在 AMPds 上的测试。

相较于 REDD, AMPds 的时间跨度更大,采样频率更低,与实际的智能电表数据更接近。选择文献[7]中的场景3对 MCE 算法和本文算法进行对比,同样选择600 h的测试数据,选择相同的7种运行频次较高的负荷,即地下室、干衣机、洗碗机、暖通空调、冰箱、热泵、壁炉。用于训练的数据时长为40 d,识别结果准确率如表2所示。

表2 各负荷识别结果准确率

Table 2 Accuracy rate of identification results for each load

| 负荷   | MCE算法            |            | 本文算法             |            |
|------|------------------|------------|------------------|------------|
|      | $a_{power} / \%$ | $F_1 / \%$ | $a_{power} / \%$ | $F_1 / \%$ |
| 地下室  | 81.1             | 86.5       | 83.1             | 93.8       |
| 干衣机  | 76.5             | 26.2       | 95.8             | 75.0       |
| 洗碗机  | 67.1             | 75.3       | 69.8             | 62.4       |
| 暖通空调 | 69.8             | 58.3       | 96.4             | 99.9       |
| 冰箱   | 16.5             | 50.7       | 79.0             | 89.1       |
| 热泵   | 88.5             | 96.1       | 94.2             | 99.8       |
| 壁炉   | 71.8             | 72.0       | 81.4             | 50.0       |
| 全部负荷 | 73.9             | 89.8       | 89.5             | 94.2       |

由表2可知,本文算法总的负荷功率分解准确率达到89.5%,且有限状态识别准确率达到94.2%,比MCE算法的计算结果有了较大幅度的提高,这表明本文算法可得到较好的负荷分解结果,其中,暖通空调主要是在夜晚使用,以保持室内温度舒适,具有较强的时间段分布特性,因此本文算法对暖通空调的识别准确率提升幅度较大。

为了验证本文算法在较长时间下的分解准确率以及算法训练集时长对准确率的影响,设计包含8种负荷(在前文7种负荷的基础上增加电视机)的4种情景:场景1—4分别利用20、40、60、80 d的历史数据训练参数,再对360 d的数据进行分解,负荷识别结果准确率如表3所示。

表3 不同时长负荷识别结果准确率

Table 3 Accuracy rate of load identification results for different time lengths

| 场景 | $a_{\text{power}} / \%$ | $F_1 / \%$ |
|----|-------------------------|------------|
| 1  | 85.4                    | 87.5       |
| 2  | 90.0                    | 90.7       |
| 3  | 88.5                    | 90.9       |
| 4  | 89.8                    | 90.3       |

由表3可知,当用于训练的数据时长达到40 d后,识别结果准确率趋于稳定,延长训练时长不会显著提高识别结果准确率,且当用于训练的数据时长为40 d时,负荷功率分解准确率能够达到90.0%,有限状态识别准确率能够达到90.7%。

将场景2下本文算法对各负荷的分解结果与真实值进行比较,结果如表4所示。由表可知:本文算法对电视机功率的分解值与真实值间有较大差别,达到3%,这主要是由于在AMPds中电视机的子数据集并非仅记录了电视机的功率数据,还包括个人录像机顶盒、音箱等设备的数据,这些数据构成一个合成的待机状态功率,在整个数据集中的第132天,待机状态功率由39 W降到19 W,并在之后保持不变,而场景2的负荷功率模板是通过数据集中前40 d的数据进行聚类而获取的,因此,无法描述该差别,从而给分解带来了误差;本文算法对其他负荷功率的分解值与真实值均较接近。

表4 场景2下各负荷消耗的功率

Table 4 Power consumption of each load under Scenario 2

| 负荷   | 功率真实值 / % | 功率分解值 / % |
|------|-----------|-----------|
| 地下室  | 8         | 8         |
| 干衣机  | 11        | 10        |
| 洗碗机  | 3         | 2         |
| 暖通空调 | 23        | 23        |
| 冰箱   | 10        | 11        |
| 热泵   | 36        | 34        |
| 电视机  | 8         | 11        |
| 壁炉   | 1         | 1         |

### 4.3 算法复杂度分析

本文算法不考虑HMM中的转移概率矩阵,且利用独热码对负荷状态进行编码缩短了算法所需时间。由于训练阶段可离线进行,且参数训练完成后可以离线调用,因此,不考虑训练阶段的计算时间。假设在负荷分解时各时间段内最大的超状态数为 $G$ ,在求解时输入观测值 $o_t$ 后,得到对应的发射矩阵中的列索引,再将发射矩阵的该列向量与初始概率矩阵进行按位乘法运算,即最多进行 $G$ 次乘法运算,因此,算法的最大复杂度为 $O(G)$ 。实际上,由于发射矩阵具有稀疏性,某一确定的观测值对应的超状态数远小于 $G$ ,因此,复杂度也小于 $O(G)$ 。本文的测试软件为MATLAB R2021b,硬件为搭载了Intel(R) Core(TM) i5-7300HQ、内存为8 GB的笔记本电脑。

在长时间数据集AMPds上进行测试,分别选取2、4、6、8种负荷,待分解的时长为1~12个月,每天的数据点数为1440。本文算法分解所需时间如附录A图A3所示。由图可知,本文算法可以在短时间内完成负荷分解。为了分析本文算法在每个测试数据点的平均分解时间,将其与文献[12]添加行为影响因子的HMM算法和文献[15]的扩展稀疏HMM算法进行对比,结果如表5所示。

表5 平均分解时间

Table 5 Average decomposition time

| 负荷数 / 种 | 平均分解时间 / ms |          |          |
|---------|-------------|----------|----------|
|         | 本文算法        | 文献[12]算法 | 文献[15]算法 |
| 2       | 0.01        | 0.21     | 0.03     |
| 4       | 0.02        | 0.24     | 0.12     |
| 6       | 0.04        | 0.27     | 0.34     |
| 8       | 0.07        | 0.29     | 0.51     |

由表5可知,本文算法相比文献[12]和文献[15]算法在平均分解时间上均具有明显优势。此外,为了分析本文所提简化HMM与传统HMM在准确率和求解时间上的差异,选择与表3中场景2相同的负荷种类及训练时长,并采用相同的负荷功率模板,采用MATLAB 2021b中集成的hmmestimate函数和hmmviterbi函数对模型进行训练和求解,测试结果如附录A图A4所示。由图可知,相较于传统HMM,简化HMM在长时间下的分解准确率更高,且在计算时间上具有显著优势。

## 5 结论

由于现有AMI的低频采集条件,利用低频有功功率进行居民家庭NILM更具可实施性。针对传统算法对于分解结果的依赖性,本文提出一种在各时刻独立求解的实时非侵入式负荷分解算法。本文算法在2种公开数据集上的负荷功率分解准确率和有

限状态识别准确率均不低于90%,并且算法的计算时间短,对含有8种负荷场景的平均分解时间仅为0.07 ms。

笔者后续将进一步考虑不同用户负荷类型以及用电行为的相似性,并进一步简化模型参数的获取过程,从而使得算法更具泛化性。

附录见本刊网络版(<http://www.epae.cn>)。

## 参考文献:

- [1] 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 Xiang, LI Linzhi, WU Hao, et al. A survey of the research on non-intrusive load monitoring and disaggregation[J]. Power System Technology, 2016, 40(10): 3108-3117.
- [2] HART G W. Nonintrusive appliance load monitoring[J]. Proceedings of the IEEE, 1992, 80(12): 1870-1891.
- [3] 邓晓平,张桂青,魏庆来,等. 非侵入式负荷监测综述[J]. 自动化学报,2022,48(3):644-663.  
DENG Xiaoping, ZHANG Guiqing, WEI Qinglai, et al. A survey on the non-intrusive load monitoring[J]. Acta Automatica Sinica, 2022, 48(3): 644-663.
- [4] KONG W C, DONG Z Y, HILL D J, et al. Improving non-intrusive load monitoring efficiency via a hybrid programming method[J]. IEEE Transactions on Industrial Informatics, 2016, 12(6): 2148-2157.
- [5] 郭红霞,陆进威,杨苹,等. 非侵入式负荷监测关键技术问题研究综述[J]. 电力自动化设备,2021,41(1):135-144.  
GUO Hongxia, LU Jinwei, YANG Ping, et al. Review on key techniques of non-intrusive load monitoring[J]. Electric Power Automation Equipment, 2021, 41(1): 135-144.
- [6] 崔亮节,孙毅,刘耀先,等. 考虑分时段状态行为的非侵入式负荷分解方法[J]. 电力系统自动化,2020,44(5):215-222,242.  
CUI Liangjie, SUN Yi, LIU Yaoxian, et al. Non-intrusive load disaggregation method considering time-phased state behavior[J]. Automation of Electric Power Systems, 2020, 44(5): 215-222, 242.
- [7] MACHLEV R, LEVRON Y, BECK Y. Modified cross-entropy method for classification of events in NILM systems[J]. IEEE Transactions on Smart Grid, 2019, 10(5): 4962-4973.
- [8] 徐伟枫,华锦修,余涛,等. 计及电器状态关联规则的非侵入式负荷分解[J]. 电力自动化设备,2020,40(4):197-203.  
XU Weifeng, HUA Jinxiu, YU Tao, et al. Non-intrusive load decomposition considering association rules of appliances' state[J]. Electric Power Automation Equipment, 2020, 40(4): 197-203.
- [9] 周勇军,吴元香,董智华,等. 基于模体挖掘与调和函数半监督学习的非侵入式负荷监测[J]. 电力自动化设备,2022,42(7):3-10.  
ZHOU Yongjun, WU Yuanxiang, DONG Zhihua, et al. Non-intrusive load monitoring based on motif mining and harmonic function based semi-supervised learning[J]. Electric Power Automation Equipment, 2022, 42(7): 3-10.
- [10] 燕续峰,翟少鹏,王治华,等. 深度神经网络在非侵入式负荷分解中的应用[J]. 电力系统自动化,2019,43(1):126-132,167.  
YAN Xufeng, ZHAI Shaopeng, WANG Zhihua, et al. Application of deep neural network in non-intrusive load disaggregation[J]. Automation of Electric Power Systems, 2019, 43(1): 126-132, 167.
- [11] KONG W C, DONG Z Y, MA J, et al. An extensible approach for non-intrusive load disaggregation with smart meter data[J]. IEEE Transactions on Smart Grid, 2018, 9(4): 3362-3372.
- [12] 蔡宇,董树锋,徐航,等. 基于行为影响因子的非侵入式负荷实时分解算法[J]. 电力自动化设备,2021,41(12):193-199.  
CAI Yu, DONG Shufeng, XU Hang, et al. Real-time disaggregation algorithm of nonintrusive load based on usage influencing factor[J]. Electric Power Automation Equipment, 2021, 41(12): 193-199.
- [13] ZOHA A, GLUHAK A, IMRAN M A, et al. Non-intrusive load monitoring approaches for disaggregated energy sensing: a survey[J]. Sensors, 2012, 12(12): 16838-16866.
- [14] 王巍,高德远. 有限状态机设计策略[J]. 计算机工程与应用, 1999, 35(7): 54-55, 73.  
WANG Wei, GAO Deyuan. Design strategy for finite state machine[J]. Computer Engineering and Applications, 1999, 35(7): 54-55, 73.
- [15] MAKONIN S, POPOWICH F, BAJIĆ I V, et al. Exploiting HMM sparsity to perform online real-time nonintrusive load monitoring[J]. IEEE Transactions on Smart Grid, 2016, 7(6): 2575-2585.
- [16] 徐青山,姜藕蝶,郑爱霞,等. 基于近邻传播聚类和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.  
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.
- [17] FREY B J, DUECK D. Clustering by passing messages between data points[J]. Science, 2007, 315(5814): 972-976.
- [18] MAKONIN S, POPOWICH F. Nonintrusive load monitoring (NILM) performance evaluation[J]. Energy Efficiency, 2015, 8(4): 809-814.
- [19] 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.
- [20] KOLTER J Z, JOHNSON M J. REDD: a public data set for energy disaggregation research[C]//Proceedings of Workshop on Data Mining Applications in Sustainability. San Diego, USA: ACM Press, 2011: 59-62.
- [21] MAKONIN S, POPOWICH F, BARTRAM L, et al. AMPds: a public dataset for load disaggregation and eco-feedback research[C]//2013 IEEE Electrical Power & Energy Conference. Halifax, NS, Canada: IEEE, 2013: 1-6.

## 作者简介:

刘凯(1998—),男,硕士研究生,主要研究方向为非侵入式负荷监测(E-mail: 780264693@qq.com);

符玲(1981—),女,副教授,博士;通信作者,主要研究方向为复杂电力系统动态监测、故障信号处理及负荷监测(E-mail: lingfu@swjtu.cn);

杨金刚(1997—),男,硕士研究生,主要研究方向为非侵入式负荷监测(E-mail: 2680391410@qq.com)。

(编辑 王锦秀)

(下转第210页 continued on page 210)

ZHANG Yunning, ZHOU Xiaomeng. Virtual synchronous generator control technology with fractional virtual inertia for grid-connected inverters[J]. Control and Decision, 2021, 36(2): 463-468.

式发电技术、虚拟同步机应用与控制技术(E-mail: shirl163@163.com);

杨桂华(1971—), 女, 副教授, 主要研究方向为电气自动化应用技术(E-mail: 954991219@qq.com);

王国斌(1996—), 男, 硕士研究生, 主要研究方向为分布式发电技术(E-mail: wang\_guob@163.com)。

作者简介:

石荣亮(1987—), 男, 副教授, 博士, 主要研究方向为分布

(编辑 陆丹)

## Grid-connected active power response strategy of energy storage VSG based on active power fractional differential correction

SHI Rongliang<sup>1,2,3</sup>, YANG Guihua<sup>1</sup>, WANG Guobin<sup>1</sup>, LAN Caihua<sup>1</sup>, HUANG Ji<sup>2</sup>, WANG Bin<sup>3</sup>

(1. College of Mechanical and Control Engineering, Guilin University of Technology, Guilin 541004, China;

2. Guangxi Special Equipment Inspection and Research Institute, Nanning 530200, China;

3. School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430081, China)

**Abstract:** In order to solve the problem that the grid-connected active power of energy storage virtual synchronous generator(VSG) is difficult to take into account the good dynamic characteristics and steady-state performance under the disturbances of active power command and power grid frequency, a grid-connected active power response optimization control strategy of active power fractional differential correction-based VSG(AFDC-VSG) is proposed. The grid-connected active power closed-loop small-signal models of typical VSG and AFDC-VSG control strategies are established respectively, and the corresponding parameter design processes are given. MATLAB/Simulink simulation software is used to compare and analyze the grid-connected active power response characteristics of energy storage VSG by using different control strategies under two types of disturbance, and the experimental platform of energy storage VSG grid-connected system is established. Simulative and experimental results jointly verify the effectiveness and superiority of the proposed AFDC-VSG strategy in optimizing the grid-connected active power response performance for energy storage VSG.

**Key words:** virtual synchronous generator; energy storage; grid-connected active power; active power fractional differential correction; small-signal model; parameter design; response performance

(上接第203页 continued from page 203)

## Non-intrusive load decomposition algorithm based on simplified HMM and time segmentation

LIU Kai<sup>1</sup>, FU Ling<sup>2</sup>, YANG Jingang<sup>1</sup>, XIONG Siyu<sup>2</sup>, HAO Baolong<sup>2</sup>, LIU Lina<sup>3</sup>

(1. Tangshan Institute, Southwest Jiaotong University, Tangshan 063000, China;

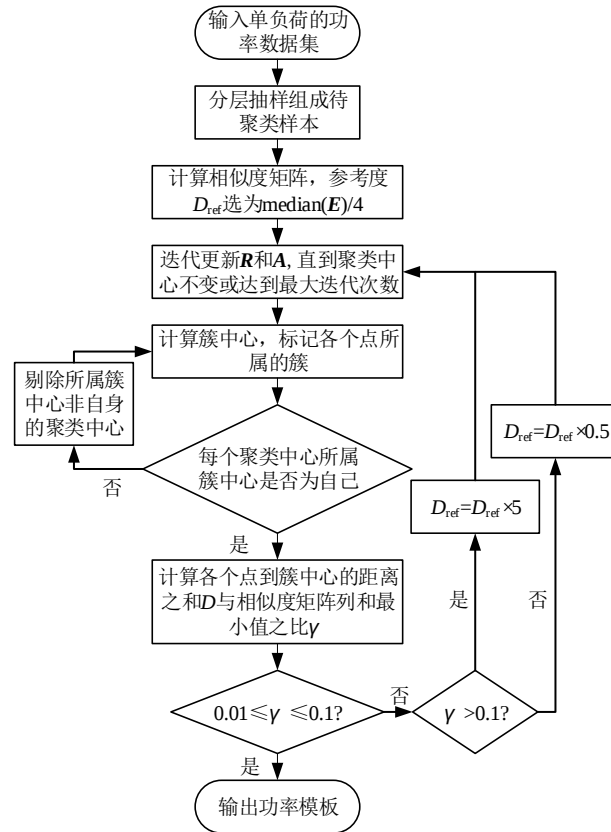
2. College of Electrical Engineering, Southwest Jiaotong University, Chengdu 611756, China;

3. Metering Center of State Grid Sichuan Electric Power Corporation, Chengdu 610045, China)

**Abstract:** Aiming at the problem that the current non-intrusive load decomposition algorithms need to be founded on the decomposition results of past moments, which causes the error accumulation, a non-intrusive load decomposition algorithm based on a simplified hidden Markov model(HMM) and time segmentation is proposed to realize the load decomposition of residential households. The stratified sampling and cluster analysis are carried out for the low-frequency power data of the load, the power template of the load is constructed and the super-state is coded and represented by the one-hot code. The time segments of household electricity consumption are divided based on the simplified HMM and universal life pattern, and the parameters are trained separately within each time segment. The load power at each moment is independently solved by combining the bus data and the parameters of each time segment. The accuracy and real-time performance of the proposed algorithm are verified by the test results based on two foreign public datasets.

**Key words:** load decomposition; hidden Markov model; affinity propagation clustering; time segmentation; super-state

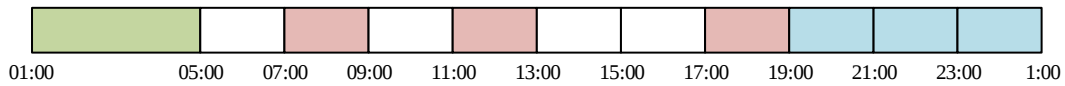
## 附录 A



注: median 表示求中位数。

图 A1 负荷功率模板获取流程图

Fig.A1 Flowchart of load power template acquisition



注: 绿色部分为睡觉时间, 人的活动较少, 负荷种类和变化也较简单; 红色部分为与饮食相关的时间段, 与烹饪相关的负荷运行较频繁; 蓝色部分为晚饭后的休闲娱乐时间, 电视等负荷更可能处于运行状态, 且考虑到不同家庭娱乐时间长短不一, 将 19:00 后按照 2 h 一段划分; 白色部分为过渡时间。

图 A2 时间分段

Fig.A2 Time segmentation

表 A1 各负荷分解准确率

Table A1 Accuracy of each load decomposition

| 负荷    | $p/r/\%$          |                   |                  |
|-------|-------------------|-------------------|------------------|
|       | MCE               | SIQCP             | 本文算法             |
| 烤箱    | 84.5/ <b>95.3</b> | <b>98.3</b> /8.0  | 11.2/16.6        |
| 冰箱    | 70.9/66.4         | 73.4/22.9         | <b>99.6/99.9</b> |
| 洗碗机   | 36.5/ <b>91.3</b> | <b>41.7</b> /79.2 | 36.5/55.8        |
| 电灯 1  | 36.3/38.2         | 55.6/87.8         | <b>89.9/94.5</b> |
| 微波炉   | <b>71.1/63.7</b>  | 28.7/21.6         | 42.3/27.3        |
| 卧室    | 17.0/50.8         | 40.3/29.6         | <b>55.5/85.1</b> |
| 厨房    | 0/0               | 15.2/20.6         | <b>43.1/70.5</b> |
| 电灯 2  | 69.7/ <b>84.0</b> | <b>94.9</b> /75.2 | 85.9/75.3        |
| 干衣机   | <b>0.99/91.9</b>  | 80.8/50.7         | 98.8/47.0        |
| 整体值   | 70.9/66.4         | 73.4/22.9         | <b>92.7/91.4</b> |
| $F_1$ | 61.9              | 50.3              | <b>92.0</b>      |

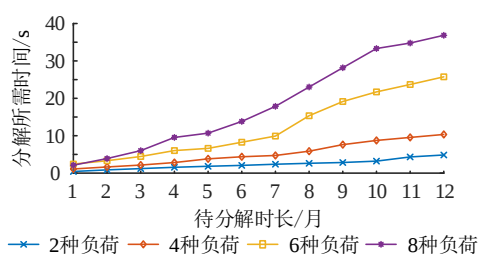


图 A3 所提方法求解时间

Fig.A3 Solution time of proposed method

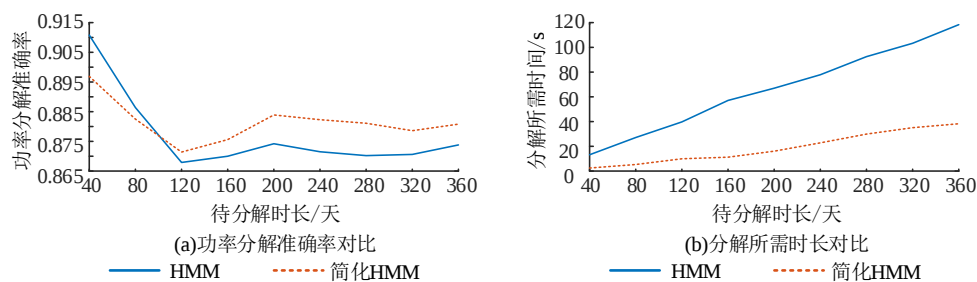


图 A4 简化 HMM 与传统 HMM 对比

Fig.A4 Comparison of simplified HMM and traditional HMM