

基于行为影响因子的非侵入式负荷实时分解算法

蔡宇¹,董树锋¹,徐航¹,毛航银²,宋永华³

(1. 浙江大学电气工程学院,浙江杭州 310027;2. 国网浙江省电力有限公司,浙江杭州 310007;
3. 澳门大学电机及电脑工程系,澳门特别行政区 999078)

摘要:针对传统非侵入式负荷分解算法准确率低、计算较耗时等问题,在隐马尔科夫模型(HMM)的基础上提出基于行为影响因子的负荷实时分解算法。使用自适应的迭代K-means聚类方法提取负荷状态,并将负荷状态组合成超状态。针对传统HMM没有考虑用电场景时间特性的缺陷,对参数进行时间分段学习。在分解阶段引入用户用电行为模式的影响因子,改进隐马尔科夫齐次假设,并利用维特比算法分解出用户的各个负荷的实时状态。通过公开数据集验证了所提算法的准确性和实时性。

关键词:负荷分解;超状态;隐马尔科夫模型;自适应K-means;行为模式

中图分类号:TM 73

文献标志码:A

DOI:10.16081/j.epae.202107005

0 引言

近年来,随着电力物联网、智能电网、智能电表等技术的迅速发展,电力系统中能够获取的数据越来越精细,如何进一步挖掘电力大数据的潜在价值成为全球的研究热点^[1]。负荷分解技术是智能电网技术之一,其不仅能够帮助用户自主调整用电行为,还能帮助售电公司了解用户的用电情况,科学制定需求响应激励政策,对推动资源优化配置和环境可持续发展具有重要意义^[2]。

非侵入式负荷分解最早由Hart G W^[2]等人在1992年提出,其目标是利用各时刻采集到的用户用电总功率或总电流等数据分解出各负荷的状态或功率。文献[3]详细阐述负荷特征提取、聚类分析、负荷建模、状态分解等流程。目前常见的分解算法包括鸡群算法^[4]、模糊聚类法^[5]、粒子群算法^[6]等,这些传统算法仅考虑负荷的电气特征,不仅耗时较长而且负荷分解的准确性较低。除了电气特征,负荷时间信息也非常关键,能体现用户的用电习惯,文献[7]提出的时间概率算法充分利用电器运行的时间特性,但忽略了相邻2个时刻之间的联系,对部分家庭用户的分解准确性较低。文献[8]使用基于序列到序列和Attention机制的深度学习框架,取得了较好的效果,但是人工智能算法具有一定的随机性,可能会收敛于局部最优,且对于不同的家庭数据均需要调整网络参数,需具备较强的调参经验。

非侵入式负荷分解是典型的时间序列分析问题,隐马尔科夫模型HMM(Hidden Markov Model)非常适合解决这类问题。文献[9]给出HMM应用于

负荷分解的一般步骤,但不同的负荷模型各自对应一条隐马尔科夫链,未考虑负荷之间的相互影响。文献[10]通过高阶隐马尔科夫模型FHMM(Factorial Hidden Markov Model)引入各负荷之间的相互影响,但模型在计算发射矩阵时使用吉布斯采样方法,虽然提高了计算效率,却降低了分解准确率。文献[11]对HMM进行简化,并采用分段二次整数约束规划进行近似求解,但是该方法会累积误差,进而降低模型的准确性。

综上,HMM框架应用于非侵入式负荷分解具有较好的效果,但仍存在一些不足。针对传统HMM未考虑时间特性这一缺陷,本文提出基于行为影响因子的负荷实时分解算法,该算法对参数按时间分段学习。此外,在状态分解阶段,对隐马尔科夫链的齐次假设进行修正,使得状态转移概率更符合实际用电场景。算例测试表明,本文算法准确率高且耗时短。

1 负荷状态与超状态

1.1 有限状态机

基于电力负荷的数据特征,可将电力负荷划分为有限状态型、开关型和连续型这3种类型。例如:有多个档位的电风扇属于有限状态型;只有开和关选择的电热壶属于开关型,开关型可看作是仅有2个状态的有限状态型;电机、锅炉等可连续改变运行功率的负荷则属于连续型。通常家用的连续型负荷较少^[12],且可通过离散化等方法将其转化为有限状态型,因此本文算法针对有限状态型负荷进行分析讨论。

1.2 自适应K-means聚类

分析有限状态型负荷,首先要通过负荷的功率序列建立有限状态机FSM(Finite-State Machine)。文献[13]采用概率质量函数PMF(Probability Mass Func-

收稿日期:2020-07-06;修回日期:2021-05-12

基金项目:国家电网公司科技项目(5211041800M)

Project supported by the Science and Technology Project of State Grid Corporation of China(5211041800M)

tion)进行聚类,每个负荷状态数等于其各自PMF曲线的波峰数,但当负荷在运行时出现2个或以上较为接近的功率数值时,PMF方法会将该状态识别为2种或以上状态,增加了负荷的状态数量,因此该方法在建模时需依据经验限制最大状态数。相较而言, K -means聚类方法具有更好的鲁棒性,文献[10]给出一种基于聚类中心数迭代的 K -means聚类方法,但其聚类中心数 K 很大程度上取决于阈值的选择,且阈值的选择需依赖一定经验。本文采用一种能自动选择聚类中心数的改进迭代 K -means方法。对 K 值进行迭代,在每一轮迭代中记录对应的聚类损失指标,得到“ K -损失指标”序列,考虑到当 K 值小于真实聚类数时, K 值的增大会大幅增加每个类的聚合度,损失指标会迅速下降,当 K 值超过真实聚类数时,损失指标下降趋缓,因此该序列中存在一个转折点,在转折点之前 K 值迅速下降,在转折点之后 K 值缓慢下降,该点即为“ K -损失指标”序列的二阶差分最大值处,该处对应的 K 值就是最佳的聚类中心数 K^* 。

图1为对REDD(the Reference Energy Disaggregation Data set)中家庭6的电热器负荷功率序列进行迭代聚类得到的“ K -损失指标”序列折线图,其中最佳聚类中心数 K^* 处于折线的转折点处。电热器的最终聚类结果如表1所示。

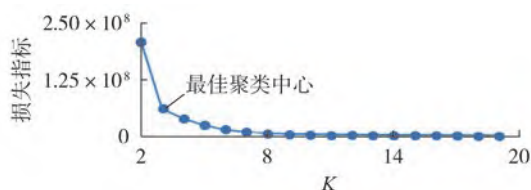


图1 电热器聚类过程的损失指标

Fig.1 Loss index of electric heater clustering process

表1 电热器FSM聚类结果

Table 1 FSM clustering results of electric heater

状态	功率均值 / W	功率标准差 / W
0	11.54	5.01
1	346.41	85.78
2	71.72	23.78

1.3 超状态

建立负荷FSM模型后,若直接建立HMM,则每个负荷都对应一条隐马尔科夫链,链中的状态转移并未考虑不同负荷之间的影响。事实上,负荷的状态总存在某些关联性^[14],如用户游戏机打开时,电视也常处于使用状态^[15];某些负荷的状态存在互斥性,如空调和电热毯同时使用的概率几乎为0。为了体现这些负荷间的影响,将每个时刻所有负荷的状态合并成一个状态,称为超状态^[9,13]。超状态是各个负荷状态的笛卡尔积,任意的状态组合都有一个超

状态与之对应,将已知各个负荷状态计算超状态的过程称为编码,反之称为解码。

设有 m 个负荷,状态数量分别为 $n_1—n_m$ 。根据组合原理,总共有 $N = \prod_{i=1}^m n_i$ 种不同组合情形,即超状态集合为 $\{s_1, s_2, \dots, s_N\}$,空间复杂度为 $O(2^N)$,但家庭用户用电场景有限,不可能遍历全部的超状态,因此 N 不是一个极大的数。REDD中家庭1的超状态理论上高达3600种,由该数据集的历史数据可知超状态总数量仅为133种。

2 基于超状态的HMM

2.1 算法流程

整个算法主要分为学习和分解2个阶段。首先将历史数据中每个时刻各负荷状态通过编码得到该时刻对应的超状态,进而得到超状态时间序列,再对该序列进行参数学习获得HMM。在分解阶段,输入当前时刻的功率读数,并利用维特比算法分解出当前时刻的超状态,该超状态通过解码过程便得到各个负荷对应的状态。算法整体流程如图2所示,图中 T 为总时刻数。

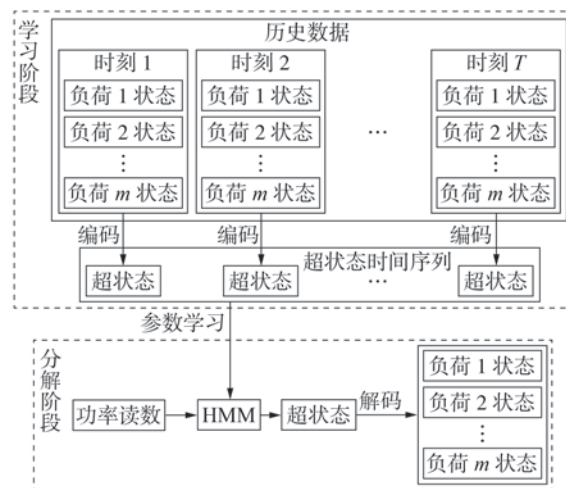


图2 基于HMM的负荷分解流程图

Fig.2 Flowchart of load disaggregation based on HMM

2.2 参数学习

常规的HMM未考虑时间节点对状态转移的影响,但在用电场景中时间节点会影响用电行为^[16],例如,在傍晚电视机通常会从关闭状态转移为开启状态。为此,需按照不同的时间节点来学习HMM参数。本文对一天24h进行HMM参数分段学习。

1) 状态转移矩阵。

$$\{A_0, A_1, \dots, A_{23}\} \quad (1)$$

$$a_{hjk} = \frac{c_{hjk}}{\sum_{s=1}^N c_{hjs}} \quad h=0, 1, \dots, 23; j, k=1, 2, \dots, N \quad (2)$$

式中:状态转移矩阵 $A_h = [a_{hjk}]$, a_{hjk} 为第 h 小时超状态 j 转移到超状态 k 的概率; c_{hjk} 为第 h 小时超状态 j 转移到超状态 k 的频数。状态转移矩阵 A_h 为 $N \times N$ 阶矩阵,该矩阵中存在大量的零值,可用稀疏矩阵技术节省内存空间。

2) 发射矩阵 B 为:

$$B = [b_{io}] \quad i \in \{s_1, s_2, \dots, s_N\} \quad (3)$$

$$b_{io} = \frac{d_{io}}{\sum_{y=1}^Y d_{iy}} \quad (4)$$

式中: b_{io} 为超状态为 i 时观测的功率值为 o 的概率; d_{io} 为超状态为 i 时观测的功率值为 o 的频数; Y 为所有可能出现的功率读数的数量。矩阵 B 同样是稀疏的。

3) 初始概率向量。

$$\{\pi_0, \pi_1, \dots, \pi_{23}\} \quad (5)$$

式中: $\pi_h = [P_{h,1}, P_{h,2}, \dots, P_{h,N}]$ ($h=0, 1, \dots, 23$), $P_{h,i}$ 为每天第 h 小时超状态为 i 的概率。

3 基于维特比算法的负荷分解

齐次假设是隐马尔科夫的基本假设之一,该假设指出任意时刻的状态只依赖于前一时刻的状态,与更早时刻的状态、观察值均无关,即在已知时刻 t 之前所有的状态值和观察值的情况下,时刻 t 的状态为 i_t 的概率公式为:

$$P(i_t | i_{t-1}, o_{t-1}, i_{t-2}, o_{t-2}, \dots, i_1, o_1) = P(i_t | i_{t-1}) \quad (6)$$

式中: o_t 为时刻 t 的输出功率。

设随机变量 d 为超状态 i 持续自转移的次数,根据齐次假设得到 d 的概率密度函数为:

$$P_i(d) = a_{ii}^d (1 - a_{ii}) \quad (7)$$

式中: $P_i(d)$ 为超状态 i 持续自转移 d 次的概率; a_{ii} 为超状态 i 在下一时刻依然保持的概率,即自转移概率。随机变量 d 为几何分布,这表明齐次假设下的 HMM 仅对无后效的过程成立^[17],因此,计算状态转移时并未考虑到上午的用电情况,由此可知用户的用电行为影响了状态转移的概率,需对此进行统计计算进而修正该偏差。文献[15]引入半隐马尔科夫模型 SHMM (Semi Hidden Markov Model) 对该缺陷进行修正,充分考虑各种电器的运行模式,但整个模型变得非常复杂,代码实现难度较大。

3.1 用户行为影响因子

根据经验,超状态的频率次数服从正态分布 $N(\mu, \sigma)$, 其中 μ 为均值, σ 为方差。令超状态统计次数 ϕ 为超状态在当天内已出现的次数,则超状态再次出现的概率为:

$$\psi_{i,i} = \int_{\phi}^{+\infty} f(x) dx = F(2\mu - \phi) \quad (8)$$

式中: $\psi_{i,i}$ 为行为影响因子,表示用户习惯对时刻 t 超状态 i 的影响系数; $f(\cdot)$ 为正态分布概率密度函数; $F(\cdot)$ 为概率密度累计函数。

将影响因子代入状态转移公式,提高状态转移的准确率,故将式(6)改为:

$$P(i_t | i_{t-1}, o_{t-1}, i_{t-2}, o_{t-2}, \dots, i_1, o_1) = P(i_t | i_{t-1}) \psi_{i_{t-1}, i_t} \quad (9)$$

3.2 超状态及工作功率的求解

维特比算法是一种动态规划算法,常用于求解已知参数和观测序列情况下的隐状态问题。定义局部状态 $\delta_t(i)$ 为时刻 t 超状态为 i 的所有可能状态转移路径中的概率最大值,如式(10)所示。

$$\delta_t(i) = \max_{s_1, s_2, \dots, s_{t-1}} P(s_t = i, s_1, s_2, \dots, s_{t-1}, o_1, o_2, \dots, o_t | \lambda) \quad (10)$$

式中: λ 为经过参数学习得到的 HMM。将式(9)代入 $\delta_t(i)$ 的递推式得到:

$$\delta_t(i) = \max_{j \in \{s_1, s_2, \dots, s_N\}} \delta_{t-1}(j) A_{t-1}(j, i) \psi_{j, i} B(i, o_t) \quad (11)$$

式中: $A_{t-1}(j, i)$ 为矩阵 A_{t-1} 的第 j 行第 i 列元素; $B(i, o_t)$ 为矩阵 B 的第 i 行第 o_t 列元素。

结合上述公式可得到超状态的求解步骤如下。

1) 取时刻 $t-1$ 对应的初始向量为初始时刻的概率向量,即:

$$\pi_{t-1} = [P_{s_1}, P_{s_2}, \dots, P_{s_N}] \quad (12)$$

2) 已知当前时刻 t 的功率读数 o_t 、上一时刻的功率读数 o_{t-1} 及当天直到时刻 $t-1$ 的超状态统计次数 ϕ_{t-1} , 有:

$$\psi_{i_{t-1}, i_t} = F(\mu - \phi_{t-1}) \quad (13)$$

$$\delta_{t-1}(i) = \pi_{t-1}(i) B(i, o_{t-1}) \quad (14)$$

将式(13)、(14)的结果代入式(11)得到 $\delta_t(i)$ 。

3) 根据式(15)计算当前时刻 t 的超状态 S_t^* 。

$$S_t^* = \operatorname{argmax}_{j \in \{s_1, s_2, \dots, s_N\}} \delta_t(j) \quad (15)$$

式中: argmax 为求参函数,表示因变量取最大值时自变量的值。根据 S_t^* 更新超状态统计次数 ϕ_t 。

再根据超状态和负荷状态的对应关系进行解码得到各个负荷的分解状态。若已知负荷 i 的状态,则根据其 FSM 可以得到状态功率 p_i , 理想情况下所有负荷的状态功率之和为总功率读数,即:

$$\sum_{i=1}^m p_i = o_t \quad (16)$$

负荷真实功率会在状态功率上下波动导致各负荷分解功率之和与总功率读数存在误差,为简便计算,将误差均分给各个负荷。首先计算波动系数 u :

$$u = \frac{o_t}{\sum_{i=1}^m p_i} \quad (17)$$

则各负荷的分解功率 $\hat{p}_i$ 为:

$$\hat{p}_i = up_i \quad (18)$$

功率分解依赖于状态分解,在状态分解较为准确时,功率分解也会准确。

3.3 算法复杂度分析

本文算法的耗时主要包括数据加载、参数学习和状态分解3个部分。前2个部分的时间复杂度取决于算法训练所用的数据长度以及参与分解的负荷数量。第3个部分的时间复杂度为 $O(N^2Z)$,其中 Z 为需要求解的隐状态数量,在进行实时负荷分解时 $Z=1$ 。因此本文算法平均每次分解需要的时间为 N^2 ,由前面的讨论可知在实际应用场景中 N 的增长较缓慢,因此本文算法有较高的时间效率。

表2对比了不同算法的时间复杂度。由表中可见:本文算法与Sparse HMM算法^[18]的时间复杂度相同;分段整数约束规划SIQCP(Segmented Integer Quadratic Constraint Programming)算法^[11]在每个数据点的平均耗时为 Q^m ,其中 Q 为每个负荷平均状态数量,在负荷较多时计算时间的数量级较大,很难保证实时性;文献[7]中的时间概率算法平均每个数据点耗时约为 D^w ,其中 D 为 S 超状态所包含的超状态数平均值, w 为 S 超状态短序列长度,当 w 变大时,计算时间会呈指数上升。

表2 时间复杂度对比

Table 2 Comparison of time complicity

算法	时间复杂度
本文算法	$O(N^2Z)$
Sparse HMM算法	$O(N^2Z)$
SIQCP算法	$O(K^mZ)$
时间概率算法	$O(D^w(Z-W+1))$

3.4 算法评估指标

1) 状态分解准确率 a_{state}

$$a_{\text{state}} = \frac{\sum_{t=1}^T \sum_{i=1}^m I(\hat{x}_{t,i} = x_{t,i})}{mT} \quad (19)$$

式中: $\hat{x}_{t,i}$ 、 $x_{t,i}$ 分别为时刻 t 负荷 i 的分解状态和真实状态; $I(\cdot)$ 为指示函数,当括号内的表达式为真时其

值为1,否则为0。状态分解准确率是关键的算法评估指标,用于判断算法对各时刻负荷工作状态的分解是否准确。

2) 功率分解准确率 a_{power}

$$a_{\text{power}} = 1 - \frac{\sum_{t=1}^T \sum_{i=1}^m |\hat{p}_{t,i} - p_{t,i}|}{2 \sum_{t=1}^T \sum_{i=1}^m |p_{t,i}|} \quad (20)$$

式中: $\hat{p}_{t,i}$ 、 $p_{t,i}$ 分别为时刻 t 负荷 i 的分解功率和真实功率。

4 算例分析

4.1 负荷FSM建模准确性对比

选取8个已知档位和功率的常见家用电器,模拟电器运行一段时间后得到各电器的功率序列。使用不同文献中的方法对功率序列进行聚类,对比各方法的聚类准确性:在测试文献[11]中的迭代 K -means聚类方法时选用3个不同的灵敏度阈值 e 作为对比;在测试文献[13]中的PMF方法时选用2个不同的限制参数 M ;在测试本文使用的改进迭代 K -means聚类方法时选用2个不同的最大迭代次数作为参数对照。

聚类结果如表3所示。由表可知,参数的选择对于迭代 K -means方法非常关键,参数过大会导致某些负荷的聚类数量不够,过小则出现过度聚类现象,因此,只有对每个负荷都分别按照经验设置参数才能获得100%的准确率。PMF方法的准确率同样取决于参数的选取。本文的改进迭代 K -means聚类方法只需达到一定迭代次数就能出现损失指标拐点,较大的迭代次数对结果影响不大,只会增加计算时间,一般负荷的状态数量都是个位数,迭代次数只要稍微大于10即可。对比结果证明,本文算法不需要选择灵敏度阈值 e ,且聚类效果更好。

4.2 分解算法结果

本文选取负荷分解中较为常用的REDD中的全部6个家庭进行测试。为了便于与其他算法对比,

表3 聚类结果

Table 3 Clustering results

负荷	真实状态数	不同方法和参数得到的状态数量						
		本文方法 ($Y=10$)	本文方法 ($Y=20$)	PMF方法 ($M=4$)	PMF方法 ($M=8$)	K -means方法 ($e=160$)	K -means方法 ($e=80$)	K -means方法 ($e=40$)
微波炉	5	5	5	4	4	5	5	5
电视	2	2	2	3	3	2	2	2
冰箱	3	3	3	2	5	3	3	3
电吹风	3	3	3	2	2	3	3	6
电热壶	2	2	2	2	2	3	5	7
吊扇	4	4	4	4	5	3	4	4
计算机	2	2	2	4	5	2	2	2
洗衣机	3	3	3	3	8	3	3	3
准确率 / %	—	100	100	37.5	12.5	75.0	87.5	75.0

家庭 1—6 的负荷数量分别为 6、7、8、7、8、7 个。模型学习与测试的长度比为 3:1。测试硬件为 Intel (R) Xeon(R) CPU E5-2630, 系统为 Centos 7.4。

本文算法、Sparse HMM 算法以及时间概率算法的负荷状态分解结果对比如表 4 所示。其中 Sparse HMM 算法在参数学习时进行的离散化处理导致部分数据缺失, 且该算法没有考虑负荷运行的时间特性, 而家庭 3 和家庭 6 用户在一天不同时间点的用电行为为差异较大, 因此该算法在这 2 个家庭的测试中表现较差。时间概率算法使用超状态短序列的时间概率最大似然估计求解最大可能超状态, 但忽略了前一时刻超状态对后一时刻超状态的影响, 而家庭 6 中超状态在时间上相互影响较强, 因此该算法出现较多的分解错误, 准确率较低。

表 4 负荷状态分解准确率对比
Table 4 Accuracy comparison of load state disaggregation

家庭	准确率 / %		
	本文算法	Sparse HMM 算法	时间概率算法
1	99.4	98.9	95.9
2	96.4	94.9	91.2
3	96.5	88.6	91.0
4	93.5	90.9	93.4
5	92.0	93.4	97.4
6	93.9	86.3	84.9
平均值 / %	95.3	92.2	92.3
最大值 / %	99.4	98.9	97.4
最小值 / %	92.0	86.3	84.9
标准差	2.7	4.5	4.4

表 5 展示了不同算法的功率分解准确率结果。从表中可看出, 本文算法、Sparse HMM 算法及 SIQCP 算法都有较高的功率分解准确率, 而时间概率算法由于在分解状态计算分解功率时使用凸二次规划模型求解, 不仅计算复杂且准确率不高。

表 5 负荷功率分解准确率对比
Table 5 Accuracy comparison of load power disaggregation

家庭	准确率 / %			
	本文算法	Sparse HMM 算法	SIQCP 算法	时间概率算法
1	94.4	89.9	78.4	83.4
2	96.4	87.0	86.4	70.1
3	94.4	95.8	83.5	75.4
4	88.7	81.3	—	76.5
5	87.7	91.3	—	90.7
6	92.1	94.8	91.6	82.1
平均值 / %	92.3	90.0	85.0	79.7
最大值 / %	96.4	95.8	91.6	90.7
最小值 / %	87.7	81.3	78.4	70.1
标准差	3.5	5.3	5.5	7.2

综上, 本文算法在负荷状态分解准确率以及负荷功率分解准确率上都有更好的表现。

4.3 算法耗时分析

为测试本文算法的时间效率, 用 REDD 中的家庭 5 和家庭 6 进行测试, 分别记录在 2—11 个不同数量负荷参与分解的情况下各个阶段的运行时间。

家庭 5 中用于参数学习的数据有 78 406 行, 测试数据有 2 011 行。图 3(a) 展示了本文算法在家庭 5 数据上的运行时间, 均不超过 25 s, 不同负荷数量情形下的运行时间相差无几。

家庭 6 中的学习数据有 367 543 行, 测试数据有 9 425 行。图 3(b) 展示了本文算法运行在家庭 6 数据上的运行时间, 可知随着负荷数量增加, 运行时间稍有增加, 但均在 100 s 内结束, 因此耗时在可以接受的范围内。

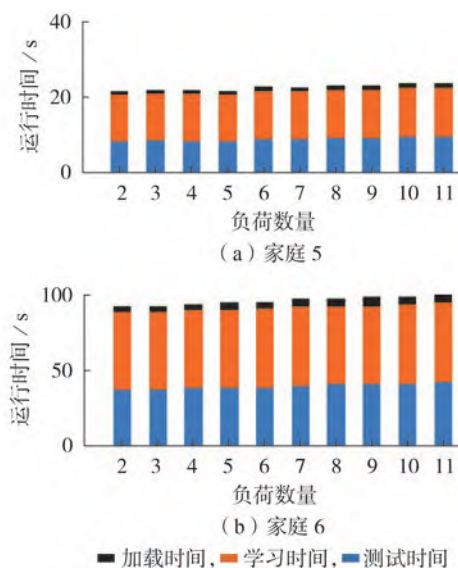


图 3 REDD 中家庭 5 和家庭 6 的运行时间

Fig.3 Running time of Family 5 and 6 in REDD

使用数据量更大(约 520 000 行)的 AMPds 数据集^[19]进行测试, 分别记录 2—18 个负荷参与负荷分解的运行时间, 如图 4 所示。可见, 本文算法的总体运行时间随着负荷数量的增加近似呈线性上升, 而计算量没有随着负荷数量的增加呈指数上升。

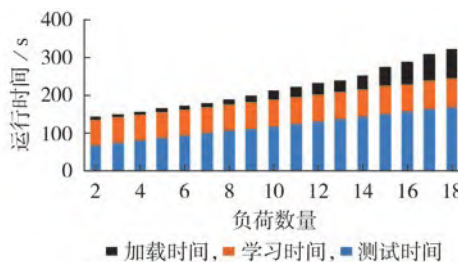


图 4 AMPds 数据集运行时间

Fig.4 Running time of AMPds data set

表 6 为不同负荷数量情况下每个测试数据点的平均分解时间, 均在 0.5 s 内完成分解。在不考虑数据传输速率等其他因素的情况下, 对于采样频率不

超过 1 kHz 的数据,本文算法都能在读数采样的周期内完成分解,实时性强。

表6 平均分解时间

Table 6 Average disaggregation time

负荷数量	平均分解时间 / ms	负荷数量	平均分解时间 / ms
2	0.21	7	0.28
3	0.23	8	0.29
4	0.24	9	0.30
5	0.25	10	0.31
6	0.27	11	0.33

5 结论

非侵入式负荷分解能够有效帮助用户以及电力公司进行配电侧的能量管理和规划。针对传统算法对参数设定要求高、分解准确率低、计算耗时较长等难点,本文在HMM的基础上,提出基于行为影响因子的负荷实时分解算法。

1)使用自适应的K-means对负荷进行FSM建模,不再依赖调参经验,且负荷状态聚类更加准确。

2)考虑用电场景的时间特性,对一天24 h进行分段参数学习,同时引入用户行为模式的影响因子,使得状态转移更加合理。

3)测试的结果表明,相较于传统算法,本文算法有更高的准确率和时间效率。

随着负荷数量的增加,用电行为越复杂,模型计算开销加大,且分解的准确性也有所下降。后续研究将寻找更快的计算方法,利用统一计算设备架构(CUDA)、高性能GPU等软硬件进一步提高计算速度。

参考文献:

- [1] 涂京,周明,宋旭帆,等. 基于监督学习的非侵入式负荷监测算法比较[J]. 电力自动化设备,2018,38(12):128-134.
TU Jing,ZHOU Ming,SONG Xufan,et al. Comparison of supervised learning-based non-intrusive load monitoring algorithms[J]. Electric Power Automation Equipment,2018,38(12):128-134.
- [2] 张潼,于鹤洋,田江,等. 基于非侵入式负荷辨识的聚合负荷需求响应能力在线评估[J]. 电力工程技术,2020,39(6):19-25,65.
ZHANG Tong,YU Heyang,TIAN Jiang,et al. Online aggregation monitoring of low-voltage power load demand response capability based on non-intrusive load identification[J]. Electric Power Engineering Technology,2020,39(6):19-25,65.
- [3] HART G W. Nonintrusive appliance load monitoring[J]. Proceedings of the IEEE,1992,80(12):1870-1891.
- [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 Xingjie,CAO Meihan,XU Yuejuan. Non-intrusive load monitoring based on improved chicken swarm optimization algorithm[J]. Electric Power Automation Equipment,2018,38(5):235-240.
- [5] 孙毅,崔灿,陆俊,等. 基于差量特征提取与模糊聚类的非侵入式负荷监测方法[J]. 电力系统自动化,2017,41(4):86-91.
SUN Yi,CUI Can,LU Jun,et al. Non-intrusive load monitoring method based on delta feature extraction and fuzzy clustering[J]. Automation of Electric Power Systems,2017,41(4):86-91.
- [6] 李如意,黄明山,周东国,等. 基于粒子群算法搜索的非侵入式电力负荷分解方法[J]. 电力系统保护与控制,2016,44(8):30-36.
LI Ruyi,HUANG Mingshan,ZHOU Dongguo,et al. Optimized nonintrusive load disaggregation method using particle swarm optimization algorithm[J]. Power System Protection and Control,2016,44(8):30-36.
- [7] 燕续峰,翟少鹏,何光宇. 基于时间概率分布与电器组合超状态的非侵入式负荷分解[J]. 电力系统自动化,2018,42(24):53-58,73.
YAN Xufeng,ZHAI Shaopeng,HE Guangyu. Non-intrusive load disaggregation based on time probability distribution and super state of appliances[J]. Automation of Electric Power Systems,2018,42(24):53-58,73.
- [8] 王轲,钟海旺,余南鹏,等. 基于seq2seq和Attention机制的居民用户非侵入式负荷分解[J]. 中国电机工程学报,2019,39(1):75-83,322.
WANG Ke,ZHONG Haiwang,YU Nanpeng,et al. Nonintrusive load monitoring based on sequence-to-sequence model with attention mechanism[J]. Proceedings of the CSEE,2019,39(1):75-83,322.
- [9] ZIA T,BRUCKNER D,ZAIDI A. A hidden Markov model based procedure for identifying household electric loads[C]//IECON 2011-37th Annual Conference of the IEEE Industrial Electronics Society. Melbourne,VIC,Australia:IEEE,2011:3218-3223.
- [10] KOLTER J Z,JOHNSON M J. REDD:a public data set for energy disaggregation research[C]//In Workshop on Data Mining Applications in Sustainability. San Diego,USA:[s.n.],2011:59-62.
- [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] 陈思运,高峰,刘烜,等. 基于因子隐马尔可夫模型的负荷分解方法及灵敏度分析[J]. 电力系统自动化,2016,40(21):128-136.
CHEN Siyun,GAO Feng,LIU Ting,et al. Load disaggregation method based on factorial hidden Markov model and its sensitivity analysis[J]. Automation of Electric Power Systems,2016,40(21):128-136.
- [13] 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.
- [14] 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 de-composition considering association rules of appliances' state[J]. Electric Power Automation Equipment,2020,40(4):197-203.
- [15] KIM H,MARWAH M,ARLITT M,et al. Unsupervised disaggregation of low frequency power measurements[C]//Proceedings of the 2011 SIAM International Conference on Data Mining. Philadelphia,PA,USA:Society for Industrial and Applied Mathematics,2011:1-40.
- [16] WELIKALA S,DINESH C,EKANAYAKE M P B,et al. Incorporating appliance usage patterns for non-intrusive load monitoring and load forecasting[J]. IEEE Transactions on Smart Grid,2019,10(1):448-461.
- [17] 郝杰,李星. 对经典隐马尔可夫模型的经验性改进[J]. 计算机工程与应用,2001,37(11):24-26.

HAO Jie, LI Xing. Empirical improvements of classical hidden Markov model[J]. Computer Engineering and Applications, 2001, 37(11): 24-26.

- [18] 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.

- [19] 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:



蔡 宇

蔡 宇(1995—),男,广东湛江人,硕士研究生,主要研究方向为负荷分解算法(E-mail: 21710206@zju.edu.cn);

董树锋(1982—),男,浙江杭州人,副教授,博士,通信作者,研究方向为状态估计、有源配电网分析(E-mail: dongshufeng@zju.edu.cn);

徐 航(1994—),女,浙江温州人,硕士研究生,主要研究方向为综合需求响应(E-mail: xuhang_1994@zju.edu.cn)。

(编辑 王锦秀)

Real-time disaggregation algorithm of nonintrusive load based on usage influencing factor

CAI Yu¹, DONG Shufeng¹, XU Hang¹, MAO Hangyin², SONG Yonghua³

(1. College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China;

2. State Grid Zhejiang Electric Power Co., Ltd., Hangzhou 310007, China;

3. Department of Electrical and Computer Engineering, University of Macau, Macau 999078, China)

Abstract: Aiming at the problems of low accuracy, time-consuming, etc., of traditional nonintrusive load disaggregation algorithms, a usage influencing factor based real-time load disaggregation algorithm is proposed based on HMM (Hidden Markov Model). The self-adaptive iterative K -means clustering method is used to extract load states, and the load states are combined into super-states. Aiming at the shortage that traditional HMM does not consider the time characteristic of electricity consumption scenes, the parameters are learned in time segments. At the disaggregation stage, the influencing factor of users' electricity consumption pattern is introduced to improve implicit Markov homogeneous hypothesis, and the real-time states of each load of users are decomposed by Viterbi algorithm. The accuracy and real-time performance of the proposed algorithm are verified through public data sets.

Key words: load disaggregation; super-state; HMM; self-adaptive K -means; usage pattern