# 计及电器状态关联规则的非侵人式负荷分解

徐伟枫,华锦修,余 涛,刘前进,蓝超凡 (华南理工大学 电力学院,广东 广州 510641)

摘要:非侵入式负荷监测与分解(NILMD)是获取电器用电信息的关键技术,针对当前NILMD缺乏考虑不同 电器关联运行的用电模式和电器状态的强波动性以致分解精度低的问题,提出一种计及电器状态关联规则 的新型负荷分解方法。通过仿射传播聚类提取电器的运行状态,基于互信息熵,运用关联规则算法挖掘电器 状态的关联性;调整含关联规则的样本权值并结合k近邻算法实现状态辨识;利用极大似然估计完成负荷功 率分解。测试算例验证了所提方法的有效性和准确性。

关键词:非侵入式负荷监测;仿射传播;互信息熵;关联规则;k近邻;极大似然估计

中图分类号:TM 714 文献标志码:A

DOI:10.16081/j.epae.202003015

## 0 引言

智能电网的发展和电力需求的持续增长促进了 智能用电技术的日益更新,电器层面海量的基础用 电信息逐渐体现其数据支撑的关键作用。监测电器 运行状态,一方面可引导用户合理制定用能计划,实 现节能降耗,另一方面,电网公司可通过用电数据挖 掘各类用户的用能模式及用能规律优化配置电力资 源,加强需求侧管理<sup>[1]</sup>。可见,电器负荷监测可促进 供需友好互动,具有广阔的研究价值及应用前景。

非侵入式负荷监测与分解 NILMD (Non-Intrusive Load Monitoring and Decomposition)是获取电 器用电数据的关键技术,其通过在用户电力入口安 装监测设备来采集总线的电气数据,利用负荷辨识 算法分析内部电器的运行状态。相比于传统的侵入 式监测方式,NILMD可大幅降低硬件成本,具有较 高的用户接受度。NILMD最早由George W. Hart于 20世纪80年代提出,伴随着智能量测技术的发展, 研究者致力于发掘更多新的负荷分解方法。文献 [2]利用近邻传播聚类细分负荷用电模式,并结合遗 传优化实现电器状态识别。文献[3]设计综合考虑 谐波电流和功率特征的正态分布度量函数,采用改 进鸡群算法对用电数据进行负荷分解。文献[4]立 足于电器启动时间差,针对非侵入式电流信号的欠 定求解问题,提出两步迭代收缩阈值算法,实现混合 电流与新投入负荷电流两路信号的分离。近年来, 国内外研究者开始关注除电气特征以外的其他负荷

#### 收稿日期:2019-06-24;修回日期:2020-01-15

基金项目:国家电网有限公司总部科技项目(能源互联网环 境下的多源互联配电网及多样化用电方式的需求策略系统 研究)

Project supported by the Science and Technology Project of State Grid Corporation of China (Research on Demand Strategies of Multi-source Interconnected Distribution Network and Diversified Power Consumption in Energy Internet) 信息,并提出相应的负荷分解方法,如针对负荷状态 转移概率的加性因子隐马尔科夫解码<sup>[5-6]</sup>,以及考虑 电器运行时间分布特性的超状态匹配法<sup>[7]</sup>等。文献 [8]对负荷辨识的4种典型监督学习算法进行分析, 对比其在训练时间、识别精度、识别速度、抗噪能力 等方面的表现。总体而言,当前大部分研究仍局限 于负荷的电气特征,鲜有文献考虑电器用电模式或 用户用电习惯对负荷分解的影响。

不同电器之间存在配合使用或关联运行的用电 模式,因关联性导致的电器状态联动切换对负荷分 解的影响却未见报道。为此,本文提出一种计及电 器状态关联规则(AR)的负荷分解方法,基于互信息 熵,运用Apriori关联规则算法挖掘不同电器之间状 态的关联性,对含关联规则的组合状态样本权值进 行优化配置,并结合 k 近邻(kNN)算法实现状态辨 识,利用极大似然估计完成功率分解。由此将不同 电器之间状态的关联特性融入非侵入式负荷分解 中,提高状态辨识的准确率,并解决电器状态波动性 强以致分解精度低的问题。实验算例验证了所提方 法的有效性。

## 1 用电负荷状态提取

## 1.1 状态聚类

各类电器因器件结构或工作特性的不同,其运 行时体现的电气特征一般有所差异,依据稳态特征 可将电器分为启/停二状态电器、有限多状态电器、 连续变状态电器3类<sup>[9]</sup>。前2类电器的运行状态可 通过历史数据直接提取,连续变状态电器的负荷特 征在有限范围内连续变动,无明显分区,一般通过聚 类离散化其状态,将其转换为有限多状态电器来处 理。鉴于现有智能电表采样频率的限制,以及采集 高频暂态特征带来的存储和运算压力,本文采用较 易获取的稳态功率来表征电器的运行状态,建立负 荷特征数据集。 本文将各电器的状态数视为未知量,采用仿射 传播聚类算法依次对各电器的功率数据进行聚类, 实现状态离散化。仿射传播聚类是 Frey 等在 2007 年首次提出的一种半监督聚类算法<sup>[10]</sup>。其基本思想 是将网络中各样本点视为潜在的聚类中心,基于消 息传递机制,通过迭代搜索更新各样本点与数据中 心之间的隶属关系,形成若干个簇集。聚类数根据 参考度和样本点相似度自动确定,无需事先指定,因 此相比于传统聚类算法其更适合处理状态数未知的 电器数据。

198

仿射传播聚类算法在样本点之间不断传递表征 样本点 $\mathbf{x}_i$ 支持样本点 $\mathbf{x}_q$ 作为其聚类中心的责任度 r(i,q),以及表征样本点 $\mathbf{x}_q$ 适合作为样本点 $\mathbf{x}_i$ 的聚类 中心的可靠度a(i,q)这2种信息<sup>[11]</sup>。其执行步骤 如下。

(1)输入样本集 $X = \{x_1, x_2, \dots, x_n\}$ ,计算样本 $x_i$ 和 $x_q$ 的欧氏距离负值作为两者的相似度s(i, q),利用 n个样本两两之间的相似度构建相似度矩阵 $K_{n \times n}$ , 参考度h用于控制聚类数,按照标准值设置h为 $K_{n \times n}$ 的中位数,初始化r(i, q) = 0、a(i, q) = 0,其中 $i, q \in \{1, 2, \dots, n\}$ 。

(2)更新各节点之间的责任度和可靠度值。

$$r(i,q) = s(i,q) - \max_{q' \neq q} \{a(i,q') + s(i,q')\}$$
(1)  
$$\left\{ \min \left\{ 0, r(q,q) + \sum \max \{0, r(i',q)\} \right\}$$

$$a(i,q) = \begin{cases} (1 + i + i + q) \\ i \neq q \\ \sum_{i' \neq q} \max \{0, r(i',q)\} & i = q \end{cases}$$
(2)

(3)通过阻尼因子λ调节聚类收敛速度,平抑迭代过程中的振荡,计算公式为:

 $r_{\iota+1}(i,q) = \lambda r_{\iota}(i,q) + (1-\lambda)r(i,q)$ (3)

$$a_{i+1}(i,q) = \lambda a_{i}(i,q) + (1-\lambda)a(i,q)$$
(4)

其中, $r_{\iota}(i,q)$ 、 $a_{\iota}(i,q)$ 分别为上一轮迭代的责任度和可靠度。

(4)迭代执行步骤(2)和步骤(3),直至聚类中心 不再变化或达到最大迭代次数,设*x<sub>i</sub>*的聚类中心为 *x<sub>q</sub>*,则*q*'满足式(5),提取各聚类中心作为电器的运 行状态,完成电器负荷的状态离散化。

$$q' = \operatorname{argmax} \left\{ r(i,q) + a(i,q) \right\}$$
(5)

其中,  $\operatorname{argmax} \{ \}$ 表示取 $\{ \}$ 中最大值对应的q值。

采集的电器历史数据属于时间序列,结合状态 聚类结果建立电器的历史状态转移链,转移链上的 每个节点为该电器在特定时刻的运行状态。定义电 器A的状态转移约束矩阵 **T**<sub>M\_A × M\_a</sub>(M<sub>A</sub>为电器A的状态 数),用于描述该电器在各状态之间转移的可行性。 遍历电器的历史状态转移链,当转移链上含有状态f 至状态v的切换时,令相应元素 $T_{f_r}$ =1,表示该状态 转移可行;反之置 $T_{f_r}$ =0,表示由于该电器受自身电 气特性约束或用户使用习惯,状态f无法直接转移至 状态v。图1所示为某电器的状态切换示意图,图中 实线表示可行的状态转移,虚线表示不可行的状态转 移,P、Q分别为状态对应的有功功率和无功功率。



Fig.1 Schematic diagram of state transition for an appliance

#### 1.2 功率波动模型

电器运行在某一特定状态时,其稳态功率并非固定值,而是在一个区间内持续变化的,可以认为是某一概率分布下的随机观测<sup>[12-13]</sup>。基于此,本文采用二维高斯分布描述每个状态下功率分布的随机性,构建功率波动模型,用于后续功率分解过程中的极大似然估计。根据样本数据和状态聚类结果,计算对应各状态的有功功率P和无功功率Q的均值和方差,并形成协方差矩阵,然后利用二维高斯分布概率密度函数建立功率波动模型,描述每个状态下功率分布的随机性,如式(6)所示。

$$F(\boldsymbol{w}) = 1/\left[ (2\pi)^{\frac{d}{2}} \boldsymbol{\Sigma}^{\frac{1}{2}} \right] \exp\left[ -\frac{1}{2} (\boldsymbol{w} - \boldsymbol{u})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{w} - \boldsymbol{u}) \right]$$
(6)

其中,观测向量 $w = [P,Q]^T; d$ 为向量维度;u为观测 向量均值; $\Sigma$ 为协方差矩阵;F(w)为w出现的概率。

## 2 电器状态关联规则挖掘

### 2.1 电器关联度

理论上各电器是独立工作的,但由于电器本身 的运行特性或用户习惯等,某些电器常常是配合使 用的<sup>[14]</sup>。例如,用户在使用洗碗机时热水器同时工 作;同为气温敏感型负荷的空调和电风扇往往存在 关联运行的用电模式等。相关性强的电器其运行状 态往往存在联动切换的现象,如果在负荷分解过程 中考虑这种关联性,对状态关联性强的电器给予更 高重视,可以改进负荷状态识别的准确率。

鉴于3台及以上电器配合使用的现象在居民用 户的实际生活中较为罕见,本文仅着眼于2台电器 之间的关联性。引入互信息熵定义为电器关联度, 用于描述2台电器关联运行的强弱程度。互信息熵 引自信息论中信息熵的概念,反映2个对象之间信 息的交叠程度,信息熵则表示1个观测对象的不确 定性,观测结果越随机,对象携带的信息量就越多, 其信息熵也越大。信息熵H和互信息熵J的计算公 式如下:

$$H(G) = -\sum_{g \in \Omega_{G}} p(g) \log_{2} p(g)$$
(7)

$$J(G, Y) = H(Y) - H(Y|G) = -\sum_{y \in \Omega_{Y}} p(y) \log_{2} p(y) + \sum_{x \in \Omega} \sum_{y \in \Omega} p(g, y) \log_{2} p(y|g)$$
(8)

其中, $G_{X}$ 为2个观测对象; $g_{X}$ 为2个对象的观测结果; $\Omega_{c}$ 、 $\Omega_{y}$ 分别为对象 $G_{X}$ 所有可能结果组成的集合;p为某一结果发生的概率。

本文将电器视为观测对象,通过历史运行数据 统计电器每个状态发生的概率,计算两两电器的互 信息熵,并设定互信息熵阈值,将超过阈值的2台电 器定为强关联度电器对。

#### 2.2 基于 Apriori 关联规则的状态相关性分析

关联规则是一种表征事物之间关联性的规则, 是形如U→Z的蕴涵式,其中U和Z分别为关联规则 的先导和后继<sup>[15]</sup>。布尔型关联规则是关联规则中的 一种,其U和Z均为离散的种类变量,因此可用其描述2台电器之间运行状态的关联性,例如电磁炉的 状态2和抽油烟机的状态3相互间有较强的依存关 系,则构成如下布尔型关联规则:电磁炉状态2→抽 油烟机状态3。

本文利用 Apriori 算法实施关联规则挖掘。 Apriori算法是挖掘布尔型关联规则的典型数据挖掘 算法,其通过重复扫描数据集并逐层剪枝与拼接,最 终过滤出满足要求的所有频繁项集<sup>[16]</sup>。利用 Apriori 算法挖掘电器状态关联规则的步骤如下。

(1)输入强关联度电器对的状态样本集,记为候选1项集 $C_1$ ,预设支持度 $D_{sup}$ 、置信度 $D_{con}$ 、Kulczynski 系数 $D_{KULC}$ (即KULC度量值)这3个变量的阈值。

(2)剪枝。对于候选l项集 $C_l$ ,统计 $C_l$ 中出现的 所有状态组合,计算 $D_{sup}$ 和 $D_{con}$ ,将低于阈值的状态 组合剔除(剪枝),得到频繁l项集 $L_l$ 。以2项集为例,  $D_{sup}$ , $D_{con}$ 的计算公式如下:

$$D_{\text{sup}}(A_f, B_v) = p(A_f \cap B_v) \tag{9}$$

$$D_{\rm con}(A_f \Longrightarrow B_v) = p\left(B_v \middle| A_f\right) \tag{10}$$

其中, $D_{sup}(A_f, B_e)$ 为电器A的状态f和电器B的状态  $v同时发生的概率; D_{con}(A_f \Rightarrow B_e)$ 为在观测到电器A处于状态f时,电器B运行于状态v的概率。

(3)拼接。将*L*<sub>l</sub>中的项目进行两两组合(拼接), 形成新一层候选集*C*<sub>l+1</sub>。

(4)过滤。计算 L<sub>i</sub>中所有状态组合的 KULC 度 量值 D<sub>KULC</sub>,超过阈值的状态组合输出为第 l 层的关 联规则(过滤)。D<sub>KULC</sub>的计算公式如下:

$$D_{\text{KULC}} = \frac{1}{2} \left[ D_{\text{con}}(A_f \Rightarrow B_v) + D_{\text{con}}(B_v \Rightarrow A_f) \right] (11)$$

(5)重复步骤(2)—(4),直至完成所有层级的关 联规则输出。

Apriori 算法挖掘电器状态关联规则的完整执行 流程如图 2 所示。算法的扫描范围限定在强关联度 电器对中,因此候选集的拼接仅上升至 C<sub>2</sub>层级,减少 了不必要的数据扫描工作。

状态样本集



图 2 Apriori 算法执行流程图 Fig.2 Execution flowchart of Apriori algorithm

## 3 基于关联规则-k近邻算法的负荷分解 方法

## 3.1 关联规则-k近邻状态辨识法

状态辨识的本质是在同一时刻识别多台电器的运行状态。本文定义在t时刻待测电器状态的组合情况为一个组合状态 $S_i = [s^{(i)}, s^{(2)}, \dots, s^{(N)}](N$ 为电器数, $s^{(A)}$ 为电器A的状态),将每个时刻的 $S_i$ 视为单个待识别实例,利用关联规则对kNN算法进行改进,提出关联规则-k近邻(AR-kNN)状态辨识法,结合组合状态样本实施分类。组合状态样本可通过对各电器的历史状态转移链进行组合,并提取组合链上各节点的组合状态及对应的负荷总功率来获取。

kNN算法是数据挖掘分类技术中的一种非参数 分类方法,现已在模式识别的各个领域中得到广泛 应用<sup>[17]</sup>。其基本思想是搜索待测实例在特征空间中 的*k*个距离最近的样本,若*k*个样本中属于某一类别 的样本数量占比最多,则判定该实例属于这一类别。 电器状态关联规则的*D*<sub>KUC</sub>是衡量关联规则强弱的 重要指标,*D*<sub>KUC</sub>越大,关联规则越强,可认为2台电 器的关联状态同时出现的可能性较高,即相比于非 关联状态对的发生具有更高的可信度。本文在kNN 算法对含关联规则的组合状态样本附乘一关联因 子,关联因子通过*D*<sub>KUC</sub>计算,公式如下:

$$\boldsymbol{\beta}_i = \prod_{z=1}^{m_i} 2e^{D_{\text{KULC}}(z)}$$
(12)

其中, $\beta_i$ 为第*i*个组合状态样本的关联因子; $m_i$ 为该 样本所含的关联规则数; $D_{\text{KULC}}(z)$ 为第*z*条关联规则。 由 $D_{\text{KULC}} \in [0,1]$ 可知,每条关联规则对关联因子的 最大增幅不会超过自然对数的2倍,因此通过式 (12)定义的关联因子可避免在算法中对含关联规则 的样本赋予过高权重导致完全忽略其余样本的 问题。

200

算法中k值的训练通过S折交叉验证来完成,即 按天将组合状态样本数据分成S份,每次选取S-1份 作为训练集,剩余1份作为测试集,若干轮训练后, 选择状态估计准确率平均值最高的k值。状态估计 准确率 $c_{\text{state}}$ 的计算如式(13)所示<sup>[18]</sup>。

$$c_{\text{state}} = \frac{\sum_{t=1A=1}^{n} \sum_{A=1}^{N} I(s_{t}^{(A)} = \tilde{s}_{t}^{(A)})}{NR}$$
(13)

其中, $s_{t}^{(A)}$ 、 $\tilde{s}_{t}^{(A)}$ 分别为电器 $A \propto t$ 时刻的状态辨识值、 实际值;R为采样时段长度, $I(\cdot)$ 为指示函数。

具体地,AR-kNN状态辨识法原理如下。

(1)输入当前采样点的非侵入式负荷稳态功率
向量*x<sub>i</sub>* = [*P<sub>i</sub>*, *Q<sub>i</sub>*]<sup>\*</sup>,依次计算其与样本集*X*′中*n*个训
练样本*x<sub>i</sub>*′的特征距离,即:

 $d_{i} = \| \mathbf{x}_{i} - \mathbf{x}_{i}' \| \quad \mathbf{x}_{i}' \in \mathbf{X}'; \ i = 1, 2, \cdots, n \quad (14)$ (2) 根据  $d_{i}$  从  $\mathbf{X}'$  中选出与  $\mathbf{x}_{i}$  距离最近的 k 个
样本。

(3)对含关联规则的样本权重乘上关联因子,关 联因子按式(12)计算,其余样本的权重保持不变。

(4)按加权占比从大到小排列 k 个近邻所属的 组合状态类作为候选。考虑负荷成分的电器稀疏 性,即当出现同一实测功率值对应多个组合状态的 情况时,更倾向于以较少数目的电器组合来拟合实 测值<sup>[19]</sup>。因此若多个候选类的占比相同,则仅保留 处于非待机状态的电器数量最少的一类,剔除其余 候选类。

(5)从第一个候选类依次往下选取,直至选出的 候选类中各电器的状态与上一采样点的状态之间满 足状态转移约束矩阵**T**的约束,此时判定待测实例 属于该候选类。

(6)输入下一采样点的非侵入式负荷稳态功率 值,返回步骤(1)继续计算,直至完成所有采样点的 辨识。

#### 3.2 功率分解

对于一个在采样时段长度*R*采集的总负荷功率 序列,非侵入式负荷分解的最终目的是得到*N*台电 器对应的功率序列  $P^{(A)} = [P_1^{(A)}, P_2^{(A)}, \dots, P_R^{(A)}](A = 1, 2, \dots, N)$ 。前述AR-kNN状态辨识的输出为电器的 运行状态序列 $s^{(A)} = [s_1^{(A)}, s_2^{(A)}, \dots, s_R^{(A)}](A = 1, 2, \dots, N)$ 由 于电器在特定状态下的功率并非固定值,而是持续 波动的,若将状态对应的聚类中心作为分解结果,忽略功率波动造成的状态偏移,则分解精度会偏低。因此,需要在每个时刻根据电器状态对应的功率波动模型估计功率大小。

根据极大似然估计求解联合概率最大化问题的 思路,同时考虑分解得到的所有电器的功率之和等 于同一时刻负荷总功率的约束条件,本文利用式(6) 构建如下优化模型求解电器的功率大小:

$$\begin{cases} \max_{P_{t}^{(1)}, Q_{t}^{(1)}, \dots, P_{t}^{(N)}, Q_{t}^{(N)}} \prod_{A=1}^{N} F_{A} \left( P_{t}^{(A)}, Q_{t}^{(A)} \middle| s_{t}^{(A)} \right) \\ \text{s.t.} \quad \sum_{A=1}^{N} P_{t}^{(A)} = P^{(L)} \\ \sum_{A=1}^{N} Q_{t}^{(A)} = Q^{(L)} \end{cases}$$
(15)

其中, $P_{i}^{(A)}$ 、 $Q_{i}^{(A)}$ 分别为t时刻电器A的有功功率值、无 功功率值; $P^{(L)}$ 和 $Q^{(L)}$ 分别为该时刻总负荷的有功功 率和无功功率; $F_{A}(P_{i}^{(A)}, Q_{i}^{(A)}|s_{i}^{(A)})$ 为式(6)所构建的功 率波动模型,即电器A处于状态 $s_{i}^{(A)}$ 时消耗功率 $P_{i}^{(A)}$ 和 $Q_{i}^{(A)}$ 的概率。

本文计及电器状态关联规则的非侵入式负荷分 解方法流程如图3所示,以电器1为例, $v_1^{(1)}$ 表示电器 1包含的第一个状态,则 $s_t^{(0)} \in \{v_1^{(1)}, v_2^{(1)}, \dots, v_{M_t}^{(0)}\}_{\circ}$ 







## 4 算例分析

#### 4.1 状态提取

本文采用加拿大学者 Stephen Makonin 等建立的 AMPds2数据集,对所提方法进行验证。AMPds2 采集了温哥华市一户家庭中的电器约2 a的用电数据,包括电压、电流、有功、无功等11种电气特征,采样频率为1次/min。由于其采集的是居民用电的真 实数据,因此相比于计算机合成数据能更好地反映 居民用户的实际用电行为,适合作为算例分析样本。

从中选取冰箱(FG)、电视机(TV)、电暖炉 (FR)、洗碗机(DW)、热泵空调(HP)、照明设备(LA) 6种电器连续10d(14400个采样点)的时序功率信 号用于训练算法。通过状态聚类得出各电器的状态 数分别为:冰箱5种、电视机5种、电暖炉3种、洗碗 机5种、热泵空调13种、照明设备7种。对每台电器 构建出状态对应的功率波动模型,如电视机4种非 待机状态的功率波动热力图见附录中图A1,图中点 A'表示电视机处于状态5时其功率波动至有功功率 370W、无功功率31 var的概率为0.006。

## 4.2 关联规则挖掘

计算6种电器两两之间的互信息熵,设定互信 息熵阈值为0.2 bit,计算结果如表1所示。由表1可 知,仅有2台电器对的互信息熵超过阈值,其中热泵 空调和电暖炉的互信息熵为0.8637 bit,电视机和照 明设备的互信息熵为0.2307 bit,此即为2个强关联 度电器对。

表1 互信息熵计算结果

Table 1 Calculation results of mutual

information entropy						
山鬼		、熵 / bit				
电位	FG	$\mathbf{D}\mathbf{W}$	HP	FR	LA	TV
FG	_	0.0017	0.0103	0.0103	0.0018	0.0009
$\mathbf{DW}$	0.0017	_	0.0060	0.0060	0.0133	0.0020
HP	0.0103	0.0060	_	0.8637	0.0089	0.0011
$\mathbf{FR}$	0.0103	0.0060	0.8637	—	0.0090	0.0011
LA	0.0018	0.0133	0.0089	0.0090	—	0.2307
TV	0.0009	0.0020	0.0011	0.0011	0.2307	_

考虑到2台电器的状态组合数量过多,而关联 性强的2台电器在运行过程中其状态一般不存在明 显的一一对应关系,因此实验中先利用Kmeans聚类 将强关联电器对中各台电器的状态聚合成若干个状 态集,再通过Apriori算法对状态集序列进行数据挖 掘,输出关联规则,如表2所示。

表2 电器状态的关联规则

Table 2 Association rules of appliances' state

规则 序号	$D_{\rm KULC}$	先导电器	先导状态	后继 电器	后继 状态
1	0.9999	HP	1,2	$\mathbf{FR}$	1
2	0.9994	HP	3—12	$\mathbf{FR}$	2
3	1.0000	HP	13	$\mathbf{FR}$	3
4	0.9365	TV	1	LA	1,2
5	0.6765	TV	2,3	LA	3,4
6	0.5186	TV	4,5	LA	3,4

## 4.3 分解结果

对单电器,定义状态辨识准确率*a*<sub>state</sub>和功率分 解准确率*a*<sub>power</sub>这2个指标,以评估算法的分解精度:

$$a_{\text{state}} = \frac{\sum_{t=1}^{R} I(s_t = \tilde{s}_t)}{R}$$
(16)

201

$$a_{\text{power}} = 1 - \frac{\sum_{i=1}^{n} \left| P_{i} - \tilde{P}_{i} \right|}{2 \sum_{i=1}^{n} \tilde{P}_{i}}$$
(17)

其中, $s_i$ 和 $P_i$ 分别为该电器在t时刻的状态辨识值和 功率分解值; $\tilde{s}_i$ 和 $\tilde{P}_i$ 分别为状态实际值和功率实际值。

选取某一天6种电器时序功率的叠加作为待分 解信号,分解结果见附录中图A2。

以文献[4]所提的针对组合优化的遗传算法 (GA)、文献[5]针对正态分布度量函数的改进鸡群 算法(ICSO)以及kNN算法作为对比,采用同样20d 的数据分别进行20次仿真实验,4种方法的平均准 确率如表3所示。可以看出,本文基于AR-kNN的负 荷分解法对总负荷的分解效果更好。

表3 4种算法平均准确率对比

Table 3	Average	accuracy	rate	comparison	among
		four algo	rithm	0	

		ioui a	igoritiniis				
北标	电器 -	平均准确率 / %					
1日 仰小		GA	ICSO	kNN	AR-kNN		
	FG	90.92	90.64	93.90	95.75		
	TV	89.11	89.18	93.28	94.47		
	$\mathbf{FR}$	92.38	94.94	97.86	98.08		
$a_{ m state}$	$\mathbf{D}\mathbf{W}$	97.44	97.38	96.33	97.64		
	HP	79.25	72.51	91.06	92.59		
	LA	78.21	77.17	95.99	97.03		
	FG	85.07	85.25	86.57	90.10		
	TV	88.91	88.53	92.82	94.73		
	$\mathbf{FR}$	97.03	97.08	97.13	97.23		
$a_{ m power}$	$\mathbf{D}\mathbf{W}$	90.79	89.95	92.43	96.66		
	HP	97.06	96.68	97.19	97.63		
	LA	89.25	89.65	93.50	95.71		

将kNN算法和GA、ICSO的平均准确率进行对 比,可见kNN算法的表现总体上更优。分析知,类 似GA和ICSO的方法单纯进行启发式状态搜索,并 将搜索结果对应的功率代表值作为电器的运行功 率,这类方法假定电器每种状态的负荷特征固定不 变,忽略了状态的波动性,当多台电器的特征固定不 变,忽略了状态的波动性,当多台电器的特征因状态 波动在特征空间内重叠时,就可能判断错误。kNN 算法则利用了组合状态样本的全局分布特性,将特 征空间内的 k个近邻作为候选,降低了误判的可能 性。在洗碗机的状态辨识准确率上,kNN算法略低 于GA和ICSO,究其原因为该用户的洗碗机属于使 用频率较低的电器,即训练集中包含洗碗机非待机 状态的组合状态样本数较少,可见kNN算法对样本 的依赖程度较高。

本文所提方法在kNN算法的基础上加入不同

电器运行状态之间的关联性,对含关联规则的组合 状态样本配置更高权重,优化了样本总体的权值分 布,因此相比于 kNN 算法取得了更好的辨识效果。 对于热泵空调和电暖炉这对气温敏感型负荷(或具 有配合使用习惯的电视机和照明设备),在候选类中 利用关联因子对包含两者联动切换状态的近邻给予 更高可信度,从而提高总体的识别准确率。利用极 大似然估计构建功率分解的优化模型,以抵消状态 波动引起的功率偏移,保证了各台电器的分解功率 之和等于总负荷功率。

由附录中图A2可见,电暖炉具有恒运行、小功 率切换的用电模式,但仍可取得较好的分解效果,因 此对于识别低功率或恒运行但与其他电器存在状态 关联性的电器,AR-kNN方法也可有效应用。在811 个采样点附近的洗碗机、电视机、照明设备及918个 采样点附近的冰箱、电视机、照明设备,均出现明显 的持续2~3 min的异常状态判断,冰箱在某些时段发 生频繁的启停判断。通过增加负荷特征、考虑电器 状态与运行时常的相关性等,有望解决这类问题。

## 5 结论

202

本文提出了一种考虑电器状态关联规则的非侵 人式负荷分解方法。采用仿射传播聚类提取电器的 运行状态,建立包含功率波动模型的负荷特征集,利 用 Apriori 算法挖掘不同电器之间运行状态的关联 性,对含关联规则的组合状态样本权值进行优化配 置,并结合 kNN 算法实现电器的状态辨识,最后基 于极大似然估计构建功率分解优化模型。算例结果 表明,相比于利用启发式算法进行状态搜索的方式, 本文所提方法取得了更高的负荷分解精度,对 NILMD 的实际应用具备一定可行性。本文工作暂 局限于考虑两两电器之间状态的关联性,下一步研 究将着眼于更高维度的关联性分析,如电器状态与 运行时长的相关性、气温敏感型电器的状态与温湿 度的相关性等,提出适用性更广泛的负荷分解方法。

附录见本刊网络版(http://www.epae.cn)。

### 参考文献:

- [1] 王蓓蓓,李扬. 面向智能电网的电力需求侧管理规划及实施机 制[J]. 电力自动化设备,2010,30(12):19-24.
   WANG Beibei, LI Yang. Demand side management planning and implementation mechanism for smart grid[J]. Electric Power Automation Equipment,2010,30(12):19-24.
- [2] 徐青山,娄藕蝶,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.

[3] 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.

 [4]武昕,韩璐,韩笑,等. 欠定分离机制下基于特征滤波的居民 负荷非侵入辨识算法[J]. 电力系统自动化,2017,41(20): 118-125.

WU Xin, HAN Lu, HAN Xiao, et al. Feature filtering based non-intrusive identification algorithm for residential load in underdetermined separation mechanism[J]. Automation of Electric Power Systems, 2017, 41(20):118-125.

- [5] KOLTER J Z, JAAKKOLA T, KOLTER J. Approximate inference in additive factorial HMMs with application to energy disaggregation [C] //Proceedings of the International Conference on Artificial Intelligence and Statistics. La Palma, The Republic of Panama: [s.n.], 2012:1472-1482.
- [6] BONFIGLI 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.
- [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]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.
- [9] HART G W. Nonintrusive appliance load monitoring[J]. Proceedings of the IEEE, 1992, 80(12):1870-1891.
- [10] FREY B J,DUECK D. Clustering by passing messages between data points[J]. Science,2007,315(5814):972-976.
- [11] 褚壮壮,梁军,张旭,等.基于仿射传播聚类算法的广义负荷稳态特性建模及其应用[J].电力自动化设备,2016,36(3): 115-123.
  CHU Zhuangzhuang,LIANG Jun,ZHANG Xu, et al. Modeling of generalized load steady-state characteristics based on affinity propagation clustering algorithm and its application[J]. Electric Power Automation Equipment,2016,36(3):115-123.
- [12] LIANG J, NG S K K, 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.
- [13] LIANG J, NG S K K, 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.
- [14] 孙毅,刘迪,李彬,等. 基于家庭用电负荷关联度的实时优化策略[J]. 电网技术,2016,40(6):1830-1834.
  SUN Yi,LIU Di,LI Bin,et al. Research on real-time optimization strategy based on correlation of household electrical load[J]. Power System Technology,2016,40(6):1830-1834.
- [15] 孙丰杰,王承民,谢宁.面向智能电网大数据关联规则挖掘 的频繁模式网络模型[J].电力自动化设备,2018,38(5): 110-116.

SUN Fengjie, WANG Chengmin, XIE Ning. Frequent pattern network model for association rule mining of big data in

smart grid[J]. Electric Power Automation Equipment,2018,38
(5):110-116.

- [16] 郭晓利,于阳. 基于云计算的家庭智能用电策略[J]. 电力系 统自动化,2015,39(17):114-119,133.
   GUO Xiaoli,YU Yang. A residential smart power utilization strategy based on cloud computing[J]. Automation of Electric Power Systems,2015,39(17):114-119,133.
- [17] 耿丽娟,李星毅.用于大数据分类的KNN算法研究[J]. 计算机应用研究,2014,31(5):1342-1344,1373.
   GENG Lijuan,LI Xingyi. Improvements of KNN algorithm for big data classification[J]. Application Research of Computers, 2014,31(5):1342-1344,1373.
- [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] 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-feature realtime power load disaggregation considering sparsity [J]. Pro-

ceedings of the CSU-EPSA, 2018, 30(2):79-84.

#### 作者简介:



徐伟枫(1994—),男,广东广州人,硕 士研究生,主要研究方向为智能配电网、 智能用电等相关技术(E-mail: fengalsk@ foxmail.com);

华锦修(1995—),男,广东韶关人,硕 士研究生,主要研究方向为智能配电网相关 技术(E-mail:hua1995220hua@163.com);

徐伟枫

余 涛(1974—),男,浙江宁波人,教 授,博士研究生导师,博士,通信作者,主要

研究方向为复杂电力系统的非线性控制理论、优化及机器学  $\Im(\mathbf{E} ext{-mail:taoyu1@scut.edu.cn});$ 

刘前进(1967—),男,广东广州人,副教授,博士,主要研究方向为保护与控制、配电网自动化及系统数字仿真(E-mail:qjliu@scut.edu.cn);

蓝超凡(1996—),男,广东河源人,硕士研究生,主要研 究方向为智能用电、负荷特性分析(E-mail: cflanscut@163. com)。

(编辑 王锦秀)

## Non-intrusive load decomposition considering association rules of appliances' state

XU Weifeng, HUA Jinxiu, YU Tao, LIU Qianjin, LAN Chaofan

(School of Electric Power, South China University of Technology, Guangzhou 510641, China)

Abstract: NILMD (Non-Intrusive Load Monitoring and Decomposition) is the key technology to obtain electricity consumption information of appliances. Aiming at the problem of low decomposition accuracy which is caused because that current NILMD lacks of concerning about the electricity consumption pattern of associated operation between different appliances and high volatility of appliances' state, a novel load decomposition method considering the association rules of appliances' state is proposed. The operation state of each appliance is extracted by affinity propagation clustering. Based on mutual information entropy, the association rule algorithm is used to mine the association of appliances' state. The sample weights with association rules are adjusted and combined with kNN(k-Nearest Neighbor) algorithm to realize the state identification. The maximum likelihood estimation is used to decompose load power. Test examples verify the effectiveness and accuracy of the proposed method.

Key words: non-intrusive load monitoring; affinity propagation; mutual information entropy; association rule; kNN; maximum likelihood estimation







Fig. A1 Thermodynamic chart of power fluctuation for television at non-standby state



Fig. A2 Results of load decomposition