

基于电器粗糙归类的无监督 NILM 结果自主标注

肖潇¹, 栾文鹏¹, 刘博^{1*}, 王岩², 杨劲男¹, 刘子帅¹, 韦尊¹

(1. 天津大学电气自动化与信息工程学院, 天津市 南开区 300072;

2. 天津求实智源科技有限公司, 天津市 西青区 300392)

Autonomous Labeling of Unsupervised NILM Results Based on Rough Classification of Appliances

XIAO Xiao¹, LUAN Wenpeng¹, LIU Bo^{1*}, WANG Yan², YANG Jinnan¹, LIU Zishuai¹, WEI Zun¹

(1. School of Electrical Engineering & Automation, Tianjin University, Nankai District, Tianjin 300072, China;

2. Tianjin Qiushi Transenergy Technologies Co., Ltd., Xiqing District, Tianjin 300392, China)

ABSTRACT: The unsupervised non-intrusive load monitoring (NILM) method can realize load decomposition by analyzing the unlabeled aggregated load data directly. However, the corresponding appliance name cannot be automatically determined generally, which affects the interpretability of NILM results and hinders its scale application. In this paper, a novel method for autonomous labeling of unsupervised NILM results based on rough classification of appliances was proposed. Both operation characteristics and usage time distribution of appliances were analyzed and summarized. Three categories of operation characteristics as periodic operation, intensive fluctuation, and fixed-duration operation were defined separately. Correspondingly, an adaptive operation characteristic inference method based on cluster analysis of appliance operation data was provided. A method for identifying the appliance usage characteristics based on the correlation between the usage of appliances and the human activities in term of time distribution was proposed. At the same time, an index to mark the periods of strong/weak human activities was given based on the change of the load component. Furthermore, the appliances were roughly classified according to the defined operation and usage characteristics based on the rough set theory, and then a bilevel decision method was proposed to realize the NILM result labeling. Experiments in private and public datasets show that the proposed method could label NILM results of common appliances in different scenarios accurately. The proposed method can be integrated with other unsupervised NILM algorithm to form a fully unsupervised NILM solution.

KEY WORDS: unsupervised non-intrusive load monitoring (NILM); autonomous appliance name labeling; general operation characteristics; adaptive discrimination; rough classification of appliances

摘要: 无监督非侵入式负荷监测 (non-intrusive load monitoring, NILM) 方法通常无法自动确定分解结果所对应的电器名称, 这影响 NILM 结果的用户可读性, 阻碍了其规模应用。为此, 该文提出一种基于电器粗糙归类的无监督 NILM 结果自主标注方法。从电器运行控制方式和使用时间分布 2 个方面, 总结分析同类电器共同具有的通用运行特性。定义周期运行、密集波动、固定时长运行 3 种控制规律特性, 给出基于聚类分析的电器相应特性自适应判别方法; 对于受人类活动影响程度不同的不同电器, 提出基于电器使用与人类活动强弱的时间分布之间相关性的电器使用规律特性判别方法, 同时给出一种基于负荷成分变化的用户个性化人类活动强弱时段划分方法。在此基础上, 基于粗糙集理论, 依据上述 2 类通用运行特性进行电器粗糙归类, 进而提出融合通用运行特性的电器名称两层决策方法, 实现 NILM 结果标注。在私有和公开数据集集中的实验表明, 该方法能在不同场景下实现常见家用电器 NILM 结果准确标注。所提方法可作为任意无监督 NILM 方法的后续步骤与之集成, 形成完全无监督 NILM 方案。

关键词: 无监督非侵入式负荷监测; 电器自主标注; 通用运行特性; 自适应判别; 电器粗糙归类

0 引言

非侵入式电力负荷监测 (non-intrusive load monitoring, NILM) 是一种成本效益高的负荷用电细节监测技术^[1]。它仅通过分析负荷总量用电数据便可获取其内部每种电器的实时工作状态 (包括启、停时间) 和累积耗电量等用电信息。所得信息能够使能用户行为精细化分析、高级电力市场化和需求侧管

基金项目: 国家自然科学基金项目(重点项目)(U2066207)。

Project Supported by National Natural Science Foundation of China (Key Program)(U2066207).

理(包括响应潜能评估、负荷精准调控与响应行为校验等)等高级应用^[1-8], 促进智能电网^[9]发展。

根据“非侵入”程度^[10]不同, 可将 NILM 分为有监督 NILM^[11-17]和无监督 NILM^[10,18-27]2 种。对于前者, 在监测系统投入使用之前, 需要人工协助获取负荷内部每个所关心电器的用电数据, 用以建立电器负荷模型, 进而形成负荷印记特性库, 这一过程通常需要侵入负荷内部; 此外, 为了避免用户新增的电器无法识别, 监测系统需要人工长期跟踪维护。相对地, 后者仅利用少量关于电器运行特性的先验知识, 便能直接处理无标注的负荷用电量数据进行电力负荷监测与分解。由于实际场景中负荷构成多样, 电器工作原理复杂且更新换代频繁, 人工协助获取负荷印记行之不便, 无监督 NILM 无疑是 NILM 技术规模应用的必然选择。

完全无监督 NILM 框架可概括为图 1, 其与有监督 NILM 的主要区别在于电器模型建立和监测结果自主标注这 2 个环节。在有监督 NILM 中, 通过有监督学习进行电器建模, 完成后可直接得到电器模型与其物理名称(如电热水器、电冰箱、空调等)的对应关系。在无监督 NILM 中, 通过无监督学习建立电器模型^[21-22], 所建立的电器模型通常没有物理名称, 为了有效支持用户行为分析、负荷精准调控等具体到电器的众多预期高级应用, 从而实现无监督 NILM 大规模的工程实用, 还需利用通用先验知识对 NILM 结果中“未知”电器的物理名称进行自主标注^[10], 简称 NILM 结果自主标注。

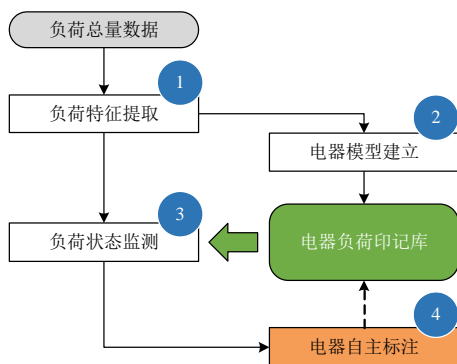


图 1 完全无监督 NILM 框架图

Fig. 1 Fully unsupervised NILM framework diagram

截至目前, 未见对 NILM 结果自主标注问题的针对性研究。绝大多数无监督 NILM 方法^[18-27]无法给出带有电器物理名称的监测结果, 在评价其准确性时, 需借助(由侵入式监测装置采集的)电器分项数据, 根据从中提取的负荷印记^[1]对监测结果中每条电器运行记录进行逐个匹配标注。此过程将会耗

费大量人力物力, 十分繁琐复杂, 不利于无监督 NILM 方法的实地部署应用。而自主标注是直接提取监测结果中电器全部运行记录所呈现的负荷印记, 根据通用先验知识实现其物理名称自动标注。这 2 种标注方案使用的负荷印记通常包括电压、电流和功率等电气信号中呈现的稳态特征与暂态特征, 以及电器正常运行时表现出的全局性统计特性等^[26]。而标注过程中与电器负荷印记有关的先验知识, 可分为经验特性和物理特性 2 种。经验特性包括容量(有功功率和无功功率)、运行(或使用)和关机时长、典型使用时段^[27]等。物理特性包括同类电器具有的相同(或类似)工作原理、功率因数^[28]、谐波特性^[29], 及由工作原理决定的负荷事件切换频率^[28]等。不过, 无论是经验特性还是物理特性, 现有研究中先验知识的表现形式通常是某些关键参数的数值区间或概率分布, 本文称之为参数特性。然而数值区间的限值或概率分布参数通常源自对一些典型场景(真实住户或实验室)的调研分析。由于同类电器的品牌和型号多种多样, 不同用户的行为千差万别, 参数特性的场景泛化能力有限。

对此, 本文提出一种基于电器粗糙归类的无监督 NILM 结果自主标注方法。从电器运行控制方式和使用时间分布 2 个方面, 总结同类型电器具有的通用运行特性。通过分析受控方式不同的电器工作循环发生规律, 定义周期运行、密集波动、固定时长运行 3 种控制规律特性, 研究基于近邻传播聚类的电器相应特性自适应判别方法; 对于受人类活动影响程度不同的不同类型电器, 建立基于电器使用与人类活动强弱的时间分布之间相关性的使用规律特性判别方法, 同时基于 NILM 获取的单个电器用电数据, 提出一种基于负荷成分变化的用户个性化人类活动强弱时段划分方法。在此基础上, 基于粗糙集理论, 依据上述两类通用运行特性进行电器粗糙归类, 进而提出融合通用运行特性的电器名称分层决策方法, 最终实现 NILM 结果标注。

本文结构如下: 首先, 第 1 节介绍可用于 NILM 结果自主标注的电器通用运行特性及其判别方法。而后, 第 2 节给出基于粗糙集的无监督 NILM 结果自主标注方法。接下来, 第 3 节通过实验验证本文所提方法的有效性。最后, 第 4 节总结全文。

1 电器通用运行特性

同类(或允许被赋予相同名称的)电器在某些运

行特性方面具有一致性,不因电器品牌和型号或用户行为的差别而不同,本文将此类特性称为通用运行特性。根据电器运行控制方式和用户使用规律,本文定义控制规律特性与使用规律特性两种通用运行特性。

1.1 控制规律特性

常见家用电器的控制方式通常为自动控制型(温度控制型、定时控制型等)或和人工控制型2种。其中不少复杂电器投、切或运行时间没有明显规律,但取决于其内部调控机制,这些电器工作过程中工作循环的持续时长或发生间隔的特征显著。电器工作循环指电器从某个起始状态经过一段时间用电回到原始状态的过程,同一电器可能有多种不同的工作循环。具体来讲,电器从开启到关停的整个过程是一个工作循环,对于有些多状态电器,中间状态之间发生转换也会形成工作循环,相对于前者可称之为子循环。以洗衣机为例,它从启动经过洗涤-漂洗-烘干等全部工序后关停的过程是一个工作循环,而在其洗涤过程中电机间歇性启停改变旋转方向会形成多个(子)工作循环。在本文中二者不做严格区分。基于对多种电器的用电数据和和工作方式分析,定义3种控制规律特性。

1) 周期运行(periodic operation, PO)电器某种工作循环有规律地重复出现称为周期运行,此时工作循环的持续时长和相邻2次发生的时间间隔都相对固定。如空调制热保温模式,其内部温控元件控制发热元件周期性启停来维持室温,实例见图2。再如洗衣机的洗涤模式,其内部定时器控制电机周期性反复旋转,实例如图0所示。

2) 密集波动运行(intensive fluctuation operation, IO)电器某种工作循环短时间内连续多次频繁发生称为密集波动运行,此时工作循环持续时长与相邻两次发生的时间间隔都相对较小。与图2空调这类间歇式运行的电器有明显区别,图3洗衣机洗涤模式和图4电饭锅保温模式(图4红框标识)

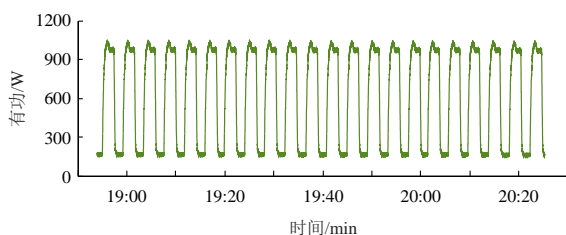


图2 空调制热保温状态下的有功功率曲线
Fig. 2 Active power curve of air conditioner in heating and holding state

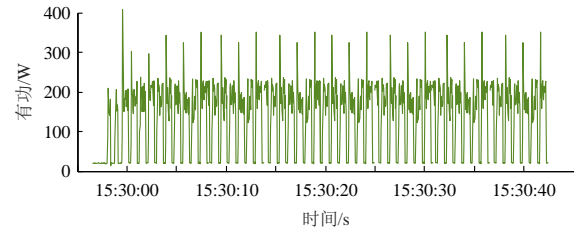


图3 洗衣机洗涤过程的有功功率曲线

Fig. 3 Active power curve of washing machine during washing process

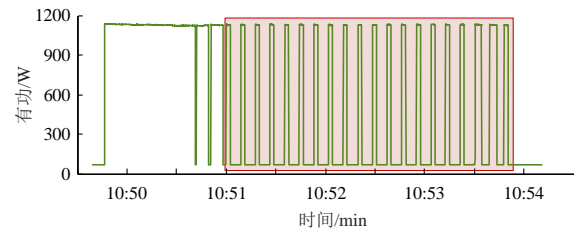


图4 电饭锅工作时的有功功率曲线

Fig. 4 Active power curve of the rice cooker

皆呈现出密集波动的特点。

3) 固定时长运行(fixed duration operation, FO)电器某种工作循环持续时长变化不大称为固定时长运行。周期运行的电器通常也具有固定时长运行特性。当用户每次烧水注入的冷水量差别不大时,电热水壶(属于温度控制型电器)具有接近固定时长运行特性,实例见图5。

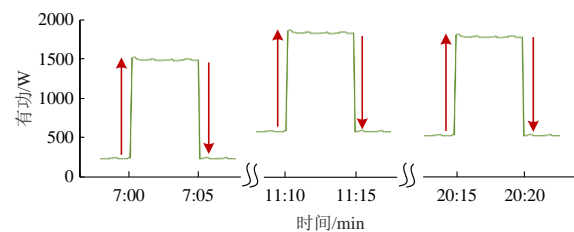


图5 电热水壶烧水过程的有功功率曲线

Fig. 5 Active power curve of kettle during boiling

对于具有不同控制规律特性的电器,其运行记录在工作循环持续时长-发生间隔二维平面上的分布规律具有明显不同,如图6所示。据此,本文提出一种基于近邻传播聚类分析的自适应判别算法,用于判断“未知”电器的控制规律特性。

为叙述方便,先给出相关符号说明。结合图7,对于“未知”电器 x 的某种工作循环,如果其任意第 i 条记录的开始和结束时刻分别由 t_1 和 t_2 表示,第 $i+1$ 条记录的开始时刻由 t_3 表示,将第 i 条记录和第 $i+1$ 条记录记为一个相邻工作循环对,那么,工作循环持续时长 $\Delta T_{on}=t_2-t_1$,工作循环的间隔 $\Delta T_{off}=t_3-t_2$ 。已积累的所有运行记录的 ΔT_{on} 共同构成集合 Ω_{on} , $(\Delta T_{on}, \Delta T_{off})$ 共同构成集合 Ω_{Δ} 。

周期运行特性判别算法的具体步骤如下:

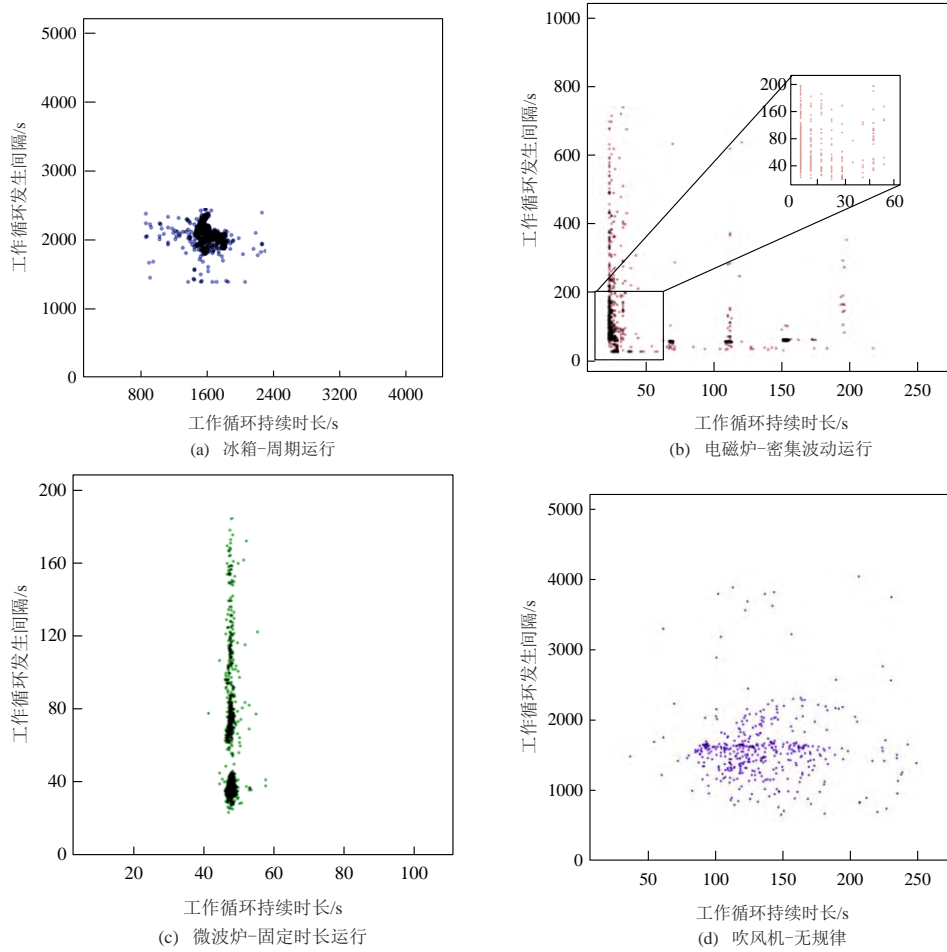


图 6 工作循环持续时长-发生间隔的二维分布图

Fig. 6 Distribution diagram of the operation cycle vs the interval of occurrence

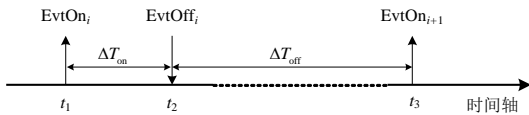


图 7 电器工作循环示意图

Fig. 7 Schematic diagram of appliance operation cycle

步骤 1: 对于“未知”电器 x 某种工作循环, 计算新增运行记录的 $(\Delta T_{on}, \Delta T_{off})$, 并更新集合 Ω_L ;

步骤 2: 若 Ω_L 中元素总数大于 N , 则执行 Step 3, 否则返回 Step 1 继续积累数据;

步骤 3: 对集合 Ω_L 中元素进行聚类分析, 共得到 K 个簇 $C_L^1, C_L^2, \dots, C_L^K$;

步骤 4: 遍历全部 K 个簇, 得到满足判别式(1)的所有簇, 并将它们的序号记录在集合 Θ_L 中:

$$\theta_L^k = |C_L^k| / |\Omega_L| \geq \theta_L^{th} \quad (1)$$

式中: $|\Omega_L|$ 表示 Ω_L 中的元素总数; $|C_L^k|$ 表示第 k 个簇所含元素个数, $\forall k \in \{1, 2, \dots, K\}$; θ_L^{th} 表示周期运行特性第一级判别阈值。

步骤 5: 对于集合 Θ_L , 若满足判别式(2), 认为“未知”电器 x 具有周期运行特性。

$$\delta_L = \sum_{r=1}^{R_L} \theta_L^{k_r} \geq \delta_L^{th} \quad (2)$$

式中: R_L 表示集合 Θ_L 中的元素总数; 对于 $\forall r \in \{1, 2, \dots, R_L\}$, 有 $k_r \in \{1, 2, \dots, K\}$; δ_L^{th} 表示周期运行特性第二级判别阈值。

对于步骤 3, 由于聚类数量未知, 因此需要选择“无参”聚类算法实现其中聚类分析。这里选择仿射传播(affinity propagation, AP)聚类算法^[30]。若步骤 4 得到的集合 Θ_L 中存在簇中心值小于 ΔT_{th} 的簇, 则认为“未知”电器 x 具有密集波动运行特性。对于固定时长运行特性, 其判断步骤与周期运行特性一致, 聚类分析的对象由“未知”电器 x 的相邻工作循环 $(\Delta T_{on}, \Delta T_{off})$ 替换成工作循环的持续时长 ΔT_{on} , 只需将上述 Step1 至 Step5 中有关符号的下标由“L”替换为“on”即可。

对于具有某种控制规律特性的电器, 考虑到其全部运行记录中可能存在有误辨识结果, 或用户随机操作产生的“噪声”记录, 因此判别阈值 δ_L^{th} 和 δ_{on}^{th} 通常取值小于 100%。根据大量实测经验, 上述判

别算法中, N 取值 5~10, θ_L^h 和 θ_{on}^h 取值 20%~25%, ΔT_{th} 取 50~60s, δ_L^h 和 δ_{on}^h 取值 50%~60%。

1.2 使用规律特性

在本文中, 电器使用规律指电器使用时间的分布规律。结合大量实测数据分析, 不同类型电器的使用受人类活动影响的程度不同, 其使用时间分布具有不同特征。据此, 本文将常见家用电器分为人类活动不相关、强相关以及弱相关 3 类。不相关电器指不受人为因素影响的连续或循环工作的电器; 强相关电器指仅受人为因素影响的电器; 弱相关电器指受环境因素和人为因素综合影响的电器, 此类电器在被用户控制开启后的一段时间内的运行状态主要取决于环境条件, 与强相关电器相比, 人类活动相关程度较弱。

通过比较电器的使用时间分布规律和用户活动强弱的时间分布规律, 便能判断其使用规律特性。现有研究通常使用家庭总口电量数据推断人员活动规律^[31], 然而家庭时段总口电量高低与相应时段在室人员活动强弱没有必然联系。而且, 不同用户的活动规律不同。借助 NILM 获取的单个电器电量数据, 本文给出了一种基于负荷成分变化的用户个性化人类活动强弱时段划分方法, 借以确定“未知”电器的使用规律特性。

假设某家庭在 D 天内学习到 M 个“未知”电器, 用 m 表示电器逻辑编号, $m \in \{1, 2, \dots, M\}$ 。电器 x_m 的使用规律特性判别方法如下:

1) 计算电器分时活动状态。

定义电器某时段累计在线次数与其在线总次数的比值为电器该时段的活动强度 I , 电器 x_m 第 i 小时的活动强度 I_i^m 为

$$I_i^m = n_i^m / (\sum_{h=0}^{23} n_h^m) \quad (3)$$

式中: n_i^m 为电器 x_m 在 D 天内第 i 小时的累计在线次数; $h \in \{0, 1, \dots, 23\}$ 。若电器某天第 i 小时的电量大于 0, 则在线次数加 1。

进而, 根据式(4)判断 M 个“未知”电器的分时活动状态:

$$s_i^m = \begin{cases} 1, & \text{if } I_i^m > I_0 \\ 0, & \text{其它} \end{cases} \quad (4)$$

式中: I_0 表示惯常在线状态判别阈值; $s_i^m = 1$ 表示电器 x_m 在第 i 小时惯常在线, 反之为非惯常在线。

2) 推断人类分时活动状态。

统计第 i 小时惯常在线的电器个数 N_i 和相邻时段惯常在线状态发生变化的电器个数, 即负荷成分变化值 ΔN_i , 计算式为

$$N_i = \sum_{m=1}^M s_i^m \quad (5)$$

$$\Delta N_i = \sum_{m=1}^M |s_{i+1}^m - s_i^m| \quad (6)$$

基于各时段的 N_i 和 ΔN_i , 从时段人类活动状态及其变化趋势两方面, 对一天 24 个小时进行划分。满足 $N_i \leq n_{th1}$ 的时段为人类活动弱时段, 满足 $N_i > n_{th2}$ 的时段为人类活动强时段; 满足 $\Delta N_i > \Delta n_{th}$ 的时段为人类活动增强/减弱时段, 满足 $\Delta N_i \leq \Delta n_{th}$ 的时段为人类活动维持时段。

这里, n_{th1} 和 n_{th2} 是活动状态判别阈值。为了避免设置经验参数, 采用四分位法进行设定, 将特定用户一天内各时段惯常在线电器个数从大到小进行排列, n_{th2} 取值为第三四分位数, n_{th1} 取值为第一四分位数。 Δn_{th} 是活动趋势判别阈值, 直接取值为 αM 。

3) 判别电器使用规律特性

计算电器在不同类型人类活动时段的平均活动强度 \bar{I}_H^m 为

$$\bar{I}_H^m = \sum_{i \in \Omega_H} I_i^m / |\Omega_H| \quad (7)$$

式中: $\Omega_H \subseteq \{0, 1, \dots, 23\}$; Ω_H 表示属于不同人类活动类型的时段集合; $|\Omega_H|$ 表示集合元素个数, $H \in \{R, Q, Z, W\}$ 为活动时段类型代号, R、Q、Z、W 分别指代人类活动弱时段、强时段、增强/减弱时段、维持时段。

按下列判别规则确定 M 个“未知”电器的使用规律特性, 式中, k_1 、 k_2 、 k_{th} 为判别阈值。

- 1) 活动不相关: $\bar{I}_R^m / \bar{I}_Q^m \in [k_1, k_2]$;
- 2) 活动弱相关: $\bar{I}_R^m / \bar{I}_Q^m > k_{th}$ 或 $\bar{I}_W^m / \bar{I}_Q^m > k_{th}$;
- 3) 若上述条件皆不满足, 则为活动强相关。

2 基于粗糙集 NILM 结果自主标注方法

无监督 NILM 结果标注的具体过程是: 依据少量关于电器运行特性的通用先验知识, 利用已积累的“未知”电器的所有运行记录, 从由常见家用电器组成的备选集合中推荐出与其最可能匹配的电器, 并以该电器的名称作为“未知”电器的物理名称, 从而标注该电器的监测结果。在数学上, 自主

标注是确定满足不同类型运行特性条件的多个电器集合的交集的过程。例如，根据“未知”电器的有功功率和无功功率特性分别确定的(候选)电器集合的交集即为根据这两种特性所得的自主标注结果。因此，基于所述通用运行特性可定义新的通用运行特性电器集合。

实际情况中同类电器的少部分实例易于表明某种通用运行特性的取值不确定，不能简单将电器视为具有或不具有此项特性。例如，不同用户的烧水习惯存在差异，烧水习惯相对固定的用户在使用电热水壶烧水时，每次注入的冷水量从统计学角度不会有太大变化，电热水壶具有固定时长运行特性，反之烧水习惯不固定的用户每次注入的冷水量变化较大，此时电热水壶不具备固定时长特性。对此，本文采用粗糙集中集合划分的思想对电器进行粗糙归类，用于确定候选电器集合。

2.1 电器粗糙归类

粗糙集理论^[32]是波兰学者 Pawlak 1982 年提出的一种处理模糊性和不确定性的数学工具，在粗糙集理论中，知识可认为是一族等价关系，它将论域分割成一系列的等价类。基本集定义为由论域中相互间不可分辨的对象组成的集合，是组成论域知识的颗粒。结合电器通用运行特性表，表中电器某列特性取值为 1、0 和 -1 分别表示该电器处于使用此通用运行特性划分得到的电器备选子集内部、边界和外部。利用多种电器通用运行特性(知识)将常见家用电器集合(论域)划分成多个常见家用电器备选子集(基本集)。当备选子集存在边界时为粗糙集，无边界时成为精确集，这是粗糙集与精确集的根本区别。以中国普通家庭中常用电器构成备选集合^[29]，基于通用运行特性的电器粗糙归类样例详见表 1。

2.2 基于粗糙集的两层决策标注方法

本文提出的电器名称自主标注的两层决策方法流程如图 8 所示。具体方案详述如下：

第一决策层利用通用运行特性判别结果，确定“未知”电器所属通用运行特性集合，即在常见家用电器备选集合的基础上缩小标注的候选范围，获得电器备选粗糙子集。

第二决策层利用多种参数特性从缩小的备选子集中确定最有可能的名称标签。针对电器备选粗糙子集，选择有功功率、功率因数、工作循环持续时长和典型使用时段这 4 种泛化能力强的参数特性，建立基于组合投票的决策方法，判别式如下：

表 1 基于通用运行特性的电器粗糙归类

Table 1 General operation characteristics of appliances

常见电器	控制规律特性			使用规律特性		
	周期运行	密集波动	固定时长	不相关	弱相关	强相关
空调	1	-1	1	-1	0	0
洗衣机	1	1	1	-1	-1	1
电暖器	1	-1	0	-1	0	0
电冰箱	1	-1	1	1	-1	-1
电饭锅	1	1	1	-1	-1	1
电磁炉	1	1	1	-1	-1	1
电热水器	-1	-1	0	-1	0	0
微波炉	-1	-1	0	-1	-1	1
电热水壶	-1	-1	0	-1	-1	1
风扇	-1	-1	-1	-1	0	0
电视机	0	0	0	-1	0	0
抽油烟机	-1	-1	-1	-1	-1	1
净水器	-1	-1	-1	-1	-1	1
吹风机	-1	-1	0	-1	-1	1

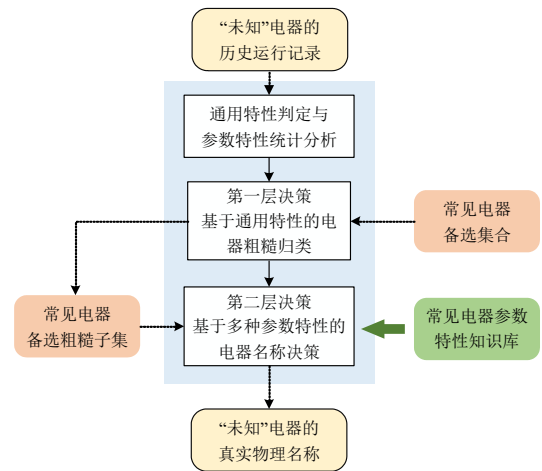


图 8 两层决策方法流程

Fig. 8 Bilevel decision method flow chart

$$A^* = \{ \arg \max (f_{P,A}(p) + f_{\phi,A}(\cos \varphi) + f_{D,A}(d) + f_{T,A}(t)) | \forall A \in \Omega_c \} \quad (8)$$

式中： $p, \cos \varphi, d$ 和 t 分别代表待标注的“未知”电器的有功功率、功率因数、工作循环持续时长及其运行时段典型值； Ω_c 表示电器备选集合；对于任意已知名称的电器 A ， $f_{P,A}(p), f_{\phi,A}(\cos \varphi), f_{D,A}(d)$ 和 $f_{T,A}(t)$ 分别代表根据上述 4 种特性；“未知”电器可能是电器 A 的投票结果； A^* 代表得票数最多的电器组成的集合。

对于 $f_{P,A}(p)$ ，若 $p \in [P_A^{\min}, P_A]$ ，则 $f_{P,A}(p)$ 取 1，得 1 票，否则 $f_{P,A}(p)$ 取 0，不得票， P_A^{\max} 和 P_A^{\min} 分别表示电器 A 的有功功率上下限值； $f_{\phi,A}(\cos \varphi)$ 和 $f_{D,A}(d)$ 同理。 $\arg \max(\cdot)$ 表示括号内表达式取得最大

值时对应自变量的取值。对于 $f_{T_A}(t)$, 若 $t \cap T_A \neq \emptyset$, 则 $f_{T_A}(t)$ 取 1, 得 1 票, 否则, $f_{T_A}(t)$ 取 0, 不得票。根据调研结果^[26-27], 电器参数特性详见表 2。得益于本文使用了电器的使用规律特性, 这里只将一天大致划分为夜间(00:00—05:59)、上午(06:00—11:59)、下午(12:00—17:59)和晚间(18:00—23:59) 4 个自然时段及餐时 5 种典型使用时段。餐时指做餐和用餐时段, 具体是 05:30—08:30, 10:30—13:30 和 16:30—19:30 这 3 个时段。 t 代表参考节 1.2 式(7) 计算得到的电器自然时段活动强度大于 28% 以及餐时活动强度大于 35% 的时段集合。

表 2 常见家用电器的参数特性统计表
Table 2 Parametric operation characteristics of common appliances

电器	有功/W	功率因数	工作循环持续时长	典型使用时段
空调	[500,4000]	(0,0.999]	[1,60]min	00:00—05:59
洗衣机	[80,1000]	(0,0.999]	[1,60]s	06:00—23:59
电暖器	[500,3000]	(0.999,1]	[1,60]min	00:00—05:59
电冰箱	[50,300]	(0,0.999]	[2,30]min	—
电饭锅	[500,3000]	(0.999,1]	[10,30]s	餐时
电磁炉	[800,2500]	(0,0.999]	[1,60]s	餐时
电热水器	[800,5000]	(0.999,1]	[1,15] min	00:00—23:59
微波炉	[500,2500]	(0,0.999]	[0.5,3.5]min	餐时
电热水壶	[800,3000]	(0.999,1]	[2,6]min	06:00—23:59
风扇	[20,100]	(0,0.999]	—	00:00—05:59
电视机	[50,500]	(0,0.999]	—	06:00—23:59
抽油烟机	[50,200]	(0,0.999]	—	餐时
净水器	[800,2000]	(0.999,1]	—	06:00—23:59
吹风机	[300,2000]	(0.999,1]	—	06:00—23:59

值得说明的是, 一般情况下, 集合 A^* 仅含一个元素, 特殊情况下, 若集合 A^* 中元素个数多于 1 个, 根据上述特性的定义, 这 2 种(或以上)电器极为相似, 需要由用户或专家干预筛选标注结果。

3 实验分析

为验证本文方法的有效性, 设计 0 所示的实验方案。图 9 中右侧为“未知”电器监测结果自主标注流程, 左侧为无监督 NILM 人工协助获取分解结果真实名称的流程。实验在私有数据集和公开数据集 REDD^[33] 上进行, 无监督 NILM 方法选用文献[21] 中的方法。据实测经验, 通用运行特性判别涉及到的参数在一定范围内取值(取值范围见 2.2 节)对最终标注结果影响不大, 此处参数 N 取值 10, θ_L^{th} 和 $\theta_{\text{on}}^{\text{th}}$ 取值 20%, ΔT_{th} 取 50s, δ_L^{th} 和 $\delta_{\text{on}}^{\text{th}}$ 取值 50%。 k_1 、

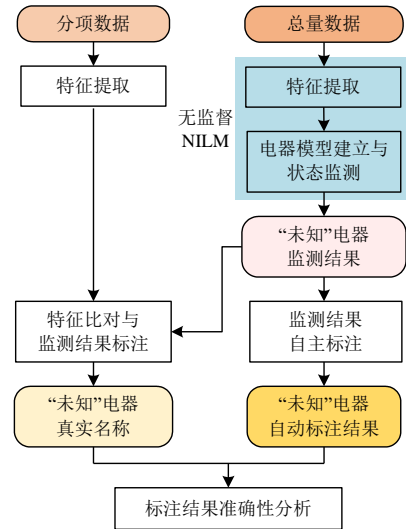


图 9 实验流程图

Fig. 9 Overall experiment flow chart

k_2 分别取值 0.8、1.2, k_{th} 取 3, α 取 0.1。下面给出测试结果及分析讨论。限于篇幅, 下面仅以一个场景为例对自主标注过程进行详细介绍, 其他场景直接给出关键中间结果和最终标注结果。

3.1 私有数据集实验分析

基于文献[21], NILM 系统通过“自学习”分别在未知场景 1 和 2 中陆续发现若干“未知”电器, 根据所积累的运行记录可推导出它们的运行特性, 本文方法据此为每个电器给出了“推荐名称”。

场景 1 是一个有老人和儿童的五口之家。下面对此场景自主标注过程进行详细分析。首先根据 1.1 节计算电器的控制规律特性, 结果详见表 3。

表 3 电器控制规律特性
Table 3 Control characteristics of appliances

逻辑名称	固定时长/周期运行		$(\Delta T_{\text{on}}, \Delta T_{\text{off}})$ 均值	控制规律特性
	$\delta_{\text{on}}/\%$	$\delta_L/\%$		
1	77.5	15.79	—	FO
2	60.45	27.18	—	FO
3	62.71	8.14	—	FO
4	79.42	65.77	(11, 10)s	PO+FO+IO
5	84.03	72.27	(17, 12)s	PO+FO+IO
6	89.89	88.06	(7, 15)min	PO+FO
7	99.89	63.05	(4, 3)s	PO+FO+IO
8	91.52	80.05	(14, 35)min	PO+FO

然后, 按照节 1.2 方法计算每小时惯常在线电器个数和负荷成分变化值, 见图 10, 图中虚线对应的纵坐标值为判别阈值。进而, 计算各电器在不同类型时段活动强度与电器使用规律特性, 详见表 4。

表 5 统计了电器在自然时段以及餐时的活动强度用于计算电器的典型使用时段, 用红色填充标识

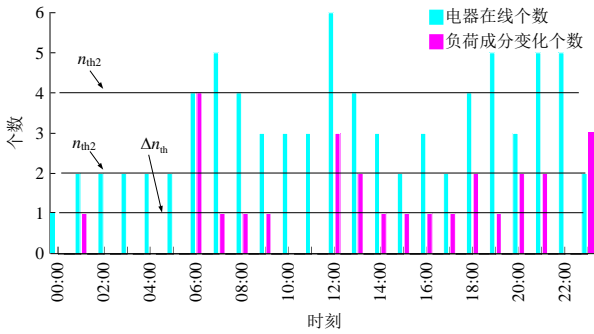


图 10 场景 1 家庭日时段电器构成柱状图

Fig. 10 Scenario 1 daily appliances histogram

表 4 电器使用规律特性

Table 4 Usage characteristics of appliances

逻辑名称	分时活动强度/%				使用特性
	强时段	弱时段	增强时段	维持时段	
1	8.03	0.00	4.76	4.01	强相关
2	5.10	2.05	2.54	6.82	弱相关
3	6.64	1.16	4.53	3.78	强相关
4	8.02	0.36	4.76	6.97	强相关
5	7.84	0.22	3.49	3.77	强相关
6	2.68	6.94	2.52	7.75	弱相关
7	7.95	0.54	10.88	2.31	强相关
8	4.36	3.93	3.98	3.89	不相关

表 5 各电器时段活动强度

Table 5 Activity indices of appliances

逻辑名称	%				
	夜间	上午	下午	晚间	餐时
1	7.68	0	15.40	76.92	56.4
2	20.52	18.18	13.60	47.7	20.94
3	7.62	22.5	19.66	50.22	44.64
4	5.88	21.18	41.20	31.74	43.92
5	7.2	31.56	18.28	42.96	37.02
6	59.34	15.06	6.73	18.78	13.56
7	3.06	16.44	15.04	65.46	62.34
8	23.04	25.92	27.16	23.88	32.28

该时段为该电器的典型使用时段之一。表 6 统计了各电器参数特性的典型值。表 7 中记录了最终标注结果，可以看出，本文对其中常用电器的推荐名称都是正确的，表中绿色填充为第一层决策得到的电器候选集合，其中的数字为候选集合中各电器在第二层决策的得票数，票数最高的电器为最终标注结果。为了方便表示，用 A~N 来表示电器代码。根据表 6 中运行时长均值的分析结果，得知此家庭中微波炉有两种常用的使用模式。

场景 2 是一个普通工薪阶层的二口之家。本文对其中大部分常用电器的推荐名称如表 8 是正确的。对于“未知电器 1”（实际是电冰箱），由于家中存在与其功率接近的电器（经核实是电灯），因此

表 6 电器参数特性

Table 6 Parametric operation characteristics of appliances

逻辑名称	功率均值/W	cosφ	ΔT _{on} 均值	典型使用时段
1	1603	0.996	30s; 90s	晚间；餐时
2	3280	1	10min	晚间
3	2057	1	5min	晚间；餐时
4	1119	1	11s	下午；晚间；餐时
5	341	0.452	17s;370s	上午；晚间；餐时
6	809	0.922	7min	夜间
7	1190	0.998	4s	晚间；餐时
8	82	0.932	14min	—

表 7 场景 1 中电器标注结果

Table 7 Labeling results of appliances in scenario 1

常见家用电器列表	逻辑名称							
	1	2	3	4	5	6	7	8
空调	×	×	×	×	×	4	×	×
洗衣机	×	×	×	2	4	×	3	×
电暖器	×	×	×	×	×	3	×	×
电冰箱	×	×	×	×	×	2	×	4
电饭锅	×	×	×	4	2	×	2	×
电磁炉	×	×	×	3	2	×	4	×
电热水器	2	4	3	×	×	×	×	×
微波炉	4	2	2	×	×	×	×	×
电热水壶	2	3	4	×	×	×	×	×
风扇	×	×	×	×	×	×	×	×
电视机	×	×	×	×	×	×	×	×
抽油烟机	×	×	×	×	×	×	×	×
净水器	×	×	×	×	×	×	×	×
吹风机	2	2	2	×	×	×	×	×
推荐名称	h	g	i	e	b	a	f	d
真实名称	h	g	i	e	b	a	f	d

负荷分解所得的“未知电器 1”的运行记录并不纯净，使得 δ_L 结果不足 50%，无法判别其为周期运行电器，即便如此其参数特性与电冰箱更相近，最终推荐名称是正确的。对于“未知电器 4”（实际是电饼铛），根据其通用运行特性和参数特性，可判断其为典型使用时段为餐时的周期运行纯阻性电器，由于候选电器表 2 中不存在电饼铛，因此最终结果为与其功能用途最相近的电器名称电磁炉。从表 8 也可以看出此家庭习惯在餐时使用电热水壶烧水。

3.2 公开数据集实验分析

本文选取 REDD 数据集进行测试，REDD 数据集包含 6 个美国家庭，选取较为复杂的 house1 和 house6 在两周内的数据进行测试，实验结果如下。

House1: 文献[21]方法学习出了 8 个“未知”

表8 场景2的标注结果
Table 8 Labeling results in scenario 2

逻辑名称	固定时长/固定间隔 $\delta_{on}/\%$ $\delta_l/\%$	$(\Delta T_{on}, \Delta T_{off})$ 均值	使用规律	功率 均值/W	$\cos\phi$	ΔT_{on} 均值	典型使用时段	推荐名称	真实名称	
1	83.06	43.29	—	不相关	55	0.936	30min	—	电冰箱	电冰箱
2	98.83	93.89	(7, 11)s	强相关	863	1	7s	晚间; 餐时	电饭锅	电压力锅
3	80.11	13.59	—	弱相关	2247	1	20min	上午; 晚间	电热水器	电热水器
4	90.33	62.77	(95, 76)s (95, 208)s	强相关	1365	1	95s	下午; 晚间; 餐时	电磁炉	电饼铛
5	32	12.5	—	强相关	538; 1082	1	-	下午; 晚间	电吹风	电吹风
6	85.62	27.6	—	强相关	1733	1	6min	上午; 下午; 餐时	电热水壶	电热水壶

电器。本文对其中部分电器的推荐名称是合理的。其中,洗碗机存在2种运行模式,分别被学习成“未知电器4”和“7”2种电器,因公开数据集中缺少功率因数信息,根据可获取的运行特性“未知电器4”被标注成在表1中与其“最相近”的电磁炉。对于“未知电器7”和其他部分电器,由于第二层决策电器候选集合内各电器组合投票的得票数少于2票,本文将它们都标注为“陌生电器”。此外,从表9典型使用时段信息可以看出此家庭电吹风存在2种典型使用模式。

House6: 文献[21]方法学习出了5个“未知”

电器。本文对其中冰箱、电热水壶、空调的推荐名称都是正确的。据典型使用时段分析结果可看出,此用户习惯在晚间使用电热水壶烧水,见表10。

在REDD数据集中测试发现,基于对中国家庭电器运行特性调研得到的特性知识库,能对部分在中国家庭同样常见的电器准确标注,不常见家用电器则未能给出其正确物理名称。推测其原因在于中西方生活习惯及地域环境气候等条件不同,常用家电种类存在差异。因此在开展无监督NILM结果标注工作时,有必要针对不同国家或地区建立不同的知识库,以保证NILM结果自主标注的准确性。

表9 REDD House 1的标注结果
Table 9 Labeling results in REDD House1

逻辑名称	固定时长/周期运行 $\delta_{on}/\%$ $\delta_l/\%$	$(\Delta T_{on}, \Delta T_{off})$ 均值	使用规律	功率 均值/W	ΔT_{on} 均值	典型使用时段	推荐名称	真实名称	
1	78.95	0	—	强相关	1572	3.4min; 1min	上午; 餐时	电吹风	bathroom_gfi
2	68.42	66.67	(1.3, 3.2)min	强相关	4139	1.3min	下午; 晚间; 餐时	陌生电器	oven
3	73.22	0	—	不相关	218	15min; 18min	—	冰箱	refrigerator
4	84.44	55.56	(4.4, 1.8)min (12, 22)s	强相关	1081	4.4min; 12s	上午; 餐时	电磁炉	dishwasher
5	57.3	0	—	强相关	1540	57.47s	上午; 晚间; 餐时	微波炉	microwave
6	66.67	40	(3.1min, 36.7s)	强相关	5224	3.1min	上午; 下午; 餐时	陌生电器	oven
7	64	52.31	(47, 22.4)s	强相关	209	47s; 4.8min	晚间	陌生电器	dishwasher
8	50	0	—	强相关	736	112.2s; 52.75s	上午; 下午; 餐时	陌生电器	washer_dryer

表10 REDD House6的标注结果
Table 10 Labeling results in REDD House6

逻辑名称	固定时长/周期运行 $\delta_{on}/\%$ $\delta_l/\%$	$(\Delta T_{on}, \Delta T_{off})$ 均值	使用规律	功率 均值/W	ΔT_{on} 均值	典型使用时段	推荐名称	真实名称	
1	52.52	74.76	(29, 37) min (33, 41) min	不相关	183	29min; 33min	—	电冰箱	refrigerator
2	83.33	0	—	强相关	1438	42s	晚间	电热水壶	kitchen_outlets
3	72.55	0	—	强相关	205	26.7s; 2.4min	晚间	陌生电器	outlets_unknown
4	60	0	—	弱相关	1937	9.8min; 3.8min	夜间	空调	air_conditioning
5	0	0	—	强相关	163	—	上午; 晚间	电视机	outlets_unknown

4 结论

无监督 NILM 方法可通过直接处理无标注的负荷用电量数据实现负荷分解, 而这类方法通常需要人工标注分解得到的电器级别用电信息是何种具体电器产生的, 阻碍了 NILM 实用化推广。本文提出一种基于电器粗糙归类的无监督 NILM 结果自主标注方法, 它依据少量关于电器运行特性的先验知识, 利用已建模“未知”电器的 NILM 结果自动为其赋予物理名称, 该方法可与任意无监督 NILM 方法集成使之变为完全无监督 NILM 方案。

针对无监督 NILM 结果自主标注, 定义控制规律与使用规律 2 种通用运行特性, 其物理含义清晰、易于理解, 采用通用运行特性有效地解决参数特性泛化能力弱的问题, 提高本文方法对同类电器品牌和型号的多样性, 以及实际环境中不同用户行为的千差万别的适应性。同时, 给出 2 种通用运行特性的自适应判别方法, 测试结果表明所提方法能对电器通用运行特性做出准确判别。再者, 为应对实际场景中同类电器关于某种通用运行特性的差异性, 利用粗糙集理论进行电器粗糙归类, 并在此基础上提出融合通用运行特性的电器名称分层决策方法, 在私有与公开数据集集中的实验测试表明本文方法能够在使用少量参数特性的情况下, 实现不同场景 NILM 结果的准确标注。

未来需深入挖掘和设计更多有效的通用运行特性, 以及拓展和丰富参数特性知识库, 进一步提高标注准确性和合理性。

致谢

本文实验所用私有数据的采集与整理工作是在天津求实智源科技有限公司的刘浩、杜伟强、马骁等工作人员的大力支持下完成的, 在此向他们表示衷心的感谢。

参考文献

- [1] 余贻鑫, 刘博, 栾文鹏. 非侵入式居民电力负荷监测与分解技术[J]. 南方电网技术, 2013, 7(4): 1-5.
YU Yixin, LIU Bo, LUAN Wenpeng. Nonintrusive residential load monitoring and decomposition technology [J]. Southern Power System Technology, 2013, 7(4): 1-5(in Chinese).
- [2] 刘博, 栾文鹏. 基于负荷分解的用电数据云架构方案及应用场景[J]. 电网技术, 2016, 40(3): 791-796.
LIU Bo, LUAN Wenpeng. Conceptual cloud solution architecture and application scenarios of power consumption data based on load disaggregation[J]. Power System Technology, 2016, 40(3): 791-796(in Chinese).
- [3] 窦迅, 王俊, 王湘艳, 等. 基于演化博弈的区域电气互联综合能源系统用户需求侧响应行为分析[J]. 中国电机工程学报, 2020, 40(12): 3775-3785.
DOU Xun, WANG Jun, WANG Xiangyan, et al. Analysis of user demand side response behavior of regional integrated power and gas energy systems based on evolutionary game[J]. Proceedings of the CSEE, 2020, 40(12): 3775-3785(in Chinese).
- [4] 陆青, 郁浩, 冷亚军, 等. 家庭智能用电任务调度优化模型及其算法研究[J]. 中国电机工程学报, 2018, 38(13): 3826-3836.
LU Qing, YU Hao, LENG Yajun, et al. Research on model and algorithm of smart electricity consumption task scheduling optimization in household[J]. Proceedings of the CSEE, 2018, 38(13): 3826-3836(in Chinese).
- [5] 别朝红, 林超凡, 李更丰, 等. 能源转型下弹性电力系统的发展与展望[J]. 中国电机工程学报, 2020, 40(9): 2735-2745.
BIE Zhaohong, LIN Chaofan, LI Gengfeng, et al. Development and prospect of resilient power system in the context of energy transition[J]. Proceedings of the CSEE, 2020, 40(9): 2735-2745(in Chinese).
- [6] 阮前途, 谢伟, 许寅, 等. 韧性电网的概念与关键特征[J]. 中国电机工程学报, 2020, 40(21): 6773-6784.
RUAN Qiantu, XIE Wei, XU Yin, et al. Concept and key features of resilient power grids[J]. Proceedings of the CSEE, 2020, 40(21): 6773-6784(in Chinese).
- [7] WANG He, HOU Yongshan, YU Huanan. Analysis of integrated energy-load characteristics based on sparse clustering and compressed sensing[J]. IET Energy Systems Integration, 2019, 1(3): 194-201.
- [8] 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(in Chinese).
- [9] 余贻鑫, 栾文鹏. 智能电网述评[J]. 中国电机工程学报, 2009, 29(34): 1-8.
YU Yixin, LUAN Wenpeng. Smart grid and its implementations[J]. Proceedings of the CSEE, 2009,

- 29(34): 1-8(in Chinese).
- [10] HART G W. Nonintrusive appliance load monitoring [J]. Proceedings of the IEEE, 1992, 80(12): 1870-1891.
- [11] CUI Gaochen, LIU Bo, LUAN Wenpeng, et al. Estimation of target appliance electricity consumption using background filtering[J]. IEEE Transactions on Smart Grid, 2019, 10(6), 5920-5929.
- [12] GIRI S, BERGÉS M. An energy estimation framework for event-based methods in Non-Intrusive Load Monitoring [J]. Energy Conversion and Management, 2015, 90: 488-498.
- [13] BASU K, DEBUSSCHERE V, BACHA S, et al. A generic data driven approach for low sampling load disaggregation[J]. Sustainable Energy, Grids and Networks, 2017, 9: 118-127.
- [14] 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.
- 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(in Chinese).
- [15] 孙毅, 崔灿, 陆俊, 等. 基于遗传优化的非侵入式家居负荷分解方法[J]. 电网技术, 2016, 40(12): 3912-3917.
- SUN Yi, CUI Can, LU Jun, et al. A non-intrusive household load monitoring method based on genetic optimization[J]. Power System Technology, 2016, 40(12): 3912-3917(in Chinese).
- [16] 祁兵, 程媛, 武昕. 基于 Fisher 有监督判别非侵入式居民负荷辨识方法[J]. 电网技术, 2016, 40(8): 2484-2490.
- QI Bing, CHENG Yuan, WU Xin. Non-intrusive household appliance load identification method based on fisher supervised[J]. Power System Technology, 2016, 40(8): 2484-2490(in Chinese).
- [17] 陈思运, 高峰, 刘焯, 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(in Chinese).
- [18] ZHAO Bochao, STANKOVIC L, STANKOVIC V. On a training-less solution for non-intrusive appliance load monitoring using graph signal processing[J]. IEEE Access, 2016, 4: 1784-1799.
- [19] WANG Lanruo, LUO Xianjue, ZHANG Wei. Unsupervised energy disaggregation with factorial hidden Markov models based on generalized backfitting algorithm[C]//Proceedings of the 2013 IEEE International Conference of IEEE Region 10(TENCON 2013). Xi'an: IEEE, 2013: 1-4.
- [20] JOHNSON M J, WILLISKY A S. Bayesian nonparametric hidden semi-Markov models[J]. The Journal of Machine Learning Research, 2013, 14(1): 673-701.
- [21] LIU Bo, YU Yinxin, LUAN Wenpeng, et al. An unsupervised electrical appliance modeling framework for non-intrusive load monitoring[C]//Proceedings of the 2017 IEEE Power & Energy Society General Meeting. Chicago: IEEE, 2017: 1-5.
- [22] KONG Weicong, DONG Zhaoyang, HILL D J, et al. A hierarchical hidden markov model framework for home appliance modeling[J]. IEEE Transactions on Smart Grid, 2018, 9(4): 3079-3090.
- [23] WELIKALA S, DINESH C, GODALIYADDA R I, et al. Robust non-intrusive load monitoring(NILM) with unknown loads[C]//Proceedings of the 2016 IEEE International Conference on Information and Automation for Sustainability. Galle: IEEE, 2016: 1-6.
- [24] KAMOTO K M, LIU Qi, LIU Xiaodong. Unsupervised energy disaggregation of home appliances[C]//Proceedings of the Third International Conference, ICCCS 2017. Nanjing: Springer, 2017: 398-409.
- [25] LIU Bo, LUAN Wenpeng, YU Yixin. Dynamic time warping based non-intrusive load transient identification [J]. Applied Energy, 2017, 195: 634-645.
- [26] 刘博. 非侵入式电力负荷监测与分解技术[D]. 天津: 天津大学, 2014.
- LIU Bo. Non-intrusive power load monitoring and disaggregation technique[D]. Tianjin: Tianjin University, 2014(in Chinese).
- [27] ELBE C, SCHMAUTZER E. Appliance-specific usage patterns for load disaggregation methods[C]//

Internationale Energiewirtschaftstagung An der TU Wien(IEWT 2013). 2013: 1-6.

- [28] WANG Zhenyu, ZHENG Guilin. Residential appliances identification and monitoring by a nonintrusive method[J]. IEEE transactions on Smart Grid, 2012, 3(1): 80-92.
- [29] DONG Ming, MEIRA P C M, XU W, et al. Non-intrusive signature extraction for major residential loads[J]. IEEE Transactions on Smart Grid, 2013, 4(3): 1421-1430.
- [30] FREY B J, DUECK D. Clustering by passing messages between data points[J]. Science, 2007, 315(5814): 972-976.
- [31] SONG K, KWON N, ANDERSON K, et al. Predicting hourly energy consumption in buildings using occupancy-related characteristics of end-user groups[J]. Energy and Buildings, 2017, 156: 121-133.
- [32] 王国胤, 姚一豫, 于洪. 粗糙集理论与应用研究综述[J]. 计算机学报, 2009, 32(7): 1229-1246.
- WANG Guoyin, YAO Yiyu, YU Hong. A survey on rough set theory and applications[J]. Chinese Journal of Computers, 2009, 32(7): 1229-1246(in Chinese).
- [33] KOLTER J Z, JOHNSON M J. REDD: a public data set for energy disaggregation research[C]//Workshop on Data Mining Applications in Sustainability(SIGKDD). San Diego: ACM, 2011: 59-62.



肖潇

在线出版日期: 2021-03-16。

收稿日期: 2021-01-28。

作者简介:

肖潇(1996), 女, 硕士研究生, 主要研究方向为非侵入式负荷监测与分解、监测结果自主标注、电力数据分析、居民用户用电行为分析等, xiaox@tju.edu.cn;

栾文鹏(1964), 男, 教授, 博士生导师, 国家特聘专家, IEEE 高级会员, 主要研究方向为电力系统规划与分析、电力数据分析、电力负荷监测、智能电网及物联网技术应用等, wenpeng.luan@tju.edu.cn;

*通信作者: 刘博(1988), 男, 讲师, 博士生导师, 主要研究方向为非侵入式负荷监测与分解、配用电大数据分析与应用、智能电网高级量测体系与电力物联网等, liubo@tju.edu.cn;

王岩(1985), 女, 硕士, 从事非侵入式负荷分解与监测方面的数据分析与技术应用, wangy@transenergy-tech.com;

刘子帅(1997), 男, 硕士研究生, 主要研究方向包括智能配电网、非侵入式负荷监测和机器学习等, zishuailiu@tju.edu.cn;

韦尊(1995), 男, 博士研究生, 主要研究方向包括负荷分解和监控、数据挖掘以及大数据分析和应用等, weizun22@163.com。

(编辑 乔宝榆, 李新洁)

Autonomous Labeling of Unsupervised NILM Results Based on Rough Classification of Appliances

XIAO Xiao¹, LUAN Wenpeng¹, LIU Bo^{1*}, WANG Yan², YANG Jinnan¹, LIU Zishuai¹, WEI Zun¹

(1. School of Electrical Engineering & Automation, Tianjin University; 2. Tianjin Qiushi Transenergy Technologies Co., Ltd.)

KEY WORDS: unsupervised non-intrusive load monitoring (NILM); autonomous appliance name labeling; general operation characteristics; adaptive discrimination; rough classification of appliances

Compared to supervised non-intrusive load monitoring (NILM), the unsupervised NILM schemes are scalable and more appropriate for large-scale implementation.

In unsupervised NILM, the appliance model established through unsupervised learning usually has no physical name. In order to achieve widespread implementation of NILM and enable the associated advanced applications, it is necessary to use general prior knowledge to autonomously label the physical names of the learnt appliances, referred to as autonomous appliance name labeling, thus to realize fully unsupervised NILM, which has a framework as summarized as Fig. 1.

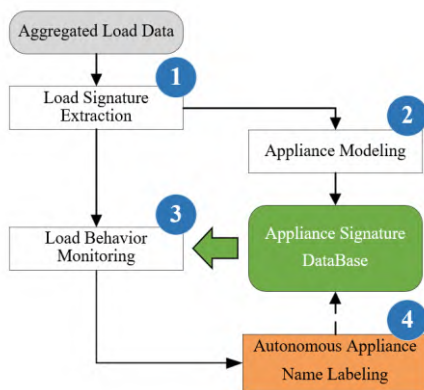


Fig. 1 Fully unsupervised NILM framework diagram

In this paper, a novel method for autonomous labeling of unsupervised NILM results based on rough classification of appliances is proposed. Both operation characteristics and usage time distribution of appliances are analyzed and summarized. Furthermore, the appliances are roughly classified according to the defined operation and usage characteristics based on the

rough set theory, and then a bilevel decision-making method is proposed to realize the NILM result labeling. The bilevel decision-making flow chart is shown in Fig. 2 with the details as follows:

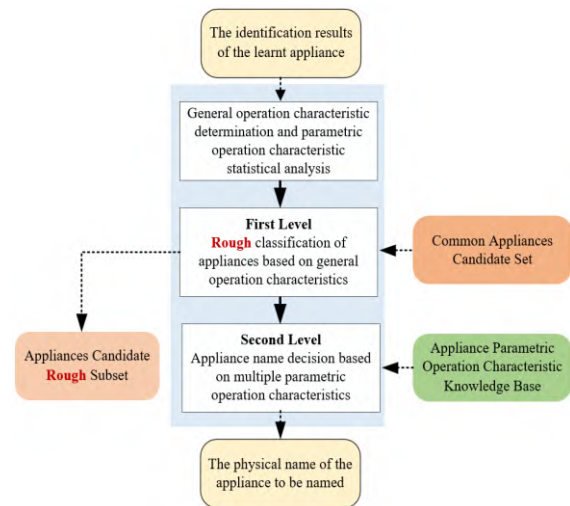


Fig. 2 Bilevel decision-making method flow chart

In the first level of decision-making, the general operation characteristics of the appliance to be named are analyzed to determine the candidate rough subset to which it belongs.

In the second level of decision-making, multiple parametric operation characteristics are used to determine the most likely appliance name tag from the candidate rough subset. Four parametric operation characteristics are chosen for decision making based on the following combined voting evaluation equation.

$$A^* = \{ \arg \max (f_{P,A}(p) + f_{\phi,A}(\cos \phi) + f_{D,A}(d) + f_{T,A}(t) | \forall A \in \mathcal{Q}_c \} \quad (1)$$

Experiments in private and public datasets show that the proposed method can label NILM results of common appliances in different scenarios accurately.