

## 第六届“泰迪杯”数据挖掘挑战赛——

### A 题：基于非侵入式负荷检测与分解的电力数据挖掘

#### 一、背景

区别于常规的连接在入户线上的电能表只能获得总能耗数据，电力分项计量可以对连接到入户线后的建筑物内各个用电设备所消耗的电能进行独立计量。电力分项计量对于电力公司准确预测电力负荷、科学制定电网调度方案、提高电力系统稳定性和可靠性有着重要意义；对用户而言可以帮助用户了解用电设备的使用情况、提高用户的节能意识、促进科学合理用电。另一方面，分项计量还可以针对能耗数据来实现设备老化、故障预警的功能，在我们的生产、生活中有非常实际的意义。同时，基于电力分项计量的一系列技术，是将电器识别作为物联网的一个重要研究方向，从物理层面去真正查看、操控物联网设备，达到绝对可信的要求，这是未来物联网的目标。

分项计量技术目前主要分为两种：一种在总负荷内部为每个用电设备配备带有数字通信功能的传感器，通过通讯网络采集各用电设备的用电信息，这种方式称为侵入式电力负荷监测（intrusive residential load monitoring, ILM）；另一种如图 1 所示，仅在电网的用户入口处安装一个传感器，通过采集和分析用户用电总功率或总电流来监测每个或每类用电设备的用电功率和工作状态，从而了解用户家中每个或每类用电设备的耗电情况和用电规律，这种方式称作非侵入式电力负荷监测与分解（non-intrusive load monitoring and decomposition, NILMD）。基于 NILMD 技术的用电分析计量具有简单、经济、可靠和易于迅速推广应用等优势，更加适用于居民用户。

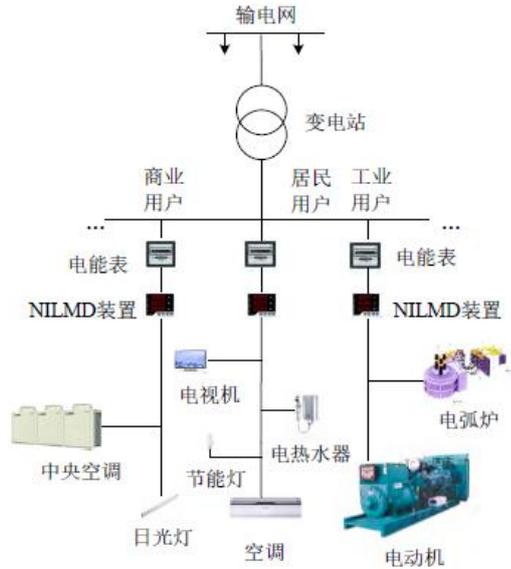


图 1 非侵入式电力负荷监测与分解系统示意图

NILMD 装置量测得到的是整个线路上的电压、电流数据，它们可以看作是各个用电设备的电压、电流数据的叠加。NILMD 的核心，是如何从采集到的整条线路的电压、电流数据中“分解”出每个用电设备独立的用电数据。

就像人类的声纹、指纹等生物特征具有唯一性可以用来实现个体识别一样，不同种类和型号的用电设备在运行过程中产生的电压、电流以及谐波等时序数据中也有相对稳定的较为显著的特征，称之为用电设备的负荷印记 (load signatures, LS)。而根据用电设备运行的过程，又可将数据分为暂态数据和稳态数据两大类，其中暂态数据主要指设备启动、设备停止、设备模式切换时的状态数据，稳态数据主要指设备稳定运行时的状态数据。

图 2 所示的电器启动暂态功率波形就是一种典型的 LS。因工作原理不同，荧光灯和电动机的暂态行为明显不同。

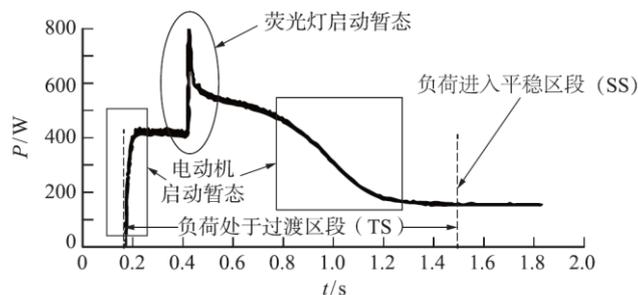


图 2 电器启动暂态功率波形是一种典型的 LS

NILMD 系统的目标是根据不同类型用电器独特的负荷印记从一个能源网关设备记录的数据中检测出接入该设备的电路中各种用电器的开关等操作，并对其用电量进行分项计量。

用电设备根据用户用电设备工作状态的不同，可分为以下 3 种类型：

(1) 启/停二状态设备 (ON/OFF)。这类用电设备只有运行和停机两种用电状态，如白炽灯、电热水壶等。

(2) 有限多状态设备。这类用电设备通常具有有限多个分立的工作状态，与之相对应的用电功率间是离散的，不同的功率水平即标志着不同的工作状态，如洗衣机、微波炉、电磁炉等。

(3) 连续变状态设备。这类用电设备的稳态区段功率无恒定均值，而是在一个范围内连续变动，如变频空调、电动缝纫机等。

## 二、 问题

请根据附件提供的某家庭用电数据完成以下任务：

1. 根据附件 1 提供的单一态数据，分析并给出各用电设备的运行特征，并估计各用电设备的实时用电量（文件名为 energy1.xlsx，表名分别为 YD1、YD2、……、YD11，结果格式参见表 6，其中各变量的含义见表 7）。

2. 根据附件 2 中的单一态数据，设计自动识别单一设备的数学模型和计算方法，并估计这一用电设备的实时用电量（文件名为 energy2.xlsx，表名分别为设备 1、设备 2、设备 3，结果格式参见表 6）。

3. 根据附件 3 提供的用电设备实测数据，设计方法确定各用电设备的状态、操作及操作时间（文件名为 operation3.xlsx，表名分别为设备组 1、设备组 2、设备组 3，结果格式参见表 5），并估计每个用电设备的实时用电量（文件名为 energy3.xlsx，表名分别为设备组 1、设备组 2、设备组 3，结果格式参见表 6）。

4. 利用问题 3 设计的方法，根据附件 4 提供的用电设备实测数据，识别出各用电设备及其状态、操作及操作时间（文件名为 operation4.xlsx，表名分别为设备组 1、设备组 2、设备组 3，结果格式参见表 5），并估计每个用电设备的

实时用电量（文件名为 energy4.xlsx，表名分别为设备组 1、设备组 2、设备组 3，结果格式参见表 6），无法识别的用电设备标记为 YD0。

### 三、数据说明

#### 1. 用电设备及工作参数

表 1 用电设备及工作参数

序号	设备 ID	设备类型	工作参数
1	YD1	奥克斯风扇	220V, 60W
2	YD2	美的微波炉	220V, 输入: 1150W, 输出: 700W
3	YD3	九阳热水壶	220V, 1800W
4	YD4	ThinkPad 笔记本电脑	20V, 3.25A/4.5A
5	YD5	白炽灯	220V, 40W
6	YD6	节能灯	220V, 5W
7	YD7	FUJI 激光打印机	220~240V, 50~60Hz, 4.6A
8	YD8	饮水机	220V, 制热: 430W, 制冷: 70W, 总: 500W
9	YD9	挂式空调	220V, 2600W
10	YD10	奔腾电吹风	220V, 50Hz, 1400W
11	YD11	创维电视机	220V, 50Hz, 150W

#### 2. 附件中表格的数据结构

数据分为用电设备单一态数据（附件 1~2）和叠加态数据（附件 3~4）两部分，两部分数据的数据结构完全相同，包括通过使用能源网关设备在入户线上测量的电流、电压、有功功率、无功功率和谐波数据。每一类数据都包含以下 4 张表：设备数据（表 2）、周波数据（表 3）、谐波数据（表 4）以及操作记录（表 5）。设备数据、周波数据和谐波数据的采样间隔是 1 秒，为了节省存储空间和数据传输时间，变化很小的数据不予传输和存储。

单一态数据记录单个用电设备的用电数据，根据用电器类型分为 2 种情况：(1) 回路中的用电设备仅有开关两种状态；(2) 回路中的用电设备具有不同挡位或者不同工作状。

叠加态数据记录分别记录回路中接入 2~5 种用电器运行一段时间的情况。

表 2 设备数据

序号	字段	说明
1	time	年月日时分秒
2	IC	电流, 单位: 0.001 A
3	UC	电压, 单位: 0.1 V
4	PC	有功功率, 单位: 0.0001 KW
5	QC	无功功率, 单位: 0.0001 KVar
6	PFC	功率因数, 单位: %
7	P	总有功功率, 单位: 0.0001 KW
8	Q	总无功功率: 单位: 0.0001 KVar
9	PF	总功率因数, 单位: %

表 3 周波数据

序号	字段	说明
1	time	年月日时分秒
2	IC001	电流一个周波的第 1 个采样点
⋮	⋮	⋮
129	IC128	电流一个周波的第 128 个采样点
130	UC001	电压一个周波的第 1 个采样点
⋮	⋮	⋮
257	UC128	电压一个周波的第 128 个采样点

注:我国交流电供电的标准频率为 50 Hz,NILMD 装置在其中一个周期内(0.02 秒)采集 128 个时间点上的数据。

表 4 谐波数据

序号	字段	说明
1	time	年月日时分秒
2	IC02	2 次电流谐波, 表示谐波的含有率, 单位: %
⋮	⋮	⋮
51	IC51	51 次电流谐波, 表示谐波的含有率, 单位: %
52	UC02	2 次电压谐波, 表示谐波的含有率, 单位: %
⋮	⋮	⋮
101	UC51	51 次电压谐波, 表示谐波的含有率, 单位: %

表 5 用电设备操作记录

序号	时间	设备	工作状态	操作
1	2018/1/25 14:55:44	YD1	关闭	
2	2018/1/25 14:56:43	YD1	1 档	启动
3	2018/1/25 14:59:38	YD1	2 档	切换
4	2018/1/25 15:03:31	YD1	3 档	切换
5	2018/1/25 15:06:59	YD1	关闭	关闭

### 3. 实时用电量结果文件的数据结构。

表 6 实时用电量

ID	time	W
YD1	2018/2/1 10:00:00	2.5000
YD1	2018/2/1 10:00:01	2.5025
YD1	2018/2/1 10:00:02	2.5050
YD1	2018/2/1 10:00:03	2.4050
YD1	2018/2/1 10:00:04	2.4015

表 7 实时用电量表中字段说明

序号	字段	说明
1	ID	识别出的设备 ID
2	time	年月日时分秒
3	W	耗电量，单位：0.001 KWh（保留 4 位小数）

## 四、名词解释及计算公式

### 1. 有功功率、视在功率、无功功率、功率因素

有功功率是保持用电设备正常运行所需的电功率，也就是将电能转换为其他形式能量（机械能、光能、热能）的电功率。例如：5.5 千瓦的电动机就是把 5.5 千瓦的电能转换为机械能，带动水泵抽水或脱粒机脱粒；各种照明设备将电能转换为光能，供人们生活和工作照明。

无功功率是用于电路内电场与磁场的交换，并用来在电气设备中建立和维持磁场的电功率。它不对外作功，而是转变为其他形式的能量。凡是有电磁线圈的电气设备，要建立磁场，就要消耗无功功率。例如：40 瓦的日光灯，除需 40 多瓦有功功率（镇流器也需消耗一部分有功功率）来发光外，还需 80 乏左右的无功功率供镇流器的线圈建立交变磁场用。由于它不对外做功，才被称之为“无功”。

视在功率等于电压有效值与电流有效值乘积，表示电源的输出能力。

功率因素定义为有功功率与视在功率的比值，由电压与电流之间的相位差角  $\varphi$  决定。

视在功率计算公式： $S = UI$ ，单位：VA（伏安）；

有功功率计算公式： $P = UI \cos \varphi = S \cos \varphi$ ，单位：W（瓦）；

无功功率计算公式： $Q = UI \sin \varphi$ ，单位：var（乏）；

其中：U 为电压，I 为电流， $\cos \varphi$  为 P/S，即功率因素。

从公式中可以看出，视在功率、有功功率、无功功率的单位都是伏安，为了分别区分这 3 个单位，无功功率使用的是英文无功伏安的缩写，即 volt Ampere Reactive，简称乏。

## 2. 基波、谐波

当供电线路中的正弦波电压施加在非线性电路上时，电流就变成非正弦波，非正弦电流在电网阻抗上产生压降，会使电压波形也变为非正弦波。非正弦波可用傅立叶级数分解，其中频率与工频相同的分量称为基波，频率大于基波的分量称为谐波。在电力行业中，谐波是指整数倍于工频频率的交流电。我国电网规定工频频率是 50 Hz，所以基波频率是 50 Hz，这样 5 次谐波电压（电流）的频率就是 250 Hz。

不同的用电器由于电路设计不同，产生的谐波也是不同的。因此谐波数据中包含有不同用电器独特的特征。