基于模糊C均值聚类算法的电力行业负荷特性分析

采用模糊C均值（FCM）聚类方法，对电力行业负荷特性进行分析。首先研究了FCM聚类算法，然后根据用户日负荷曲线走势，使用FCM算法分析同行业内不同用户的日负荷曲线特征。通过对实际用户负荷数据的计算分析，表明该方法可以有效区分不同用户的负荷特征，将负荷曲线走势相近的用户归为一类，实现电力行业内大量用户的类型识别。

周国鹏　孙志杰　傅　军　方茂益／华北电力科学研究院有限责任公司
于　爽／国网冀北电力有限公司经济技术研究院

周国鹏／工程师

关键词／Keywords
模糊C均值聚类·
电力负荷·
特性分析·

0　引言

随着电力事业的不断发展，各行业电力用户越来越多，电力企业的用户服务和管理工作也越来越复杂。研究电力用户的负荷特性，对具有大量性、多样性特点的电力负荷进行分析，具有现实意义和实用价值[1-6]。当前，传统的用户负荷分类方式主要是按照行业、用户电压等级和电价目录类别等方式进行分类，这些分类方法虽然在一定程度上反映了用户用电特性，但是不能满足深层次分析用户负荷特性的需求。基于聚类的用户负荷分析方法，可以根据大量电力用户的用电负荷曲线特性进行分类。通过将用电负荷曲线变化趋势相近的用户归为一类，深入分析其用电量变化规律，为负荷管理、电网规划、电力调度和电力营销等工作提供理论依据和技术支持[7-8]。

目前用于电力用户负荷特性分析的方法有多种，比较常见的有K-means聚类、层次聚类、高斯混合模型聚类和自组织特征映射神经网络等[9-16]。模糊C均值（FCM）聚类建立了样本对类别的不确定性描述，可以量化地表征样本与类别的隶属度关系，能够更加客观地反映聚类结果。因此，FCM聚类算法可用于电力负荷特性分析。但是，当前多数关于FCM用于负荷分析的研究主要集中在电力负荷曲线特性识别方面，即不考虑用户行业类别，对大量混杂的用户数据进行聚类分析，将负荷曲线形状相似的归为一类。这种处理方法虽然一定程度上能够反映用户的负荷特征，但脱离了电力用户的行业特征，实际应用意义不大。因此，在考虑用户行业分类的基础上，对某一具体行业的电力用户进行聚类分析，获得该行业内用户的用电特征，探究FCM聚类在电力行业负荷特性识别中的作用。

1　FCM聚类算法

1.1　FCM聚类算法原理

FCM聚类算法是一种基于划分的聚类算法，其思想就是使得被划分到同一簇的对象之间相似度最大，而不同簇之间的相似度最小。1973年，Bezdek提出了该算法，作为早期硬C均值（HCM）聚类方法的一种改进。HCM算法对于数据的划分是硬性的，而FCM则是一种柔性的模糊划分[17-20]。

FCM把n个向量$X_i(i=1,2,\cdots,n)$分为c个模糊组（类），并求取每组的聚类中心，使得非相似性指标的价值函数达到最小。通过模糊划分，FCM可以将

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

每个给定数据点用值在[0,1]之间的隶属度U来确定其属于各个类的程度。矩阵$\boldsymbol{U}$的元素U_{ij}表示第$j(j=1,2,\cdots,n)$个数据点属于第$i(i=1,2,\cdots,c)$类的隶属度，U_{ij}须满足如下条件[21-22]，即

$$\begin{aligned}&\forall j,\sum_{i=1}^{c}u_{ij}=1\\&\forall ij,u_{ij}\in[0,1]\\&\forall i,\sum_{j=1}^{n}u_{ij}>0\end{aligned}\tag{1}$$

FCM聚类算法就是求使聚类的目标函数最小化的隶属度矩阵$\boldsymbol{U}$和聚类中心C，目标函数是所有各点隶属度与该点到中心的欧氏距离的乘积之和，即

$$\min\quad J_m(U,C)=\sum_{j=1}^{n}\sum_{i=1}^{c}u_{ij}^{m}d_{ij}^{2}(x_j,c_i)\tag{2}$$

$$d_{ij}(x_j,c_i)=\left\|c_i-x_j\right\|=\sqrt{(x_j-c_i)^{\mathrm{T}}(x_j-c_i)}\tag{3}$$

式中，n为样本数据集的个数；c为聚类中心数；m为加权指数；d_{ij}为样本点和聚类中心距离。

隶属度U和聚类中心C可以由如下公式计算，即

$$u_{ik}=\frac{1}{\sum_{j=1}^{c}(\frac{d_{ik}}{d_{jk}})^{2}}\tag{4}$$

$$c_i=\frac{\sum_{k=1}^{n}(u_{ik})^{2}x_k}{\sum_{k=1}^{n}(u_{jk})^{2}}\tag{5}$$

FCM算法的具体步骤描述如下。

1）确定聚类数目c，初始化模糊指标及聚类中心C。

2）建立模糊相似矩阵，初始化隶属矩阵$\boldsymbol{U}$。

3）算法开始迭代，直到目标函数J收敛到极小值。

4）根据迭代结果，由最后的隶属矩阵确定数据所属的类，显示最后的聚类结果。

1.2 FCM聚类效果评价

一般而言，有效的FCM聚类结果应该具有如下特点。

1）类内紧凑度大，同类曲线间距离较小。

2）类与类之间分离度大，非同类曲线间距离较大。

3）在聚类生成的隶属度矩阵中，每条曲线的最大隶属度的平均值较高。

常用的FCM有效性评价指标公式如下，即

$$V(U,C,c)=\frac{\sum_{j=1}^{n}\sum_{i=1}^{c}u_{ij}^{m}d_{ij}^{2}(x_j,c_i)}{\sum_{j=1}^{n}\sum_{i=1}^{c}u_{ij}^{m}d_{ij}^{2}(c_j,\overline{c})}\tag{6}$$

$$\overline{c}=\sum_{i=1}^{c}c_i/c$$

上述有效性函数的分子部分是FCM的目标函数，可以衡量聚类后类内的紧凑度，分子越小，表示类内越紧凑；分母部分可以衡量类与类之间的分离度，分母越大，表示类与类之间分离度越大。显然，好的聚类结果应当是类内紧凑度大、类与类之间分离度大，对应于有效性函数的最小值。

2 FCM用于电力负荷特性分析

2.1 数据预处理

各行业用户的电力负荷数据可以从电网企业的用电信息采集系统中获得，但是不同用户的用电量存在较大差异。为去除原始数据数量级对聚类分析的影响，需要对数据进行预处理。

首先筛选、剔除错误数据，然后通过极值将序列集归一化，以提取数据形态。归一化处理公式为

$$x'_{ij}=\frac{x_{ij}}{\max\{x_{ij}\}}\quad i=1,2,\cdots,m;\ j=1,2,\cdots,n\tag{7}$$

归一化后，将数据限定在[0,1]区间内，然后处理数据，进行负荷特性分析。

2.2 算法流程

将数据归一化之后，可以开始FCM聚类分析，对某一个电力行业内的大量用户进行分类，分析其负荷特性。FCM聚类算法用于电力行业负荷特性分析的计算过程如图1所示。

3 算例分析

3.1 算例设计

电力用户涉及的行业种类较多，包括农业、

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

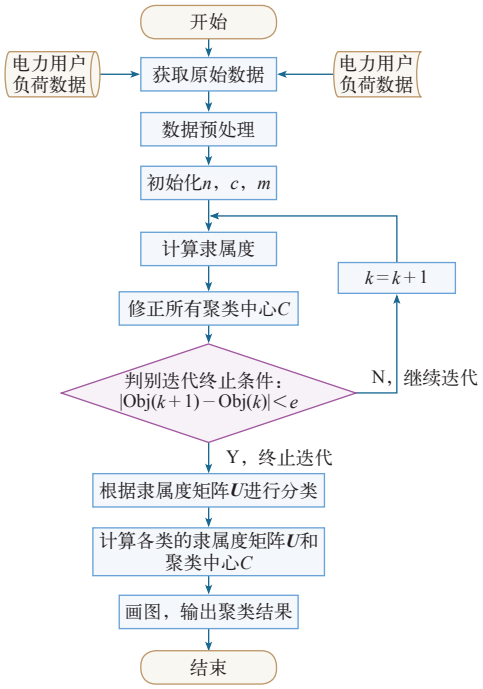

图1 FCM用于电力负荷特性分析的流程

工业、建筑业、商业和公共事业等，可以应用FCM聚类算法针对电力用户的每个行业进行聚类分析，得到该行业用户用电负荷的特性，以此作为标准开展各行业负荷特性研究、负荷管理工作。在此，选取住宿业、餐饮业两个行业作为研究案例。

以我国北方某城市实际电网为研究对象，从该电网辖区内选取日负荷数据较为完整的14户住宿行业客户、11户餐饮业客户进行分析。从电网企业的用电信息采集系统可以得到各个用户典型工作日0：00～23：00的用电量数据，数据采集时间间隔为1 h，每天有24个数据点。两个行业的用户日用电曲线分别如图2、图3所示（用电量进行了归一化处理）。

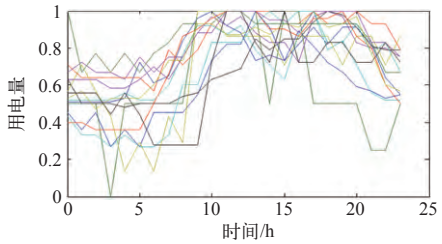

图2 住宿行业电力用户典型日用电量曲线

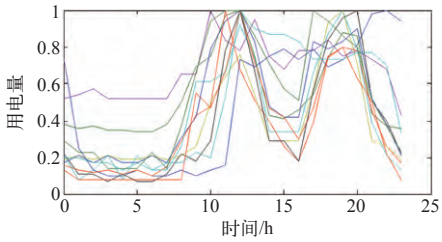

图3 餐饮行业电力用户典型日用电量曲线

从图2中可以看出，住宿行业电力用户的负荷特性具有一定的规律性，用电高峰一般为6：00～22：00，低谷时段为0：00～6：00，这与住宿行业的营业性质有关。但是，每户的用电量曲线又存在一定差异性，根据酒店级别、规模和营业性质（旅游性酒店、一般酒店）等，又可以划分为几种典型类别。这样，采用FCM聚类算法可以找出负荷曲线相近的住宿行业用户，并把它们归为一类，作为日后分析该类用户负荷特征的参考标准。

类似地，从图3中也可以看出，餐饮行业负荷曲线的高峰集中在10：00～14：00和16：00～20：00两个时段，对应于午餐和晚餐两个餐饮行业客流量较大的时间。因此，可以同样应用FCM聚类算法对大量用户负荷曲线进行分析，得到该行业用户的电力负荷模式。

3.2 聚类结果及分析

（1）聚类结果

针对住宿、餐饮两个行业电力用户的典型日用电量曲线，采用FCM方法对负荷曲线进行聚类分析。根据所选两个行业用户数量及特点，将住宿行业用户分为三类，餐饮行业用户分为两类。按照图1所示流程，在Matlab中编写程序，对两个行业电力负荷曲线聚类处理，得到聚类结果如图4、图5所示。

（2）结果分析

对于住宿行业，按照用户负荷曲线走势特点聚为三类，类别一、类别二分别包含五户，类别三包含四户。三类曲线用电高峰基本集中在9：00～21：00（白天活动时间），低谷时段在22：00～次日6：00（夜晚休息时间），用电量走势与住宿行业特点相符。进一步分析发现，类别一与类别二

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

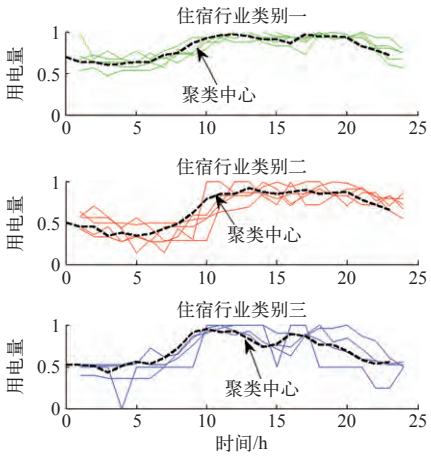

图4 住宿行业电力用户日用电量曲线聚类结果

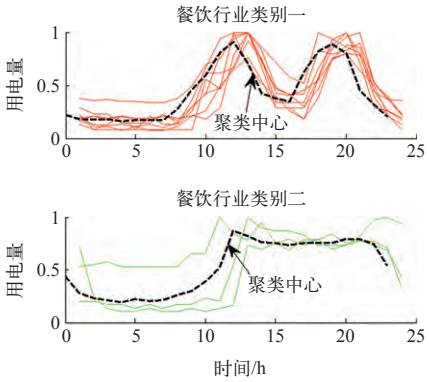

图5 餐饮行业电力用户日用电量曲线聚类结果

走势基本相同，区别在于类别一整体用电量高于类别二，对应于规模较大、业务量较多的酒店。类别三虽然整体走势与行业一致，但不规则变化较多，对应于规模一般、业务量不确定性较大的酒店。

对于餐饮行业，按照用户负荷曲线走势特点划分为两类，类别一包含八户，类别二包含三户。类别一所属用户的用电高峰，基本集中在11：00～13：00（午餐时间）和18：00～21：00（晚餐时间）两个时段，其他时间由于客源较少，用电量较低，用电量走势基本符合餐饮行业特点。因此，类别一对应于多数客流量、业务较为稳定的餐饮公司。类别二用电高峰基本集中在11：00～21：00整个时段，对应于营业时间长的冷饮、快餐等业务为主的餐饮行业电力用户。

概括而言，经过FCM聚类之后，可以对同行业内的用户用电特性进一步进行识别、聚类，将业务类型相似、用电走势趋近的电力用户分为一类，有助于加强对相关电力用户的用电服务和管理。

4 结束语

研究了FCM聚类算法，并将其应用于电力行业用户负荷特性分析中。选取某电力公司住宿、餐饮两个行业实际负荷数据，使用FCM对其进行聚类分析，分别得到两个行业的聚类结果。研究表明，对于所选行业，FCM聚类算法能够有效地区分同行业内的不同用户类型，对于电力负荷研究、电力用户服务和管理以及电力运行调度等工作的开展都具有一定的参考意义。

参考文献

[1] 宋易阳，李存斌，祁之强．基于云模型和模糊聚类的电力负荷模式提取方法 [J]．电网技术，2014，38 (12)：3378-3383．

[2] 陈刚．基于数据挖掘的电力营销决策支持系统的结构原理及算法研究 [D]．重庆：重庆大学，2004．

[3] 张忠华．电力系统负荷分类研究 [D]．天津：天津大学，2007．

[4] 韩丽娜．杭州地区电力负荷特性的分析及预测 [D]．杭州：浙江大学，2008．

[5] 陈伟，乐丽琴，崔凯，等．典型用户负荷特性及用电特点分析 [J]．能源技术经济，2011，23 (9)：44-49．

[6] 朱晓清．电力负荷的分类方法及其应用 [D]．广州：华南理工大学，2012．

[7] 张军微．聚类算法在电力大客户行为分析中的研究与应用 [D]．北京：华北电力大学，2013．

[8] 程慈．聚类分析算法在电力营销决策支持系统中的应用研究 [D]．阜新：辽宁工程技术大学，2008．

[9] 周开乐，杨善林．基于改进模糊C均值算法的电力负荷特性分类 [J]．电力系统保护与控制，2012，40 (22)：58-63．

[10] 杨浩，张磊，何潜，等．基于自适应模糊C均值算法的电力负荷分类研究 [J]．电力系统保护与控制，2010，38 (16)：111-115，122．

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

[11] 曾博，张建华，丁蓝，等．改进自适应模糊C均值算法在负荷特性分类的应用 [J]．电力系统自动化，2011，35 (12)：42-46．
[12] 蒋国栋．基于K-means聚类算法的负荷模型研究 [D]．哈尔滨：哈尔滨工业大学，2009．
[13] 刘建华，王进，孟颖，等．基于模拟退火的粗糙集K均值电力负荷聚类分析 [J]．现代电力，2012 (1)：10-14．
[14] 郭良峰．基于遗传算法的电力系统过电压分层模糊聚类识别 [D]．重庆：重庆大学，2009．
[15] 李培强，李欣然，陈辉华，等．基于模糊聚类的电力负荷特性的分类与综合 [J]．中国电机工程学报，2005，25 (24)：73-78．
[16] 张红斌，贺仁睦，刘应梅．基于KOHONEN神经网络的电力系统负荷动特性聚类与综合 [J]．中国电机工程学报，2003，23 (5)：1-5．
[17] 段钿，张彩庆，刘爱芳．模糊聚类在电力用户分类中的应用 [J]．电力需求侧管理，2005，7 (5)：18-20．
[18] 冯晓蒲，张铁峰．基于实际负荷曲线的电力用户分类技术研究 [J]．电力科学与工程，2010 (9)：18-22．
[19] 刘丽轻．电力用户负荷模式识别系统研究与设计 [D]．北京：华北电力大学，2012．
[20] 李小芳．变电站负荷特性分类与综合负荷建模研究 [D]．长沙：湖南大学，2013．
[21] 刘永光，孙超亮，牛贞贞，等．改进型模糊C均值聚类算法的电力负荷特性分类技术研究 [J]．电测与仪表，2014，51 (18)：5-9．
[22] 许翠娟．负荷特性聚类分析方法的研究 [D]．北京：华北电力大学，2007．

（收稿日期：2015－07－30）EA

(上接第84页)

参考文献

[1] 宋杲，李伟，宋竹生，等．国网公司系统组合电器运行情况分析 [J]．高压电器，2009，45 (6)：78-82．
[2] 梁之林，朱人铭，张英杰，等．两起GIS组合电器事故原因分析及建议 [J]．电力设备，2007，8 (7)：63-65．
[3] 孙茁，薛源，郭宏伟．126 kV GIS设备典型故障分析与处理 [J]．高压电器，2010，46 (11)：95-98．
[4] 黎斌．SF_6高压电器设计 [M]．北京：机械工业出版社，2010．
[5] 张勇平，刘鸿斌，邓春，等．沽太500 kV双回线路覆冰舞动故障分析 [J]．中国电力，2010，43 (3)：27-30．
[6] 杜小平，李涛，李瑞林．一起220 kV GIS闪络故障分析及建议 [J]．电力系统保护与控制，2010，38 (3)：128-130．
[7] 李继红，戴彦，王超．大电网连锁故障的风险分析及对策 [J]．电网技术，2011，35 (12)：43-49．
[8] DL/T 555—2004 气体绝缘金属封闭开关设备现场耐压及绝缘试验导则 [S]．
[9] International Electro technical Commission．IEC 62278 Railway applications－the specification and demonstration of reliability，availability，maintainability and safety (RAMS) [S]．2002．
[10] 孔凡，张耀坤，宋彬．基于WSR系统方法论的DHDF综合评价方法 [J]．兵工自动化，2007，26 (1)：29-31．

（收稿日期：2015－04－13）EA

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net