文章编号:1007-6735(2015)01-0013-05 DOI:10.13255/j.cnki.jusst.2015.01.003

基于直觉模糊C-均值的客户聚类和识别方法

耿秀丽， 尤星星， 吕文元
（上海理工大学 管理学院,上海 200093）

摘要：客户聚类和识别是大规模客户化生产中产品/服务快速有效设计的基础.考虑客户需求信息的不确定性,提出了基于直觉模糊C-均值的客户聚类算法.针对传统基于欧式距离的C-均值聚类方法无法计算直觉模糊数组间距离的缺点,采用直觉模糊交叉熵方法处理算法中的距离计算问题.同时,直觉模糊交叉熵还用来计算新客户和各客户类间的偏好相似度,进行客户识别.最后以某工程机械企业服务开发中的客户聚类和识别为例,验证了所提方法的有效性.

关键词：大规模客户化生产；客户聚类；C-均值；直觉模糊集；交叉熵
中图分类号：TH 122;N 94 **文献标志码**：A

Customer Clustering and Pattern Identification Approach Based on Vague C-means

GENG Xiuli, YOU Xingxing, LV Wenyuan
(*Business School, University of Shanghai for Science of Technology, Shanghai 200093, China*)

Abstract: In the mass customization production, customer clustering and identification are the basis of quick and effective product/service design. Considering the uncertainty of customer requirements, a customer clustering and pattern identification approach based on vague C-means was proposed. Aiming at the problem that the traditional fuzzy C-means based on Euclidean distance cannot deal with the distance between vague sets, a vague cross-entropy approach was adopted to deal with the distance calculating problem in the C-means clustering algorithm. At the same time, the vague cross-entropy was also applied in calculating the similarity between new customer and different customer groups, and then the customer identification was realized. Finally, a case study of customer clustering and identification in a mechanical company's service development was presented to illustrate the effectiveness of the proposed approach.

Key words: *mass customization*; *customer clustering*; *C-means*; *vague set*; *cross-entropy*

收稿日期：2014-03-09
基金项目：国家自然科学基金资助项目(71301104,71271138)；上海市教委科研创新基金资助项目(14YZ088)；上海市一流学科建设资助项目（S1201YLXK)；高等学校博士学科点专项科研基金资助项目(20133120120002,20120073110096)；沪江基金资助项目(A14006)
第一作者：耿秀丽(1984-),女,讲师.研究方向:产品服务工程、工业工程.E-mail:xiuliforever@163.com

大规模客户化生产是基于客户需求生产定制产品和服务,同时能保持大规模生产高质量和高效率的生产方式.大规模客户化生产强调满足客户多样化的偏好和需求,根据客户对产品或服务的偏好,分析产品或服务功能,最终获得个性化的设计方案.客户关系管理(customer relationship management, CRM)系统和累积的设计知识为客户化生产提供了支持.通过对市场客户需求偏好进行聚类分析,建立客户类与产品/服务方案类间的映射关系,可以快速分析客户需求,提高产品/服务设计效率.Shao 等[1]指出产品客户化设计的两个基本问题是需求配置和配置设计,其中需求配置是建立客户类和产品功能需求类间的依赖关系.Hong 等[2]采用模糊 C-均值(fuzzy C-means, FCM)聚类方法进行客户需求聚类,通过挖掘客户类与产品结构类间的关联关系,支持窗户产品的大规模客户化生产.

大规模客户化产品/服务设计中,客户聚类的依据是客户对产品/服务不同属性的偏好差异.客户对需求属性的偏好表达通常是不确定的.常用的传统模糊集方法仅用一个隶属度函数来表达决策者判断的信心度,包含的信息量少,难以全面反映评价信息的模糊性和不确定性.为了克服传统模糊集的缺点,Gau 等对模糊集理论进行了拓展,提出了直觉模糊集(vague set)的概念[3].直觉模糊集在处理决策者评价信息时,同时考虑了正隶属度、负隶属度和犹豫度3个方面的信息,提高了处理模糊语义信息的能力.客户间偏好的差异体现为两组偏好信息的差异,如采用直觉模糊数表达客户对各产品/服务属性的偏好,客户间的偏好差异体现为两组直觉模糊数集间的距离大小.

常见的聚类方法有层次聚类、划分聚类、基于网格聚类、基于密度聚类及模糊聚类等[4].文献[5]采用模糊聚类法来分析大学生网络行为.文献[6-7]分别提出了改进的层次谱聚类算法和改进的 FCM 聚类算法.FCM 是常用的客户聚类方法,通过优化目标函数得到每个样本点对所有类中心的隶属度,从而决定样本点的类属以达到自动对样本数据进行分类的目的.传统 FCM 算法是针对特征空间中的点集设计的,通常采用欧式距离计算两点之间的距离,无法处理采用直觉模糊数表达条件下的多属性客户偏好聚类.直觉模糊交叉熵可以解决这一问题.交叉熵是模糊集理论中的一个重要课题,它是度量两个系统间差异程度的重要工具[8].文献[9]根据概率分布交叉熵的概念,提出了计算两个直觉模糊数集间熵的方法.该熵称为直觉模糊交叉熵,用于计算两组直觉模糊数集间的信息相似度.相似度越大表明两个直觉模糊数集之间的距离越小.文献[10]将直觉模糊交叉熵与 TOPSIS 方法相结合用于确定方案属性评价值和正负理想解之间的距离.此外,直觉模糊交叉熵已成功应用在模式识别、疾病诊断等领域.

本文在客户化服务开发背景下,提出了基于直觉模糊 C-均值的客户聚类和识别方法.在客户聚类中,采用直觉模糊数处理和表达客户个性化的服务需求偏好信息,将直觉模糊交叉熵引入 C-均值聚类算法中,提出了直觉模糊 C-均值聚类方法.在客户识别中,通过采用直觉模糊交叉熵方法计算新客户和典型客户类间的信息相似度,进行客户类型识别.最后以某工程机械企业服务开发中的客户聚类分析为例,验证了所提方法的有效性.

1 客户需求信息表达和研究框架

当前我国正处于从“产品经济”向“服务经济”的转型过程中,很多制造企业开始加大服务设计和供应力度.但是,企业往往在产品使用过程中向客户销售服务.这些服务的设计和提供并没有依据客户个性化的需求,没有与客户使用情景及产品特征相结合,不能有效提升客户满意度,也难以产生规模效益.目前研究较热的产品服务系统理念强调企业在客户购买初期为其提供产品和服务组合的完整解决方案.为客户提供个性化的系统服务方案需要根据已有的服务设计和供应经验及数据分析客户对服务需求偏好及其满意的服务方案,从而进行客户需求配置分析.需求配置分析包括客户需求聚类、服务方案聚类和需求类与方案类间依赖关系的识别.通过客户需求类的识别,根据获取的需求类与方案类间的依赖关系,可有效地获得客户化的服务方案,如图1所示.

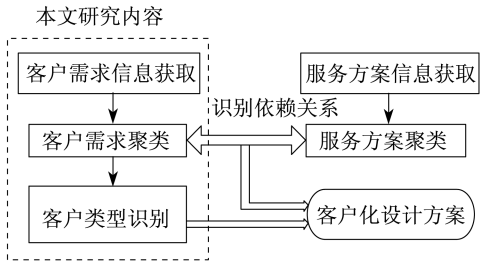

图1 客户需求配置分析流程

Fig.1 Process of customer requirements configuration analysis

假设企业获取的有效客户需求偏好信息有 n 条，共有 t 个服务属性，S_i 表示第 i 个客户（$1\leqslant i\leqslant n$），A_j 表示服务的第 j 个属性（$1\leqslant j\leqslant t$），x_{ij} 即表示第 i 个客户对第 j 个服务属性的偏好程度. 本文中客户的需求偏好共有 7 个等级，分别为非常不重要、不重要、较不重要、一般、较重要、重要、非常重要. 将客户需求偏好的语义信息转化为直觉模糊数，则第 i 个客户的需求偏好可用 $X(i)=\{\hat{x}_{i1},\hat{x}_{i2},\cdots,\hat{x}_{it}\}$ 这一组直觉模糊数来表示.

直觉模糊集的定义：设 X 是一个论域，则称 $V=\{(x,[t_V(x),1-f_V(x)])\mid x\in X\}$ 为 X 上的一个直觉模糊集. $t_V(x)$ 为 x 属于集合 V 的正隶属度函数，是支持 x 的证据所导出的 x 的肯定隶属度下界；$f_V(x)$ 为 x 属于集合 V 的负隶属度函数，是从反对 x 的证据所导出的 x 的否定隶属度下界，$0\leqslant t_V(x)+f_V(x)\leqslant 1$. 当 $t_V(x)=1-f_V(x)$ 时，直觉模糊集退化为模糊集，$1-(t_V(x)+f_V(x))$ 即为判断 x 是否属于集合 V 的犹豫度，定义为 $\pi_V(x)$，满足条件

$$t_V(x)+f_V(x)+\pi_V(x)=1,\quad 0\leqslant t_V(x),\pi_V(x),f_V(x)\leqslant 1$$

2　考虑属性离散程度的改进直觉模糊交叉熵

设 $X=\{X(1),X(2),\cdots,X(n)\}$ 是 n 组直觉模糊数集合，其中 $X(i)$ 表示第 i 组直觉模糊数集，$X(i)=\{\hat{x}_{ij}\mid 1\leqslant j\leqslant t\}$. 令 $X(a)$ 和 $X(b)$ 是 X 中两组直觉模糊数集，$\hat{x}_{aj}=[t_a(j),1-f_a(j)]$，$\hat{x}_{bj}=[t_b(j),1-f_b(j)]$，则 $X(a)$ 和 $X(b)$ 间的直觉模糊交叉熵 $D(X(a),X(b))$ 定义为[9]

$$\begin{aligned}D(X(a),X(b))=&\sum_{j=1}^{n}w_j\frac{t_a(j)+1-f_a(j)}{2}\cdot\\&\mathrm{lb}\frac{\frac{(t_a(j)+1-f_a(j))}{2}}{\frac{1}{4}[(t_a(j)+1-f_a(j))+(t_b(j)+1-f_b(j))]}+\\&\sum_{j=1}^{n}w_j\frac{1-t_a(j)+f_a(j)}{2}\cdot\\&\mathrm{lb}\frac{\frac{(1-t_a(j)+f_a(j))}{2}}{\frac{1}{4}[(1-t_a(j)+f_a(j))+(1-t_b(j)+f_b(j))]}\end{aligned}\tag{1}$$

式中，w_j 表示属性 A_j 的权重.

属性权重确定的典型方法有：a. 专家直接打分法. 该方法简单直观，但是主观性强且不能处理语义评价信息. b. AHP 方法[11]. 该方法适用面广，但需要调查大量顾客需求信息，并进行两两比较，但难以保证方法所要求的评判信息的一致性. c. 信息熵方法[12]. 该方法的原理是根据属性值的离散程度确定属性权重，但是该方法难以处理直觉模糊数. 本文考虑属性偏好离散程度确定属性权重，属性偏好分布越离散，该属性的权重越大；反之，该属性的权重越小. 本文采用改进加权最小平方法[10]，计算出各个属性偏好程度的中心值；然后分别针对各个属性，计算所有客户需求偏好与中心值间的距离和；最后得到各个属性的权重. 属性权重确定的具体步骤如下：

步骤 1　针对每个属性，采用改进加权最小平方法确定该属性所有客户需求偏好的中心值 $\hat{x}_j=[x_{jl},x_{ju}]$，$x_{jl}$ 表示区间值的下限，x_{ju} 表示区间值的上限，则

$$\begin{cases}\min D_j=\dfrac{1}{2}\displaystyle\sum_{i=1}^{n}[(x_{ijl}-x_{jl})^2+\\\qquad(x_{iju}-x_{ju})^2]\\\text{s.t. }0\leqslant x_{jl}\leqslant x_{ju}\leqslant 1\end{cases}\tag{2}$$

步骤 2　针对每个属性，分别计算所有客户需求偏好和中心值间的距离和，两直觉模糊数间距离计算定义如下：

假设 $x=[t_x,1-f_x]$，$y=[t_y,1-f_y]$，则 x 和 y 之间的距离 $d(x,y)$ 为[13]

$$\mathrm{d}(x,y)=\sqrt{(t_x-t_y)^2+(1-f_x-(1-f_y))^2}\tag{3}$$

步骤 3　将得到的各个属性偏好的组内距离和归一化，即得到各个属性的权重.

3　基于直觉模糊 C－均值方法对客户进行聚类和识别

3.1　基于直觉模糊交叉熵的模糊 C－均值客户聚类算法步骤

模糊 C－均值算法是一种以局部代价函数最小化为目标的聚类算法，将数据集划分为 $c(c>1)$ 类. c 类确定后，选取第一个点作为第一个聚类中心；接着选取离第一个点距离最远的点，作为第二个聚类中心；至于第三个聚类中心，选取离第一、二两个点距离之和最远的点；以此类推，直到选出 c 个聚类中心. 假设从 CRM 中提取了 n 个客户的需求偏好信息，样本集为 $X=(X(1),X(2),\cdots,X(n))$，$X(i)(1\leqslant i\leqslant n)$ 表示第 i 个客户对产品不同属性偏好的一组直觉模糊数. 因此，本文中选取聚类中心

时，聚类中心不是一个点，而是一组直觉模糊数，设 $R=\{R(1),R(2),\cdots,R(k),\cdots,R(c)\}$ 为 c 个聚类中心.通过迭代方法不断修正聚类中心，迭代过程以极小化所有的数据点到各个聚类中心的距离与隶属度值的加权和为优化目标，其目标函数为

$$J_m=\sum_{i=1}^{n}\sum_{k=1}^{c}u_{ik}^{m}d_{ik}^{2} \tag{4}$$

式中，d_{ik} 表示两组直觉模糊数 $X(i)$ 与 $R(k)$ 之间的距离；直觉模糊交叉熵用来计算直觉模糊数序列之间的关联度，则 $d_{ik}=D(X(i),R(k))$；m 为模糊系数，用于调节聚类模糊程度，通常取 $m=2$；$\boldsymbol{U}=\{u_{ik}\in[0,1]\}_{n\times c}$ 为模糊划分矩阵，又称隶属度矩阵；u_{ik} 为第 i 个样本数据 $X(i)$ 对第 k 组聚类中心 $R(k)$ 的隶属度，且满足 $\sum_{k=1}^{c}u_{ik}=1$，$u_{ik}\in[0,1]$，$0<\sum_{i=1}^{n}u_{ik}<n$，$1\leqslant i\leqslant n$，$1\leqslant k\leqslant c$.

直觉模糊 C－均值聚类算法的具体步骤如下：

步骤 1 给定聚类类别数 c，初始化聚类中心，$R^0(1)=X(1)=\{\hat{x}_{11},\hat{x}_{12},\cdots,\hat{x}_{1t}\}$，设置迭代条件 $\varepsilon>0$，设置迭代次数 $l=1$.

步骤 2 利用 R^{l-1} 更新 $\boldsymbol{U}^l$，则 $u_{ik}^{l}=1/\sum_{r=1}^{c}\left[\left(\frac{d_{ik}^{l}}{d_{ir}^{l}}\right)^{\frac{2}{m-1}}\right]$，$1\leqslant i\leqslant n$，$1\leqslant k\leqslant c$，$1\leqslant r\leqslant c$；$d_{ik}^{l}$ 和 d_{ir}^{l} 分别表示第 l 次迭代时第 i 组聚类中心与第 k 组聚类中心间的距离以及第 i 组聚类中心与第 r 组聚类中心间的距离.若 $X(i)=R^l(k)$，则 $u_{ir}(l)=1$.

步骤 3 利用 $\boldsymbol{U}^l$ 更新 R^l，$R^l(k)=\dfrac{\sum_{i=1}^{n}u_{ik}^{m}(l)X(i)}{\sum_{i=1}^{n}u_{ik}^{m}(l)}$，$1\leqslant k\leqslant c$.

步骤 4 如果 $\max_k D(R^l(k),R^{l-1}(k))<\varepsilon$，算法停止；否则令 $l=l+1$，回到步骤 2.

该算法最终得到 c 种客户类型和各自的聚类中心：$R(1)$，$R(2)$，$\cdots R(c)$，其中 $R(k)=\{\hat{\rho}_{k1},\hat{\rho}_{k2},\cdots,\hat{\rho}_{kj},\cdots,\hat{\rho}_{kt}\}$，$1\leqslant k\leqslant c$. $\hat{\rho}_{kj}$ 为第 k 个聚类中心第 j 个客户需求偏好的取值，$1\leqslant j\leqslant t$.

3.2 基于直觉模糊交叉熵的客户类识别

客户类识别是根据该客户的需求，将其与已有的客户类进行匹配，从而将其归到已有的客户类当中，再根据已经建立的客户类与产品/服务功能结构类间的映射关系，最终设计出客户所需的产品/服务.

新客户需求可以表达为一组反映客户需求偏好信息的直觉模糊数集.此时，利用直觉模糊交叉熵方法依次计算该组直觉模糊数集与典型客户类聚类中心的直觉模糊集之间的距离.距离越小，即相似度越大，从而将其归到某一类当中，实现客户类型的识别.

4 案例分析

A公司是国内一家著名的工程机械制造企业.近几年该公司一直致力于满足客户要求，向客户提供各种服务.然而这些服务的设计与提供并没有根据客户个性化的需求，没有与客户使用情景及产品特征相结合，不能有效提升客户满意度并难以产生规模效益.公司拟采用本文所提方法对该公司的典型客户进行聚类，从而提高该公司的售后服务设计的效率和有效性，提升客户满意度.

从A公司的客户关系管理系统中，收集到50名客户对该型号挖掘机5种售后服务项目的需求偏好信息.该5种售后服务项目为：节能环保、响应性、金融需求、再制造、低成本，分别用 A_1，A_2，A_3，A_4，A_5 表示.客户对每个售后服务项目的指标评价值用非常不重要、不重要、较不重要、一般、较重要、重要、非常重要表示.限于篇幅，文中列举了部分客户需求信息，如表1所示.

表 1 客户需求偏好信息

Tab.1 Customer requirements information

ID	A_1	A_2	A_3	A_4	A_5
1	非常重要	一般	非常不重要	重要	较不重要
2	较重要	重要	非常不重要	一般	重要
3	一般	非常重要	非常不重要	一般	重要
4	较重要	一般	非常不重要	非常重要	较重要
5	不重要	较重要	不重要	较重要	非常重要
6	不重要	重要	不重要	重要	非常重要
7	非常重要	较不重要	非常不重要	较重要	一般
8	非常不重要	较重要	重要	较重要	较重要
⋮	⋮	⋮	⋮	⋮	⋮
50	非常不重要	一般	非常重要	较重要	较重要

采用本文所提基于直觉模糊 C－均值方法对客户需求信息进行分析，实现客户聚类，步骤如下：

步骤 1 利用表2将语言术语形式的偏好评价信息转换为直觉模糊数，结果如表3所示.

步骤2 利用改进加权最小平方法确定5个属

表 2　语言变量和相应的直觉模糊数

Tab.2　Linguistic variables and the corresponding vague set numbers

语言变量	直觉模糊数
非常不重要	[0.0,0.1]
不重要	[0.1,0.3]
较不重要	[0.3,0.4]
一般	[0.4,0.5]
较重要	[0.5,0.6]
重要	[0.6,0.9]
非常重要	[0.9,1.0]

表 3　采用直觉模糊数表达的客户需求偏好信息

Tab.3　Customer requirements information expressed in vague set numbers

ID	A_1	A_2	A_3	A_4	A_5
1	[0.9,1.0]	[0.4,0.5]	[0.0,0.1]	[0.6,0.9]	[0.3,0.4]
2	[0.5,0.6]	[0.6,0.9]	[0.0,0.1]	[0.4,0.5]	[0.6,0.9]
3	[0.4,0.5]	[0.9,1.0]	[0.0,0.1]	[0.4,0.5]	[0.6,0.9]
4	[0.5,0.6]	[0.4,0.5]	[0.0,0.1]	[0.9,1.0]	[0.5,0.6]
5	[0.1,0.3]	[0.5,0.6]	[0.1,0.3]	[0.5,0.6]	[0.9,1.0]
6	[0.1,0.3]	[0.6,0.9]	[0.1,0.3]	[0.6,0.9]	[0.9,1.0]
7	[0.9,1.0]	[0.3,0.4]	[0.0,0.1]	[0.5,0.6]	[0.4,0.5]
8	[0.0,0.1]	[0.5,0.6]	[0.6,0.9]	[0.5,0.6]	[0.5,0.6]
⋮	⋮	⋮	⋮	⋮	⋮
50	[0.0,0.1]	[0.4,0.5]	[0.9,1.0]	[0.5,0.6]	[0.5,0.6]

性的权重.首先利用式(2)确定 5 个属性偏好程度的中心值,结果为:$x_1=[0.4,0.53]$,$x_2=[0.56,0.71]$,$x_3=[0.19,0.36]$,$x_4=[0.61,0.80]$,$x_5=[0.62,0.77]$.

步骤 3　针对每个属性,利用式(3)计算所有客户需求偏好和中心值的距离和,然后将各个属性偏好的组内距离和归一化,得到 5 个属性的权重分别为:$w_1=0.365$,$w_2=0.051$,$w_3=0.380$,$w_4=0.122$,$w_5=0.082$.

步骤 4　利用直觉模糊 C－均值聚类算法,依据表 3 所列客户需求偏好信息,对该 50 名客户进行聚类.基于 Matlab 软件进行聚类运算,最终得到 5 种不同的客户类型,分别为环保型、效率型、成长型、可持续发展型、经济型客户.该 5 种客户类型的聚类中心分别为

$R(1)=\{[0.89,0.97],[0.44,0.58],[0.00,0.11],[0.61,0.87],[0.25,0.36]\}$;

$R(2)=\{[0.34,0.40],[0.91,1.00],[0.00,0.10],[0.39,0.48],[0.42,0.53]\}$;

$R(3)=\{[0.01,0.12],[0.46,0.59],[0.99,1.00],[0.40,0.51],[0.64,0.91]\}$;

$R(4)=\{[0.51,0.63],[0.38,0.47],[0.00,0.12],[0.92,0.99],[0.37,0.49]\}$;

$R(5)=\{[0.01,0.11],[0.48,0.59],[0.29,0.38],[0.51,0.62],[0.86,0.97]\}$

该公司一名新顾客对该型号挖掘机 5 种售后服务项目的需求偏好信息为:$I=\{[0.0,0.1],[0.5,0.6],[0.9,1.0],[0.5,0.6],[0.6,0.9]\}$.根据已获得的 5 种客户类型和聚类中心,采用直觉模糊交叉熵方法对该客户进行需求类型识别.

利用式(1)计算该客户与上述获取的 5 种典型客户类间的信息相似度,两者之间的距离越小,相似度越高.计算得到的该客户需求偏好信息 I 与 5 类典型客户需求偏好信息 $R(1)$,$R(2)$,$R(3)$,$R(4)$,$R(5)$间的距离分别为:$D(I,R(1))=1.058$,$D(I,R(2))=0.662$,$D(I,R(3))=0.014$,$D(I,R(4))=0.772$,$D(I,R(5))=0.263$.显然,该客户与第三种典型客户类聚类中心的距离最小,因此该客户应划分到第三种客户类,即成长型客户类中.下一步即可针对成长型客户类的特点和相关的服务方案属性,快速设计适合该客户的服务方案.

5　结束语

客户聚类和识别是大规模客户化生产的重要前提.本文提出了一种基于直觉模糊交叉熵的直觉模糊 C－均值客户聚类和识别方法.该方法的特点有:a.在客户需求偏好信息获取和表达方面,采用可处理正负隶属度信息的直觉模糊集方法,相比于传统模糊集方法提高了处理需求信息模糊性和不确定性的能力;b.针对直觉模糊数的特点,提出了采用直觉模糊交叉熵方法计算不同客户需求偏好信息间的距离,用于构建 C－均值聚类算法函数;此外,还将直觉模糊交叉熵用于计算新客户需求偏好和已知客户类需求偏好间的距离来进行客户类识别.所提方法已用于某公司对挖掘机售后服务项目的客户聚类和识别分析,通过实证分析,验证了所提方法的有效性和可行性.下一步工作将在现有研究工作的基础之上,研究客户需求类与产品/服务方案类之间的映射关系获取,实现产品/服务客户化方案的设计.

(下转第 35 页)

参考文献：

[1] 曾文良，梁启明. 蓄热热泵空调器：中国，01240625.2 [P]. 2002－03－20.

[2] 曾文良，曲景年. 几种蓄热热泵空调器的性能比较 [J]. 暖通空调，2004，34(11)：45－50.

[3] 曾文良，高日新. 不间断制热的运行方法及空调系统：中国，01127805.6 [P]. 2003－03－19.

[4] 齐琦，姜益强，马最良. 相变蓄热材料及其在暖通空调领域中的应用[J]. 建筑热能通风空调，2006，25(3)：18－19.

[5] 颜慧磊，张华，邵秋萍. 一种太阳能与空气源双热源热泵系统的性能研究[J]. 上海理工大学学报，2014，36(2)：177－180.

[6] Cabeza L F, Mehling H, Ziegler F. Heat transfer enhancement in water when used as PCM in thermal energy storage [J]. Applied Thermal Engineering, 2002，22(10)：1141－1151.

[7] O'Brieen J M，Bansal P K，Raine R R. An integrated domestic refrigerator and hot water system [J]. International Journal of Energy Research, 1998, 22(8)：761－776.

[8] Farid M M，Khudhair A M，Razack S A K，et al. A reviewon phase change energy storage：materials and applications[J]. Energy Conversion and Management，2004，45(9/10)：1597－1615.

[9] 林小茁，倪龙，江辉民，等. 公共建筑用空气源热泵热水机组初探[J]. 建筑科学，2009，25(2)：55－59.

[10] 韩志涛，姚杨，马最良，等. 空气源热泵蓄能热气除霜新系统与实验研究[J]. 哈尔滨工业大学学报，2007，39(6)：901－903.

[11] 周峰，马国远. 空气能热泵热水器的现状及展望[J]. 节能，2006(7)：13－17.

（编辑：石　瑛）

（上接第 17 页）

参考文献：

[1] Shao X Y，Wang Z H，Li P G，et al. Integrating data mining and rough set for customer group-based discovery of product configuration rules [J]. International Journal of Production Research，2006，44(14)：2789－2811.

[2] Hong G, Xue D, Tu Y. Rapid identification of the optimal product configuration and its parameters based on customer-centric product modeling for one-of-a-kind production[J]. Computers in Industry，2010，61：270－279.

[3] Gau W，Buehrer D. Vague sets[J]. IEEE Transactions on Systems，Man and Cybernetics，1993，23(2)：610－614.

[4] Mitra S, Pal S K, Mitra P. Data mining in soft computing framework：a survey[J]. IEEE Transactions on Neural Networks，2002，13(1)：3－14.

[5] 李云先，彭敦陆. 大学生网络行为方式的模糊分析 [J]. 上海理工大学学报，2013，35(2)：107－112.

[6] 杨晓慧，王莉莉，李登峰. 一种新的层次谱聚类算法 [J]. 上海理工大学学报，2014，36(1)：49－52.

[7] 曹易，张宁. 一种改进的模糊 C－均值聚类算法[J]. 上海理工大学学报，2012，34(4)：351－354.

[8] 李香音. 区间直觉模糊连续交叉熵及其多属性决策方法[J]. 计算机工程与应用，2013，49(15)：234－237.

[9] Zhang Q S，Jiang S Y. A note on information entropy measures for vague sets and its applications [J]. Information Sciences，2008，178(21)：4184－4191.

[10] Geng X, Chu X, Zhang Z. A new integrated design concept evaluation approach based on vague sets[J]. Expert Systems with Applications, 2010, 37: 6629－6638.

[11] Lin M，Wang C，Chen M，et al. Using AHP and TOPSIS approaches in customer-driven product design process [J]. Computers in Industry，2008，59：17－31.

[12] Chan L, Wu M. A systematic approach to quality function deployment with a full illustrative example [J]. Omega，2005，33(1)：119 － 139.

[13] Zhang D，Huang S，Li F. Approach to measuring the similarity between vague sets[J]. Journal of Huazhong University of Science and Technology (Natural Science Edition)，2004，32(5)：59－60.

（编辑：丁红艺）