

改进混合蛙跳算法优化的产品族模糊C均值聚类设计方法

崔文华^{*1,2}, 刘晓冰¹, 王伟¹, 王介生²

(1. 大连理工大学 控制科学与工程学院, 辽宁 大连 116024;

2. 辽宁科技大学 电子与信息工程学院, 辽宁 鞍山 114044)

摘要: 研究了基于改进混合蛙跳算法优化的模糊C均值聚类解决模块化产品族设计中产品平台的确定问题. 建立了该产品开发过程中的部件关联矩阵, 采用变个体长度的混合蛙跳算法同时优化模糊聚类数和聚类中心, 求得产品构成部件的最优模糊划分. 切断算子和拼接算子用来对个体进行重新组合而形成新个体, 采用 ISODATA 迭代算法进行局部寻优. 通过对纸币清分机进行的产品族设计的仿真研究, 表明所提方法为产品族模块化设计提供了定量数学分析和快速配置的理论依据.

关键词: 纸币清分机; 产品族; 产品平台; 混合蛙跳算法; 模糊C均值聚类

中图分类号: TP12 **文献标志码:** A

0 引言

在国内外竞争日益激烈的市场环境当中, 针对不同层次和规模的客户群体设计提供合适的金融机具产品, 是金融机具企业面临的一个严重挑战. 基于产品平台(product platform)的面向产品族(product family)的设计方法在产品设计中得到了广泛应用^[1], 该方法扩展了面向变形的设计方法(design for variety, DFV), 以提高企业的敏捷定制水平.

模糊C均值(FCM)聚类算法^[2]是非监督模式识别中的一种基本方法, 目前在图像处理、模式识别、模糊控制等众多领域被广泛采用. 判定聚类算法是否高效需要考虑如下问题: 能否自动确定聚类的数目(聚类有效性问题^[3]); 是否对初始参数和噪声敏感; 是否具有全局优化能力和较快的收敛速度. 诸如粒子群算法^[4]、微分进化算法^[5]等智能优化算法被应用在模糊聚类算法中的优化领域. 这些算法在某种程度上有效地克服了传统模糊聚类算法易陷入局部极小值的缺点, 而且有效避免了对初始化选值敏感性的问题. 另一方面, 众多聚类有效性函数被提出, 如 Xie-Beni 的 V_{XB} ^[6]、Fukuyama-Sugeno 的 V_{FS} ^[7]、Kwon 的 V_K ^[8] 和

Chen-Linkens 的 V_p ^[9], 但基于聚类有效性函数的 FCM 聚类算法易陷入局部极小值和对初始化选值敏感的弱点没有得到有效解决. 文献[10]则采用神经网络算法对获取最优聚类数进行了有效尝试. 在面向产品族的产品开发设计领域, 国内外学者采用聚类算法也进行了有效探索^[1,11]. 文献[1]采用模糊聚类分析方法解决饮水机的模块化产品族设计中模块和核心平台的确定问题. 文献[11]采用聚类分析方法对电子产品功能模块进行提取, 为产品族设计提供了一定思路.

混合蛙跳算法(SFLA)就是一种基于群体的进化搜索智能算法^[12]. 2009年, Amiri等^[13]采用 SFLA 对 K 均值聚类问题进行了求解; 杜长海等^[14]提出 SFLA-FCM 优化交通时段划分问题, 但没有考虑聚类有效性问题.

本文应用基于改进混合蛙跳算法优化的模糊C均值聚类解决模块化产品族设计中产品平台的确定问题, 以期对产品族模块化设计提供定量数学分析和快速配置的理论依据.

1 FCM 聚类算法

FCM 聚类算法将特征空间 $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots,$

$\mathbf{x}_n$ 中的特征点分为 c 类 ($2 \leq c \leq n$), 第 i 类的聚类中心用 $\mathbf{v}_i$ 表示, 其中任意特征点 $\mathbf{x}_j$ 属于第 i 类隶属度 u_{ij} ($0 \leq u_{ij} \leq 1$), 且 u_{ij} 满足如下条件:

$$\sum_{i=1}^c u_{ij} = 1; j = 1, 2, \dots, n \quad (1)$$

$$0 < \sum_{j=1}^n u_{ij} < n; i = 1, 2, \dots, c \quad (2)$$

FCM 聚类算法的目标函数^[15] 为

$$J_m(\mathbf{U}, \mathbf{V}) = \sum_{i=1}^c \sum_{j=1}^n (u_{ij})^m \|\mathbf{x}_j - \mathbf{v}_i\|^2 \quad (3)$$

其中 $m > 1$. FCM 聚类算法的目的就是获取最小化目标函数(3)的 $\mathbf{U} = (u_{ij})_{c \times n}$ 和 $\mathbf{V} = (\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_c)$.

ISODATA 算法流程如下(算法流程 1):

Step 1 设定 $c_{\max}$ ($c_{\max} \leq \sqrt{n}$), T, m 和 $\epsilon > 0$;

Step 2 对于 $c = 2, 3, \dots, c_{\max}$, 初始化 $\mathbf{V}_0 = (\mathbf{v}_{10} \ \mathbf{v}_{20} \ \dots \ \mathbf{v}_{c0})$;

Step 3 针对迭代次数 $t = 1, 2, \dots, T$, 根据以下公式计算隶属度和聚类中心:

$$u_{ij,t} = \left[\frac{1}{\|\mathbf{x}_i - \mathbf{v}_j\|^2} \right]^{1/(a-1)} / \sum_{j=1}^n \left[\frac{1}{\|\mathbf{x}_i - \mathbf{v}_j\|^2} \right]^{1/(a-1)} \quad (4)$$

$$\mathbf{v}_{j,t} = \frac{\sum_{i=1}^c (u_{ij,t})^m \mathbf{x}_i}{\sum_{i=1}^c (u_{ij,t})^m} \quad (5)$$

式中: $i = 1, 2, \dots, c$; $j = 1, 2, \dots, n$. 如果 $\|\mathbf{V}_t - \mathbf{V}_{t-1}\| < \epsilon$, 则停止迭代, 否则重复 Step 3.

2 基于变个体长度的 SFLA 优化 FCM 参数

2.1 SFLA 算法

SFLA 算法的基本思想是^[12]: 随机生成 S 维解空间中由 N 只蛙组成的初始种群 $\mathbf{P} = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N\}$, 第 i 只蛙表示为 $\mathbf{X}_i = (x_{i1} \ x_{i2} \ \dots \ x_{iS})$; 然后将种群内个体按适应度值进行降序排列, 种群中最优个体记为 $\mathbf{X}_g$; 接下来将蛙群个体均分为 m 个模因组, 每组有 n 只蛙, 即 $N = m \times n$. 设蛙的集合 $\mathbf{M}^k$ 为第 k 个模因组, 则蛙的分配过程可用下式进行描述:

$$\mathbf{M}_k = \{\mathbf{X}_{k+m(l-1)} \in \mathbf{P} \mid 1 \leq l \leq n\}; 1 \leq k \leq m \quad (6)$$

记录每个模因组中最优个体和最差个体适应度为 $\mathbf{X}_b$ 和 $\mathbf{X}_w$, 而种群中最优个体记为 $\mathbf{X}_g$. 然后对

$\mathbf{X}_w$ 进行局部搜索, 采用如下蛙跳规则进行更新:

$$\mathbf{D} = r \cdot (\mathbf{X}_b - \mathbf{X}_w) \quad (7)$$

$$\mathbf{X}'_w = \mathbf{X}_w + \mathbf{D}; \|\mathbf{D}\| \leq D_{\max} \quad (8)$$

式中: $r \in [0, 1]$, $D_{\max}$ 为蛙所跳最大距离. 如 $\mathbf{X}'_w$ 优于 $\mathbf{X}_w$, 则进行个体替换, 如无改进, 则用 $\mathbf{X}_g$ 取代 $\mathbf{X}_b$, 根据式(7)和(8)重新进行局部搜索; 如个体依然没有得到优化, 则随机产生新个体替代 $\mathbf{X}_w$. 重复上述优化过程 $L_{\max}$ 次, 进行蛙群的混洗、排序、模因组划分和局部搜索, 如此反复直到定义的收敛条件结束为止, 例如达到混合迭代次数 $G_{\max}$.

2.2 编码和适应度函数

本文对聚类中心向量 $\mathbf{V}$ 进行编码, 设 n 个样本要分成 c 类, 为了合理有效地缩小 SFLA 的搜索空间, 本文中 c 的取值范围为 $[2, c_{\max}]$ ($c_{\max} \leq \sqrt{n}$). 根据各自的取值范围, 采用浮点数编码, 则一个个体 $\mathbf{I}$ 可表示为^[16]

$$\mathbf{I} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_c\} = \{\underbrace{\beta_1, \beta_2, \dots, \beta_k}_{\mathbf{v}_1}, \dots, \underbrace{\beta_i, \dots, \beta_l, \dots, \beta_{k_c}}_{\mathbf{v}_c}\} \quad (9)$$

在本文中, 采用 SFLA 对聚类中心向量 $\mathbf{V}$ 进行优化, 式(9)中的聚类数 c 是变化的, 所以个体 $\mathbf{I}$ 的长度不是固定的. 本文借助 $J_m(\mathbf{U}, \mathbf{V})$ 定义 SFLA 的适应度函数:

$$ff(\mathbf{U}, \mathbf{V}) = \frac{1}{J_m(\mathbf{U}, \mathbf{V}) + \xi} \quad (10)$$

适应度函数采用如下步骤进行求解(算法流程 2):

Step 1 从个体 $\mathbf{I}$ 中提取 $\mathbf{v}_i$ ($i = 1, 2, \dots, c$);

Step 2 依据式(4)求取 u_{ij} ;

Step 3 根据式(5)重新计算聚类中心参数;

Step 4 根据式(3)和(10)计算 $ff(\mathbf{U}, \mathbf{V})$.

确定了编码方式和 $ff(\mathbf{U}, \mathbf{V})$ 后, 就可用混合蛙跳算法对模糊聚类数和聚类中心同时优化.

2.3 切断算子和拼接算子

由于编码方式的特殊性, 由式(7)、(8)所定义的蛙跳规则不适用于变个体长度的 SFLA, 所以切断算子 (cut operator) 和拼接算子 (splice operator)^[16] 被用来将群体中各个个体进行重组 (如图 1 所示).

2.4 算法流程

基于变个体长度的 SFLA 同时优化模糊聚类数和聚类中心的算法步骤如下(算法流程 3).

Step 1 选定如下算法初始参数: $N, c_{\max}, S, m, n$ ($N = m \times n$), $L_{\max}$ 和 $G_{\max}$.

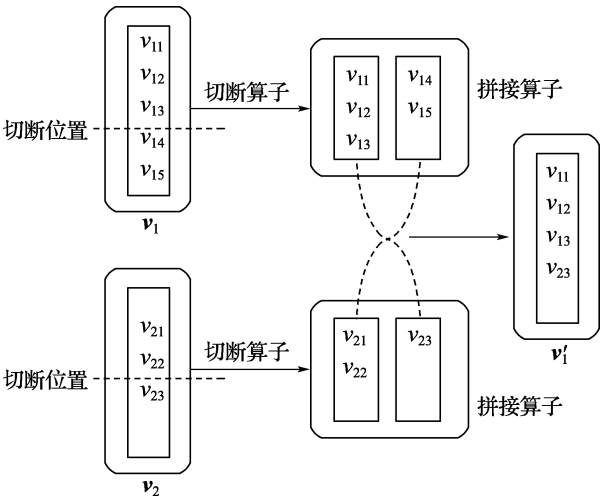


图 1 基于切断算子和拼接算子的新蛙跳规则
Fig. 1 New frog-leaping rule based on cut operator and splice operator

Step 2 随机生成包含 N 只蛙的 $P = \{X_1(t), \dots, X_k(t), \dots, X_N(t)\}$; 迭代计数值 $t = 0$; 将每只蛙 $X_k(t)$ 作为聚类中心参数, 并根据算法流程 2 求出每个个体的适应度值 $F_k(t) = F(X_k(t))$; 对蛙群按照适应度值从大到小进行排列, 令 $U_k(t) = \{X_k(t), F_k(t)\}$ 和 $X_g(t) = U_1(t)$.

Step 3 根据式(1)对 U 进行 m 个模因组 $M^1(t), \dots, M^j(t), \dots, M^m(t)$ 的形成, 令最优个体和最差个体分别为 $X_b^i(t)$ 和 $X_w^j(t)$.

Step 4 对 $M^j(t)$ 中的 $X_w^j(t)$ 与 $X_b^i(t)$ 做切断和拼接运算, 产生新蛙, 并采用算法流程 1 进行局部搜索; 然后根据算法流程 2, 计算其适应度值, 采用上文所述的蛙跳规则进行局部寻优, 得到进化后的模因组 $M^1(t)', M^2(t)', \dots, M^m(t)'$.

Step 5 将更新后的模因组 $M^1(t)', M^2(t)', \dots, M^m(t)'$ 内的蛙重新混合, 记 $U(t+1) = (M^1(t)' \ M^2(t)' \ \dots \ M^m(t)'),$ 对 $U(t+1)$ 中的蛙按照适应度值从大到小进行排列, 记录 $X_g(t+1) = U_1(t+1)$.

Step 6 $t = t + 1$, 如 $t < G_{\max}$, 转 Step 3, 否则输出最优蛙.

3 基于聚类有效性函数的模糊聚类数优化

3.1 聚类有效性函数

(1)Bezdek 提出的 V_{PC} 和 V_{PE}

Bezdek 提出的 V_{PC} 和 V_{PE} 如下:

$$V_{PC} = \frac{1}{n} \sum_{i=1}^c \sum_{j=1}^n u_{ij}^2 \quad (11)$$

$$V_{PE} = -\frac{1}{n} \sum_{j=1}^n \sum_{i=1}^c [u_{ij} \log_a u_{ij}] \quad (12)$$

V_{PC} 在 $1/c$ 和 1 之间有效, 如果期待一个良好的划分, 目标就是选取使 V_{PC} 最大的模糊划分. 它最主要的优势是分配系数简便, 缺点是它的单调下降, 缺乏直接连接到数据本身的性能.

V_{PE} 在 0 和 $\log_a c$ 之间有效. 对所有的概率群 C , 事实证明它有如下关系: $0 \leq 1 - PC(C) \leq PE(C)$. 它基本上是一个衡量模糊聚类划分的函数, 与 V_{PC} 类似.

(2)Xie-Beni 提出的 V_{XB}

V_{XB} 是基于 J_m 的客观功能, 通过确定平均数据和最小距离来确定聚类中心的.

$$V_{XB} = \frac{\sum_{i=1}^c \sum_{j=1}^n u_{ij}^2 \|x_j - v_i\|^2}{n(\min_{i \neq k} \|v_i - v_k\|^2)} \quad (13)$$

(3)Fukuyama-Sugeno 的 V_{FS}

V_{FS} 如下:

$$V_{FS} = \sum_{j=1}^n \sum_{i=1}^c u_{ij}^m \|x_j - v_i\|^2 - \sum_{j=1}^n \sum_{i=1}^c u_{ij}^m \|v_i - \bar{v}\|^2 \quad (14)$$

(4)Kwon 的 V_K

V_K 如下:

$$V_K = \frac{\sum_{j=1}^n \sum_{i=1}^c u_{ij}^2 \|x_j - v_i\|^2 + \frac{1}{c} \sum_{i=1}^c \|v_i - \bar{v}\|^2}{\min_{i \neq k} \|v_i - v_k\|^2} \quad (15)$$

(5)Chen-Linkens 的 V_P

V_P 如下:

$$V_P = \frac{1}{n} \sum_{j=1}^n \max_i(u_{ij}) - \frac{1}{\sum_{i=1}^c i} \sum_{i=1}^{c-1} \sum_{j=i+1}^c \left(\frac{1}{n} \sum_{k=1}^n \min(u_{ik}, u_{jk}) \right) \quad (16)$$

3.2 算法流程

结合 FCM 聚类算法和聚类有效性函数可得到最优模糊划分, 其算法描述如下(算法流程 4):

Step 1 指定最大聚类中心个数 $c_{\max}$ ($c_{\max} \leq \sqrt{n}$), 聚类有效性函数 $V_w(U, V, c)$ (如式(11) ~ (16)), 初始化 T, m 和 $\epsilon > 0$;

Step 2 对于 $c = 2, 3, \dots, c_{\max}$, 进行 $V_0 = (v_{10} \ v_{20} \ \dots \ v_{c0})$ 的初始化;

Step 3 针对迭代次数 $t = 1, 2, \dots, T$, 根据式(4)和(5)计算 $u_{ij,t}$ 和 $v_{j,t}$. 如果 $\|V_t - V_{t-1}\| < \epsilon$,

转下一步,否则重复 Step 3;

Step 4 计算 $V_W(U,V,c)$;如果 $c < c_{\max}$,转步骤2;否则停止迭代,最优聚类数 $c = c_b, c_b$ 满足下式:

$$V_W(U,V,c_b) = \min\{V_W(U,V,c)\};$$
$$c = 2,3,\cdots,c_{\max}$$

(17)

4 仿真实验

4.1 聚类有效性实验

聚类有效性准则要求模糊划分必须在紧密性和分离性间达到一定的平衡. 为与基于聚类有效性函数的模糊聚类优化算法进行比较,本文采用类中距和类间距两个性能指标进行对比实验:

$$J_{lz} = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^2 \|x_j - v_i\|^2$$

(18)

$$J_{lj} = \frac{1}{c} \sum_{i=1}^c \sum_{j=i+1}^c \|v_i - v_j\|^2$$

(19)

为了与基于聚类有效性函数对 FCM 算法的聚类中心个数寻优进行比较,本文采用图 2 中的数据集合开展研究.

表 1 为聚类数为 2~6 时,6 个有效性指标在 $m=2$ 下的验证结果. 从表 1 可以看出针对图 2 这个空间分布极为简单的数据集只有指标 V_{XB} 、 V_{FS} 和 V_K 找到正确的模糊数 4. 这就说明并不是所有的有效性函数对给定的数据集均为有效. 采用本文所提的基于变个体长度的 SFLA-FCM 算法,在 10 次随机运行中均得到正确聚类中心数 4,并且类中距和类间距平均值分别为 5.118 2 和 1.737 4,较好地 在模糊划分的紧密性和分离性之间取得折中.

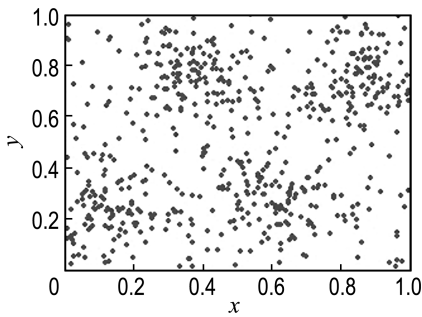


图 2 具有 4 个聚类中心的含噪声测试数据集
Fig. 2 Test data set with noises having four cluster centers

4.2 基于改进的 SFLA-FCM 的产品族设计

产品平台的选择和设计是产品族设计的核心,其中模块化的设计涉及实现产品各项功能结构的最终整合,而在 p 个功能模块之间大多具有如下关系^[11]: (1) 每个功能模块对其余功能模块

提供支撑;(2) 每个功能模块从其余功能模块接受支撑. 这样,任意两个功能模块之间的关系可采用 $p \times p$ 的模块关联矩阵进行描述.

表 1 不同聚类有效性函数数据集结果($m=2$)
Tab. 1 Results of data set under different clustering valid function ($m = 2$)

c	V_{PC}	V_{PE}	V_{XB}	V_{FS}	V_K	V_P
2	0.711 3	0.449 7	0.212 9	-60.885 9	101.844 4	0.298 3
3	0.640 8	0.593 4	0.148 6	-85.259 9	71.828 5	0.455 3
4	0.693 8	0.567 0	0.069 0	-114.713 8	34.232 1	0.569 2
5	0.616 7	0.731 2	0.342 6	-104.554 7	171.116 2	0.579 1
6	0.570 6	0.855 0	0.342 5	-97.932 7	173.389 1	0.562 8

本文所研究的纸币清分机是银行系统对回笼纸币进行自动清分处理的高端设备. 典型纸币清分机主要功能模块有电机模块(a)、嵌入式模块(b)、网络模块(c)、操作模块(d)、直线-滚轮式走钞模块(e)、进钞模块(f)、辅助模块(g)、UV 传感器(h)、磁性传感器(i)、红外光学传感器(j)、荧光传感器(k)、图像传感器(l)、测厚传感器(m)和电解质传感器(n).

本文采用文献[11]所提的部件关联矩阵标准化处理方法对部件关联矩阵进行标准化:

$$r_{ij} = w_i \left(\frac{x_{ij} - \min_j x_{ij}}{\max_j x_{ij} - \min_j x_{ij}} \right)$$

(20)

式中: $\max_j x_{ij}$ 和 $\min_j x_{ij}$ 分别表示某一部件同其他部件的最大和最小关联特征值; w_i 为功能模块 i 的关联权重变量. 这样,对纸币清分机进行标准化处理后的模糊矩阵如表 2 所示.

表 2 部件量化后关联模糊矩阵表
Tab. 2 Coefficient fuzzy matrix table after component quantity

	a	b	c	d	e	f	g	h	i	j	k	l	m	n
a	0	0	0	6	25	20	3	0	3	3	0	0	5	0
b	5	0	9	30	5	5	0	15	10	20	10	24	15	15
c	0	25	0	25	0	0	5	0	0	0	0	0	0	0
d	5	25	25	0	10	10	15	6	10	9	10	9	6	6
e	25	15	0	30	0	35	10	0	0	0	0	0	0	0
f	35	10	0	15	15	0	10	0	0	0	0	0	0	0
g	5	10	15	25	15	15	0	10	6	6	9	10	6	9
h	5	5	0	6	0	0	10	0	6	10	6	9	6	10
i	10	9	0	5	0	0	10	3	6	6	9	6	6	9
j	15	5	0	3	0	0	9	3	0	0	9	6	6	10
k	0	5	0	6	0	0	6	6	6	0	3	10	6	6
l	6	25	0	5	0	0	10	6	3	5	6	9	9	9
m	10	10	0	6	0	0	9	4	6	6	6	6	0	10
n	0	9	0	6	0	0	10	6	3	3	9	9	3	0

针对如表 3 所示的标准化后的部件关联模糊矩阵表,采用本文所提出的基于变个体长度的

SFLA 来同时优化 FCM 的聚类数和聚类中心,得到的最优聚类数为 3. 同时得到最优模糊矩阵 U . 状态变量 h 以相应的相对隶属度值作为权重,基于式(21)求得各个功能模块归属各个类别的特征值 $H(u)$ 表,这样就得出了最优的功能模块如表 4 所示分类结果.

$$H(u)=\sum_{h=1}^cU_{Ah}\cdot h$$

(21)

表 3 标准化模糊矩阵表

Tab.3 Standardized fuzzy matrix table

	a	b	c	d	e	f	g	h	i	j	k	l	m	n
a	1	0	0	0.24	1	0.80	0.12	0	0.12	0.12	0	0	0.20	0
b	0.17	1	0.30	1	0.17	0.17	0	0.50	0.33	0.67	0.33	0.80	0.50	0.50
c	0	1	1	1	0	0	0.20	0	0	0	0	0	0	0
d	0.20	1	1	1	0.40	0.40	0.60	0.24	0.40	0.36	0.40	0.36	0.24	0.24
e	0.71	0.43	0	0.86	1	1	0.29	0	0	0	0	0	0	0
f	1	0.29	0	0.43	0.43	1	0.29	0	0	0	0	0	0	0
g	0.20	0.40	0.60	1	0.60	0.60	1	0.40	0.24	0.24	0.36	0.40	0.24	0.36
h	0.50	0.50	0	0.60	0	0	1	1	0.60	1	0.60	0.90	0.60	1
i	1	0.90	0	0.50	0	0	1	0.30	1	0.60	0.90	0.60	0.60	0.90
j	1	0.33	0	0.20	0	0	0.60	0.20	0	1	0.60	0.40	0.40	0.67
k	0	0.50	0	0.60	0	0	0.60	0.60	0.60	0	1	1	0.60	0.60
l	0.24	1	0	0.20	0	0	0.40	0.24	0.12	0.20	0.24	1	0.36	0.36
m	1	1	0	0.60	0	0	0.90	0.40	0.60	0.60	0.60	0.60	1	1
n	0	0.90	0	0.60	0	0	1	0.60	0.30	0.30	0.90	0.90	0.30	1

表 4 c 为 3 时最优模糊矩阵和相对隶属度表
Tab.4 Optimum fuzzy matrix and relative membership
table under $c=3$

部件	$H(u)$			分类结果
	聚类中心 1	聚类中心 2	聚类中心 3	
a	0.887 5	0.045 8	0.066 6	1.179 1
b	0.112 0	0.275 6	0.612 4	2.500 4
c	0.211 8	0.160 8	0.627 4	2.415 6
d	0.081 3	0.098 8	0.819 9	2.738 6
e	0.877 3	0.041 8	0.080 9	1.203 6
f	0.918 6	0.031 1	0.050 4	1.131 8
g	0.203 9	0.194 7	0.601 3	2.397 4
h	0.069 3	0.772 3	0.158 5	2.089 2
i	0.073 7	0.779 7	0.146 6	2.073 0
j	0.200 2	0.532 1	0.267 7	2.067 4
k	0.111 3	0.570 8	0.317 9	2.206 6
l	0.160 5	0.388 8	0.450 8	2.290 3
m	0.068 3	0.787 9	0.143 7	2.075 4
n	0.079 7	0.661 5	0.258 8	2.179 1

5 结 论

(1)提出用变个体长度的混合蛙跳算法对模糊 C 均值聚类算法的模糊聚类数和聚类中心同时进行优化. 针对 FCM 自身特点,采用切断算子和拼接算子对个体进行重新组合而形成新个体,

分析 $c=3$ 时的功能模块聚类结果,所研究金融机具纸币清分机的功能模块可归为 3 类: $\{(a, e, f); (b, c, d, g); (h, i, j, k, l, m, n)\}$. 这样就可以根据金融机具生产企业的实际情况和特点,选取某些模块作为产品族设计的核心平台,例如将电机模块、直线-滚轮式走钞模块以及进钞模块等作为构成产品平台的部件.

然后采用 ISODATA 迭代算法进行个体的局部寻优以加快收敛速度.

(2)建立了纸币清分机产品开发过程中的部件关联矩阵,采用所提改进的 SFLA-FCM 算法来进行产品族的设计. 仿真研究结果表明所提方法为产品族模块化设计提供了定量数学分析和快速配置的理论依据.

参考文献:

[1] Tseng M M, Jiao J X. A variant approach to product definition by recognizing functional requirement patterns [J]. **Journal of Engineering Design**, 1997, **8**(4):329-340.

[2] Arbelaiz O, Gurrutxag I, Muguerza J, *et al.* An extensive comparative study of cluster validity indices [J]. **Pattern Recognition**, 2013, **46**(1):243-256.

[3] Falasconi M, Gutierrez A, Pardo M, *et al.* A stability based validity method for fuzzy clustering [J]. **Pattern Recognition**, 2010, **43**(4):1292-1305.

[4] LI Chao-shun, ZHOU Jian-zhong, KOU Pan-gao, *et al.* A novel chaotic particle swarm optimization based fuzzy clustering algorithm [J]. **Neurocomputing**, 2012, **83**(15):98-109.

[5] HE Yao-yao, ZHOU Jian-zhong, KOU Pan-gao,

- et al.* A fuzzy clustering iterative model using chaotic differential evolution algorithm for evaluating flood disaster [J]. **Expert Systems with Applications**, 2011, **38**(8):10060-10065.
- [6] XIE Xuan-li, Beni G. A validity measure for fuzzy clustering [J]. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, 1991, **13**(8): 841-847.
- [7] Fukuyama Y, Sugeno M. A new method of choosing the number of clusters for the fuzzy *c*-means method [C] // **Proceedings of the International Conference of the Fifth Fuzzy System Symposium**. Kobe: The Japan Society of Applied Physics, 1989:247-250.
- [8] Kwon S H. Cluster validity index for fuzzy clustering [J]. **Electronics Letters**, 1998, **34**(22): 2176-2177.
- [9] Chen M Y, Linkens D A. Rule-base self-generation and simplification for data-driven fuzzy models [J]. **Fuzzy Sets and Systems**, 2004, **142**(2):243-265.
- [10] Alp Erilli N, Yolcu U, E ğrio ğlu E, *et al.* Determining the most proper number of cluster in fuzzy clustering by using artificial neural networks [J]. **Expert Systems with Applications**, 2011, **38**(3):2248-2252.
- [11] 王爱民, 孟明辰, 黄靖远. 聚类分析法在产品族设计中的应用研究[J]. 计算机辅助设计与图形学学报, 2003, **15**(3):343-347.
- WANG Ai-min, MENG Ming-chen, HUANG Jing-yuan. Research on clustering analysis for product family design [J]. **Journal of Computer-Aided Design & Computer Graphics**, 2003, **15**(3):343-347. (in Chinese)
- [12] Eusuff M M, Lansey K E. Optimization of water distribution network design using the shuffled frog leaping algorithm [J]. **Journal of Water Resources Planning and Management**, 2003, **129**(3):210-225.
- [13] Amiri B, Fathian M, Maroosi A. Application of shuffled frog-leaping algorithm on clustering [J]. **The International Journal of Advanced Manufacturing Technology**, 2009, **45**(1-2):199-209.
- [14] 杜长海, 黄席樾, 杨祖元, 等. 改进的 FCM 聚类在交通时段自动划分中的应用[J]. 计算机工程与应用, 2009, **45**(24):190-193.
- DU Chang-hai, HUANG Xi-yue, YANG Zu-yuan, *et al.* Application of improved fuzzy C-means clustering in automatic programming traffic intervals [J]. **Computer Engineering and Applications**, 2009, **45**(24):190-193. (in Chinese)
- [15] Kannan S R, Devi R, Ramathilagam S, *et al.* Some robust objectives of FCM for data analyzing [J]. **Applied Mathematical Modelling**, 2011, **35**(5): 2571-2583.
- [16] WANG Jie-sheng, GAO Xian-wen. Optimization of fuzzy C-means clustering by genetic algorithms based on sizable chromosome [C] // **Proceedings of 2009 Chinese Conference on Pattern Recognition**. New York:IEEE Press, 2009:1-5.

Product family design method based on fuzzy C-means clustering method optimized by improved shuffled frog leaping algorithm

CUI Wen-hua^{*1,2}, LIU Xiao-bing¹, WANG Wei¹, WANG Jie-sheng²

(1. School of Control Science and Engineering, Dalian University of Technology, Dalian 116024, China;

2. School of Electronic and Information Engineering, University of Science & Technology Liaoning, Anshan 114044, China)

Abstract: The decision of product platform in the modularized product family design is proposed based on fuzzy C-means (FCM) clustering algorithm optimized by the improved shuffled frog leaping algorithm (ISFLA). The component coefficient matrix based on the data collected in the product exploitation process is set up. The number of fuzzy clustering and cluster centers are optimized by sizable-individual SFLA to obtain the optimized fuzzy partition of the components of the product. Cut operator and splice operator are used to combine the individuals to form a new individual. ISODATA iterative algorithm is adopted to carry through the local optimization. The simulation results of the product family design of the paper currency sorter show that the proposed method provides the theoretical basis of the quantitative mathematic analysis and fast configuration for the product modularized design.

Key words: paper currency sorter; product family; product platform; shuffled frog leaping algorithm; fuzzy C-means clustering