

## 新的小生境萤火虫模糊聚类\*

李 丹<sup>1</sup>, 罗 可<sup>1</sup>, 孙 振<sup>2</sup>

(1. 长沙理工大学计算机与通信工程学院, 湖南 长沙 410114;

2. 郑州轻工业学院计算机与通信工程学院, 河南 郑州 450002)

**摘要:**模糊 C 均值算法因其简单、快速得到了广泛应用,但仍存在对初始值敏感和容易陷入局部最优的不足。提出了一种新的小生境萤火虫模糊聚类算法。该算法使用遍历性较好的立方混沌映射序列初始化萤火虫种群,并将随机惯性权重引入萤火虫算法,改变了基本萤火虫算法的位置更新公式,不仅减少了迭代次数,而且平衡了算法局部搜索和全局搜索的能力;并在迭代过程中合适时机实施小生境算法,进而增加了种群的多样性并加快了算法运算速度。仿真实验结果表明,该算法有效地抑制了早熟,并保证了种群的多样性和避免陷入局部最优,取得了较好的稳定性及良好的聚类结果。

**关键词:**立方映射;随机惯性权重;萤火虫;小生境技术

**中图分类号:**TP301

**文献标志码:**A

**doi:**10.3969/j.issn.1007-130X.2017.05.028

## New fuzzy clustering based on niching firefly

LI Dan<sup>1</sup>, LUO Ke<sup>1</sup>, SUN Zhen<sup>2</sup>

(1. School of Computer & Communication Engineering, Changsha University of Science & Technology, Changsha 410114;

2. School of Computer & Communication Engineering, Zhengzhou Institute of Light Industry, Zhengzhou 450002, China)

**Abstract:** The fuzzy C-means algorithm is widely used due to its simplicity and speediness. However, it is sensitive to the initial value and easy to fall into local optimum. We propose a new fuzzy clustering based on niching firefly. The algorithm utilizes the chaotic sequence to initialize the firefly population so as to obtain the initial population. The introduction of random inertia weight not only decreases the number of iterations, but also balances the global search ability and the local search ability of the algorithm. By implementing the niche in the process of the iteration algorithm, the diversity of population is increased and the algorithm's speed is accelerated. Simulation results show that the proposed algorithm can suppress precociousness effectively and ensure population diversity. It can also avoid falling into the local optimum and achieve good clustering performance.

**Key words:** cube mapping; random inertia weight; firefly; niche technology

### 1 引言

聚类分析是一个把数据对象(或观测)划分成子集的过程,其作为一种数据挖掘工具已广泛地应用于商务智能、生物学、Web 搜索等领域<sup>[1]</sup>。模糊

C-均值 FCM(Fuzzy C-Means)是一种经典的“软”聚类划分算法,然而其存在不足,如对初始值敏感和易陷入局部最优等。因此,众多学者对其研究并改进,部分学者将各种群智能优化算法如粒子群算法<sup>[2,3]</sup>、蚁群算法<sup>[4]</sup>、人工蜂群算法<sup>[5]</sup>等融合模糊 C 均值算法,取得了良好的效果。

\* 收稿日期:2015-06-23;修回日期:2016-02-22

基金项目:国家自然科学基金(71371065,11671125);湖南省自然科学基金衡阳联合基金(10JJ8008);湖南省科技计划(2013SK3146)

通信地址:410114 湖南省长沙市长沙理工大学计算机与通信工程学院

Address: School of Computer & Communication Engineering, Changsha University of Science & Technology, Changsha 410073, Hunan, P. R. China

萤火虫算法<sup>[6,7]</sup>是一种新型的启发式智能算法,该算法由自然界生物萤火虫的生活特性启发而来,2008年由剑桥学者 Yang 提出,由于其具有可调参数少、收敛速度快等优点得到较广泛的应用,如目标跟踪<sup>[8]</sup>、多目标函数问题、调度问题<sup>[9]</sup>、图像处理<sup>[10]</sup>等。与其他智能算法类似,基本萤火虫算法存在易陷入局部最优、发生早熟等不足。文献[11]提出了将文化算法融合萤火虫算法,取得了收敛速度和求解质量上的较好效果;文献[12]在萤火虫个体之间构建了信息交互网络,利用进化机制改善了萤火虫算法过早进化停滞的问题。

鉴于此,本文提出了一种新的小生境萤火虫模糊聚类,将性能较优的立方映射引入以获得初始化萤火虫种群,并为种群多样性奠定基础;利用随机惯性权重修改萤火虫个体位置更新公式,使探索能力和开发能力得到较好的平衡;适当时机引入的小生境技术不仅能摆脱陷入局部最优,而且可以提高种群的多样性。

## 2 模糊 C 均值算法

已知数据集  $X = \{x_1, x_2, \dots, x_n\}$ , 其包含  $n$  个元素,每个元素含  $d$  个属性。将  $X$  划分为  $c$  类( $2 \leq c \leq n$ ), 聚类中心为  $V = \{v_1, v_2, \dots, v_c\}$ , 则 FCM 的目标函数定义为:

$$\min J_F(U, V) = \sum_{i=1}^n \sum_{j=1}^c u_{ij}^m \|x_i - v_j\|^2 \quad (1)$$

其中,  $m$  为模糊加权指数,且其值一般取为 2;  $u_{ij}$  是样本  $i$  对样本  $j$  的隶属度 ( $i = 1, 2, \dots, n; j = 1, 2, \dots, c$ ), 并满足  $\sum_{j=1}^c u_{ij} = 1, u_{ij} \in [0, 1]; \|x_i - v_j\|$  为样本  $x_i$  与聚类中心  $v_j$  之间的欧氏距离。

FCM 算法的思想即通过对目标函数的不断迭代优化而得到对样本集的模糊分类并使目标函数最小<sup>[13]</sup>。其算法流程为:

**Step 1** 初始化及参数设置,设定模糊加权系数  $m$ , 聚类数目  $c$ , 初始化聚类中心  $V$ 。

**Step 2** 计算隶属度矩阵  $U$ :

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left( \frac{\|x_i - v_j\|^2}{\|x_i - v_k\|^2} \right)^{\frac{2}{m-1}}} \quad (2)$$

**Step 3** 更新聚类中心  $V$ :

$$v_j = \frac{\sum_{i=1}^n u_{ij}^m x_i}{\sum_{i=1}^n u_{ij}^m} \quad (3)$$

**Step 4** 当新得到的聚类中心与原聚类中心距离小于或等于给定阈值  $\epsilon$  ( $\epsilon > 0$ ), 则迭代停止;否则,转 Step 2。

FCM 算法迭代停止后,即可输出聚类中心  $V$

及隶属度  $U$ , 该算法收敛速度快且局部搜索能力强,其迭代本质即为利用梯度下降方法得到最优解,因此,聚类中心初始值的选择很重要,否则会陷入局部最优。这是经典 FCM 算法的不足,容易陷入局部最优及对初始聚类中心敏感等。

## 3 基本萤火虫算法

### 3.1 算法原理

萤火虫算法 FA (Firefly Algorithm) 是一种源于生物界的随机优化算法,利用萤火虫的发光并吸引同伴的特性,根据各萤火虫分布的位置赋予其目标函数值并计算相对吸引度,然后依据相对吸引度的大小决定各萤火虫移动方向,原则是荧光亮度弱的萤火虫个体将根据位置更新公式向荧光亮度较强的个体移动,萤火虫个体不断移动的过程即为问题优化的过程。算法中,优化过程基于以下三个原则:

(1) 萤火虫的性别因素忽略,任何两只萤火虫个体均可以互相吸引。

(2) 萤火虫个体的吸引度与亮度成正比,亮度强的萤火虫吸引亮度弱的萤火虫向其移动,亮度最强的个体随机移动;个体的吸引度与距离成反比,距离增大,吸引度与亮度减小。

(3) 萤火虫个体的亮度由其所在位置的目标函数值决定。

### 3.2 算法的数学描述

实现寻优的关键因素为亮度和吸引度,亮度的强弱由萤火虫个体的位置体现,位置越好目标函数值越好亮度越强,即可吸引亮度弱的个体向其移动。总之,移动大小由吸引度决定,移动方向由亮度决定。亮度和吸引度又与距离成反比,随着距离的增大,亮度和吸引度减小。算法实现过程中的几个定义:

(1) 荧光亮度:

$$I \propto -f(x_i), 1 \leq i \leq n \quad (4)$$

$$I = I_0 \exp(-\gamma r_{ij}) \quad (5)$$

其中,  $I$  为荧光亮度,  $f(x_i)$  为所求问题的目标函数,  $x_i$  为萤火虫  $i$  的空间位置;  $I_0$  为最大荧光亮度,  $\gamma$  为光强吸收因子,其值为常数,  $r_{ij}$  为  $x_i$  与  $x_j$  之间的距离,用公式表示为:

$$r_{ij} = \|x_i - x_j\| = \sqrt{\sum_{k=1}^d (x_{i,k} - x_{j,k})^2} \quad (6)$$

其中,  $d$  表示具体问题维数。

(2)吸引度:

$$\beta = \beta_0 \exp(-\gamma r_{ij}^2) \quad (7)$$

其中,  $\beta_0$  为最大吸引度。

(3)位置更新:

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \beta(\mathbf{x}_j(t) - \mathbf{x}_i(t)) + \alpha \boldsymbol{\varepsilon}_i \quad (8)$$

其中,  $\alpha$  表示初始化的步长因子;  $\boldsymbol{\varepsilon}_i$  表示服从高斯分布的随机因子<sup>[6,7]</sup>。

## 4 新的小生境萤火虫模糊聚类算法

### 4.1 小生境技术

小生境技术<sup>[14,15]</sup>是来源于生物学的小生境的理念,其生物学概念为,生活在一起的同种类的生物形成的小的生活环境,进而使不同种类的个体产生分离。这种思想应用到计算机科学领域,即为具有相同数据特征的数据被归为同一类,不同特征的数据将会被分离,因此避免了大量数据集于局部最优点周围,提高了种群多样性。提出小生境思想的是Holland,之后不同的小生境技术相继被提出,即排挤、共享、标签等。

其中,基于排挤机制的选择策略为:

(1)设定排挤因子  $CF$ (取值为 2 或 3);

(2)从种群中随机选择种群规模的  $1/CF$  个个体的参与排挤运算;

(3)得到新个体,计算新个体与排挤个体之间的海明距离;

(4)依据新个体与排挤个体的相似性排挤掉部分与排挤个体类似的个体。

小生境技术不仅具有跳出局部最优的能力,而且能够保持种群多样性,提高求解精度。然而,基本萤火虫算法在迭代后期容易陷入局部最优且存在种群多样性下降的问题,因此,利用小生境技术的优点来弥补基本萤火虫算法的缺点,本文将基于排挤选择策略的小生境技术融入基本萤火虫算法,提出了小生境萤火虫模糊聚类算法。

### 4.2 算法中的关键问题

(1)混沌初始化种群。

一般来说,萤火虫个体位置的初始化是随机分布的,这可能会导致其位置分布不均匀。因此,本文引入具有规律性和遍历性特点的混沌映射<sup>[16]</sup>,得到混沌初始化种群,不仅避免算法陷入局部最优,同时提高了种群多样性。本文选用的混沌映射为立方映射,而没有选用较常用的 Logistic 映射,因为相比 Logistic 映射,立方映射具有较优的序列

均匀性。

混沌初始化的步骤为:

①得到第一个个体。随机产生一个  $d$  维向量,即为第一个萤火虫个体,  $\mathbf{Y} = \{y_1, y_2, \dots, y_d\}$ , 其中,  $y_i \in [-1, 1], 1 \leq i \leq d$ 。

②得到其余  $n-1$  个个体。对  $\mathbf{Y}$  的每一维用式(9)进行  $n-1$  次迭代;

$$y(n-1) = 4y(n)^3 - 3y(n) \quad (9)$$

其中,  $-1 \leq y(n) \leq 1, n = 0, 1, 2, \dots$ 。

③将混沌变量映射到解空间。利用式(10)进行映射。

$$x_{id} = O_d + (1 + y_{id}) \frac{R_d - O_d}{2} \quad (10)$$

其中,  $x_{id}$  表示第  $i$  个萤火虫在解空间第  $d$  维坐标;  $O_d$  表示解空间第  $d$  维的下限;  $y_{id}$  表示由式(9)得到的第  $i$  个个体的第  $d$  维;  $R_d$  表示解空间第  $d$  维的上限。

(2)将随机惯性权重引入位置更新公式。

惯性权重<sup>[17]</sup>的概念是由 Shi 和 Eberhart 首次提出的,其是一个影响算法性能的关键参数,不但能够平衡开发能力和探索能力,而且具有保持种群多样性的作用。故惯性权重的选择极为重要,在大多数改进中,线性递减方法是一种常见的选择,虽然迭代初期的惯性权重较大值有利于全局搜索,但是容易导致搜索效率低下且增大算法开销;迭代后期的惯性权重较小值有利于加快算法收敛,却容易陷入局部最优而难以跳出。然而,若将惯性权重设为服从某种分布的随机数,随机变量的随机性将会使惯性权重在运算初期取得较小或较大的值。若某个体离最优个体较近,则会产生较小的权重值,进而加快收敛速度;假定产生了较大值,则得到了较大的适应函数值,劣于最优值,该权重值即被淘汰。同理,当个体离最优个体较远时,原理相同。因此,本文将随机惯性权重引入到位置更新公式中,即:

$$\mathbf{x}_i(t+1) = \omega \cdot \mathbf{x}_i(t) + \beta(\mathbf{x}_j(t) - \mathbf{x}_i(t)) + \alpha \boldsymbol{\varepsilon}_i \quad (11)$$

$$\omega = \mu_{\min} + (\mu_{\max} - \mu_{\min}) \times \text{rand}() + \sigma \times \text{randn}() \quad (12)$$

其中,  $\mu_{\min}$ 、 $\mu_{\max}$  分别为权重值的最小值和最大值;  $\text{rand}()$  为  $[0, 1]$  上均匀分布的随机数;  $\sigma$  是衡量随机变量权重  $\omega$  与其数学期望(即均值)之间的偏离程度,它是为了控制取值中的权重误差,使权重  $\omega$  向期望权重方向进化,因为正常情况下实验误差服从正态分布;  $\text{randn}()$  为正态分布的随机数。

(3)算法的终止条件。

当达到预先设定的最大迭代次数或者参与迭代的个体适应函数值连续多次变化范围均小于某值(常数  $\chi$ , 可设为 0.01%), 即可终止迭代。

(4)引入小生境技术的时机选择。

为了萤火虫算法和小生境技术较好地结合, 就需要选择合适的操作时机。迭代初期, 基本萤火虫算法进行全局搜索, 没有必要引入小生境技术。然而, 一旦个体出现集聚现象, 可以看出多数个体适应函数值比较接近, 故个体之间的集聚程度需要确定一个变量来表示, 如下:

$$\varphi^2 = - \sum_{j=1}^m \left( \frac{f_i - f_{\text{avg}}}{f} \right)^2 \quad (13)$$

其中,  $m$  为个体数,  $f_i$  为第  $i$  个个体的适应函数值,  $f_{\text{avg}}$  为萤火虫群当前的平均适应度。设定一足够小阈值  $q$ , 当  $\varphi^2 < q$  时, 说明萤火虫群集聚程度较高, 是引入小生境技术的适当时机。即随机生成当前群个体数目 20% 的新个体, 然后进行小生境技术排挤策略, 进而排挤掉不符合条件的个体, 从而种群的多样性得到保持。

#### 4.3 算法流程

**Step 1** 初始化算法的各参数, 利用混沌初始化得到的个体作为初始种群;

**Step 2** 利用式(1)计算每个个体目标值作为其最大亮度  $I_0$ ;

**Step 3** 根据式(5)和式(7)分别计算各个萤火虫之间的相对亮度  $I$  和吸引度  $\beta$ , 依照  $I$  的大小确定萤火虫的移动方向, 利用式(11)对萤火虫的位置进行更新;

**Step 4** 根据更新位置后的萤火虫的位置, 重新计算萤火虫的亮度;

**Step 5** 判断种群集聚程度, 若达到设定阈值则执行小生境算法, 否则转 Step 6;

**Step 6** 判断算法是否已经达到终止条件, 若达到则输出最优划分聚类, 否则转至 Step 3。

#### 4.4 聚类效果判断指标

聚类效果是判断一个算法优劣的重要指标, 本文采取了基于香农信息熵的聚类有效性函数<sup>[18]</sup>作为衡量指标, 该指标越小, 说明聚类效果越好。其定义为:

$$Hp(U; c) = H_1(U; c) - H_2(U; c) \quad (14)$$

其中,

$$H_1(U; c) = - \sum_{k=1}^n \sum_{i=1}^c \frac{u_{ik}}{n} \ln \frac{u_{ik}}{n},$$

$$H_2(U; c) = - \sum_{i=1}^c \sum_{k=1}^n \frac{u_{ik}}{(c \sum_{i=1}^c u_{ik}) \ln(u_{ik} / (c \sum_{k=1}^n u_{ik}))} \quad (15)$$

其中,  $c$  为聚类中心数目,  $U$  为模糊隶属度矩阵。

## 5 实验分析

本文选取了 UCI 标准数据库的三个数据集 Iris、Wine、Glass 对本文算法(以下简称为 INAFACM 算法)的性能进行验证, 并将其实验结果与 FCM 算法、文献[3]算法(以下简称为 MP-SOFCM 算法)和基于改进萤火虫的模糊聚类算法(以下简称为 MFAFCM 算法)进行比较, 涉及的数据集特征如表 1 所示。

**Table 1 Characteristics of the data sets in the experiments**  
表 1 实验涉及数据集特征

数据集名称	类别数	属性维数	样本数目
Iris	3	4	450
Glass	6	9	214
Wine	3	13	178

实验环境为: 软件: 操作系统 Windows 7, Microsoft Visual C++ 6.0, Matlab 2012; 硬件: Intel Core i5 CPU, 内存 4 GB。算法中参数设置为: 各算法中模糊加权指数为  $m=2$ ; MP-SOFCM 算法中参数设置为: 粒子群体规模  $N=20$ ,  $\omega \varphi_1$  从 0.9 线性减小到 0.4,  $\eta_1 = \eta_2 = 0.5$ ; INAFACM 算法的参数设置为: 种群规模  $N=20$ , 最大迭代次数  $maxT=500$ , 步长因子  $\alpha=0.5$ , 最大吸引度  $\beta_0=1$ , 介质吸收因子  $\gamma=1$ ,  $\mu_{\text{max}}=0.95$ ,  $\mu_{\text{min}}=0.5$ 。将算法在各数据集上运行 20 次, 得到各项结果, 如表 2~表 4 所示。

**Table 2 Indices comparison among algorithms on the Iris data**  
表 2 各算法在 Iris 数据集上各指标比较

算法	$J_F(\text{means})$	准确率	迭代次数	$HP(U, c)$
FCM	67	89.42	26	0.005
MP-SOFCM	53	91.08	52	-0.002
MFAFCM	54	90.36	48	-0.002
INAFACM	49	92.53	23	-0.003

**Table 3 Indices comparison among algorithms on the Glass data**

表 3 各算法在 Glass 数据集上的各指标比较

算法	$J_F(\text{means})$	准确率	迭代次数
FCM	156	47.69	53
MP-SOFCM	179	76.12	101
MFAFCM	175	80.28	95
INAFACM	153	81.79	49

**Table 4 Indices comparison among algorithms on the Wine data**

**表 4 各算法在 Wine 数据集上各指标比较**

算法	$J_F(\text{means})$	准确率	迭代次数	$HP(U, c)$
FCM	610	70.11	43	0.712
MPSOFCM	589	74.02	75	0.649
MFAFCM	581	76.43	79	0.634
INAFACM	573	78.51	68	0.621

由各实验数据来看,各表包含了目标函数值的平均值、聚类准确率、所用迭代次数和聚类有效性值。

因此,总的来说,INAFACM 在三个数据集上在各指标(目标函数均值、迭代次数、准确率、聚类有效性值)均得到了较优的结果,且针对某数据集甚至有更优秀的表现。在 Iris 数据集上各算法指标上差距不大,尤其是准确率指标,但 INAFACM 用了比其他三种算法少的迭代次数,体现了算法的优越性。从 Glass 数据集上来看,改进算法在准确率指标上较原始算法都得到了很大提高,INAFACM 算法所用迭代次数优势虽不是特别明显,但较其他三种算法来说还是有所改善。在 Wine 数据集上,虽然各算法在各指标上差距不明显,但是在迭代次数和聚类有效性值指标仍优于其他三种算法。

本文给出了各算法在 Iris、Glass、Wine 数据集上的迭代曲线图,如图 1~图 3 所示。

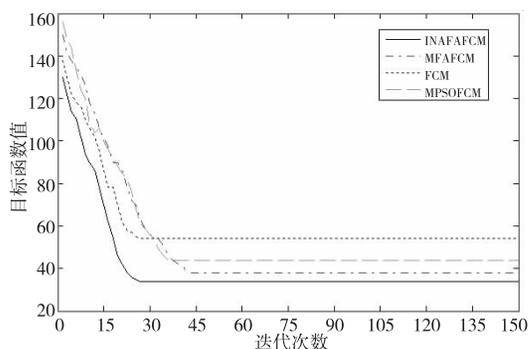


Figure 1 Comparison of the curves of the objective function for Iris data set using the four algorithms

图 1 各算法在 Iris 数据集上的迭代曲线图

由图 1~图 3 可以看出,本文算法在三个数据集上的目标函数值均比其他三种算法小,较快地找到局部极值,并跳出局部最优且随着迭代次数增加,逐渐趋于稳定,最终找到全局最优。通过图中可以看到本文算法的稳定性得到验证,由此可见,本文算法具有稳定的性能以及平衡的开发和探索能力,使算法避免陷入局部最优。

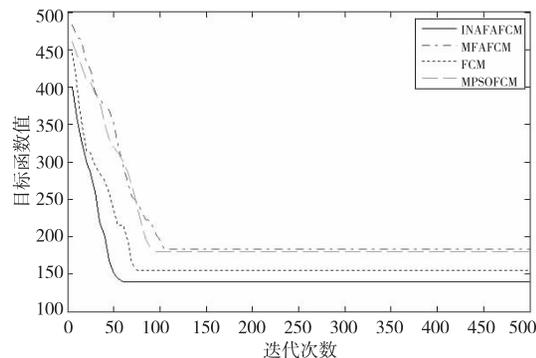


Figure 2 Comparison of the curves of the objective function for Glass data set using the four algorithms

图 2 各算法在 Glass 数据集上的迭代曲线图

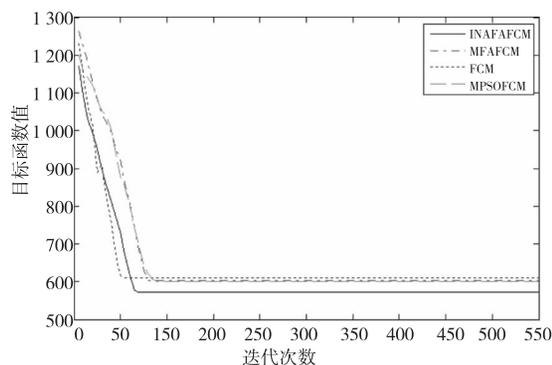


Figure 3 Comparison of the curves of the objective function for Wine data set using the four algorithms

图 3 各算法在 Wine 数据集上的迭代曲线图

## 6 结束语

本文提出了一种新的小生境萤火虫模糊聚类算法,弥补了基本萤火虫算法后期收敛慢及容易陷入局部最优的不足,解决了模糊 C-均值算法对初始值敏感的问题。通过对三个 UCI 标准数据集进行测试,并与 FCM 算法、MPSOFCM 算法及 MFAFCM 算法进行对比,实验结果表明,本文提出的新的小生境萤火虫模糊聚类算法取得了较好的聚类效果,并且具有避免陷入局部最优的能力及有效平衡探索和开发的能力。

### 参考文献:

- [1] Han Jia-wei, Kan Bo, Pei Jian. Data mining concepts and techniques [M]. 3rd Edition. Beijing, China Machine Press, 2012. (in Chinese)
- [2] Wang Lin, Luo Ke, Luo Yong-hong. Fuzzy clustering based on predator prey particle swarm optimization [J]. Computer Engineering and Applications, 2008, 44 (35): 129-131. (in Chinese)
- [3] Wang Zong-hu, Liu Zhi-jing, Chen Dong-hui. Research of PSO-based fuzzy C-means clustering algorithm [J]. Computer

- Science,2012,39(9):166-169. (in Chinese)
- [4] Niknam T, Amiri B. An efficient hybrid approach based on PSO, ACO and  $k$ -means for cluster analysis[J]. Applied Soft Computing,2010,10(1):183-197.
- [5] Luo Ke, Li Lian, Zhou Bo-xiang. Artificial bee colony rough clustering algorithm based on mutative precision search[J]. Control and Decision,2014,29(5):838-842. (in Chinese)
- [6] Yang Xin-she. Nature-inspired metaheuristic algorithms [M]. [S. l.]:Luniver Press,2008:83-96.
- [7] Yang Xin-she. Firefly algorithms for multimodal optimization [C]//Proc of the 5th International Conference on Stochastic Algorithms, Foundations and Applications,2009:169-178.
- [8] Gao Ming-liang, He Xiao-hai, Luo Dai-sheng, et al. Object tracking using firefly algorithm[J]. IET Computer Vision, 2013,7(4):227-237.
- [9] Nithya G, Jayapratha G. A multi-agent brokering approach and jumper firefly algorithm for job scheduling in cloud computing[C]//Proc of International Conference of Intelligent Computing Applications,2014:52-58.
- [10] Napoli C, Pappalardo G, Tramontana E, et al. Simplified firefly algorithm for 2d image key-points search [C]//Proc of Computational Intelligence for Human-like Intelligence (CIHLI),2014:1-8.
- [11] Yuan Feng, Chen Shou-qiang, Liu Hong, et al. An improved clutural firefly algorithm[J]. Computer Simulation,2014,31(6):261-265. (in Chinese)
- [12] Fu Qiang, Tong Nan, Zhong Cai-ming, et al. Firefly algorithm based on evolutionism[J]. Computer Science,2014,41(3):228-231. (in Chinese)
- [13] Bezdek J C, Ehrlich R, Full W. FCM: The fuzzy  $c$ -means clustering algorithm[J]. Computers & Geosciences,1984,10(2):191-203.
- [14] Zhang Jun, Huang De-shuang, Lok T M, et al. A novel adaptive sequential niche technique for multimodal function optimization[J]. Neurocomputing,2006,69(16/18):2396-2401.
- [15] Tang Chao-li, Huang You-rui, Li Ji-jun. Niche particle swarm algorithm and application study[M]//Future Control and Automation. Berlin, Springer Berlin Heidelberg,2012:57-65.
- [16] Xu Xiao-bo, Zheng Kang-feng, Li Dan, et al. New chaos-particle swarm optimization algorithm[J]. Journal on Communications,2012,33(1):24-30. (in Chinese)
- [17] Arasomwan M A, Adewumi A O. On adaptive chaotic inertia weights in particle swarm optimization[C]//Proc of Swarm Intelligence (SIS),2013:72-79.
- [18] Qing Ming, Sun Xiao-mei. A new clustering effectiveness

function; Fuzzy entropy of fuzzy partition[J]. CAAI Transactions on Intelligent Systems,2015,10(1):75-79. (in Chinese)

### 附中文参考文献:

- [1] 韩家炜, 堪博, 裴健. 数据挖掘: 概念与技术[M]. 第3版. 北京: 机械工业出版社, 2012.
- [2] 王琳, 罗可, 罗永红. 基于捕食-被捕食粒子群优化的模糊聚类[J]. 计算机工程与应用, 2008, 44(35):129-131.
- [3] 王纵虎, 刘志镜, 陈东辉. 基于粒子群优化的模糊  $C$ -均值聚类算法研究[J]. 计算机科学, 2012, 39(9):166-169.
- [5] 罗可, 李莲, 周博翔. 基于变异精密搜索的蜂群聚类算法[J]. 控制与决策, 2014, 29(5):838-842.
- [11] 袁锋, 陈守强, 刘弘, 等. 一种改进的文化萤火虫算法[J]. 计算机仿真, 2014, 31(6):261-265.
- [12] 符强, 童楠, 钟才明, 等. 基于改进型进化机制的萤火虫优化算法[J]. 计算机科学, 2014, 41(3):228-231.
- [16] 胥小波, 郑康锋, 李丹, 等. 新的混沌粒子群优化算法[J]. 通信学报, 2012, 33(1):24-30.
- [18] 卿铭, 孙晓梅. 一种新的聚类有效性函数: 模糊划分的模糊熵[J]. 智能系统学报, 2015, 10(1):75-79.

### 作者简介:



**李丹**(1988-), 女, 河南封丘人, 硕士生, 研究方向为数据挖掘。E-mail: lid2014happy@126.com

**LI Dan**, born in 1988, MS candidate, her research interest includes data mining.



**罗可**(1961-), 男, 湖南长沙人, 博士, 教授, 研究方向为数据挖掘和计算机应用。E-mail: luok@csust.edu.cn

**LUO Ke**, born in 1961, PhD, professor, his research interests include data mining, and computer application.



**孙振**(1989-), 男, 河南延津人, 硕士生, 研究方向为智能信息处理。E-mail: zz-sunzhen@163.com

**SUN Zhen**, born in 1989, MS candidate, his research interest includes intelligent information processing.