

引用格式:靳雁霞,齐欣,张晋瑞,等.一种改进的简化均值粒子群 K-means 聚类算法[J].微电子学与计算机,2020,37(5):69-74. [JIN Y X,QI X,ZHANG J R, et al. An improved simplified mean particle swarm optimization K-means clustering algorithm[J]. Microelectronics & Computer,2020,37(5):69-74.]

一种改进的简化均值粒子群 K-means 聚类算法

靳雁霞,齐欣,张晋瑞,程琦甫

(中北大学 大数据学院,山西 太原 030051)

摘 要:针对粒子群优化算法容易陷入局部最优且 K-means 算法受聚类数及初始聚类中心的选取影响较大,提出了一种改进的简化均值粒子群 K-means 优化聚类算法(ISMPSO-AKM)。一方面,在简化粒子群算法的基础上,加入邻域最优粒子,由个体最优位置、全局最优位置及邻域最优位置线性组合改进位置公式。另一方面,构造一种基于余弦函数和对数函数的惯性权重,实现对惯性权重的动态调整。此外,引入 AKM 聚类算法确定聚类数,动态获取初始中心,进一步提高算法的准确性。仿真实验表明,改进的 ISMPSO-AKM 算法具有更快的收敛速度,更高的求解精度及更稳定的聚类结果。

关键词:粒子群优化算法;简化粒子群;邻域最优粒子;K-means 聚类;聚类数;初始聚类中心

中图分类号: TP18;TP391.41

文献标识码: A

文章编号: 1000—7180(2020)05—0069—06

An improved simplified mean particle swarm optimization K-means clustering algorithm

JIN Yan-xia,QI Xin,ZHANG Jin-rui,CHENG Qi-fu

(College of Bigdata,North University of China, Taiyuan 030051, China)

Abstract: To figure out the problems such as the fact that particle swarm optimization algorithm is easy to fall into local optimum and K-means algorithm is greatly influenced by the number of clusters and the selection of initial cluster centers, an improved simplified mean particle swarm optimization K-means clustering algorithm (ISMPSO-AKM) is proposed. On the one hand, on the basis of simplified particle swarm optimization, the neighborhood optimal particle is added to improve the position formula by linear combination of individual optimal position, global optimal position and neighborhood optimal position. On the other hand, an inertia weight based on cosine function and logarithmic function is constructed to realize dynamic adjustment of inertia weight. In addition, AKM clustering algorithm is introduced to determine the number of clusters and dynamically obtain the initial center, which further improves the accuracy of the algorithm. The simulation results show that the improved ISMPSO-AKM algorithm has faster convergence speed, higher accuracy and more stable clustering results.

Key words: particle swarm optimization algorithm; simplified particle swarm optimization; neighborhood optimal particle;k-means clustering; cluster number; initial cluster center

1 引言

数据挖掘^[1](data mining, DM),即数据库的知

识发现过程,是从数据库大量的数据中,提取出未知的、有潜在价值的信息和知识,并用这些信息解决相关问题。同物以类聚,聚类分析是按照一定的规律,

将数据集中未标记类别的数据划分为多个不相交的分类或簇的过程. 其中, K-means 聚类算法 (K-means clustering algorithm, KM) 作为经典聚类算法之一, 因其具有原理简单、易于实现、可解释度较强等优点, 被广泛应用于统计学、机器学习、市场营销等众多领域. 由于传统的 KM 算法存在 K 值很难确定、对噪声数据敏感、过于依赖初始中心的选取等缺点, 导致算法很难得到全局最优解, 对不同类别规模差异较大的数据或非凸数据集聚类效果不佳. 针对上述情况, 近年来研究者们提出了一系列的改进 KM 算法^[2-4].

粒子群优化算法 (Particle Swarm Optimization, PSO) 是由 Kennedy 和 Eberhart^[5] 通过研究鸟类行为规律得到的一种全局优化算法. PSO 算法因可调参数少、结构简单、效率较高等优点已经广泛应用于多目标优化、决策调度、机器人实时路径规划等各类优化问题. 由于基本粒子群算法存在早熟收敛、局部寻优能力较差、对离散的优化问题处理不佳等缺点, 近年来对 PSO 优化聚类算法的改进陆续被提出^[6-8].

为了提高 PSO 算法的探测开发能力, 进一步改善粒子群聚类算法的稳定性, 本文提出了一种改进的简化均值粒子群 K-means 优化聚类算法 (improved simplified mean particle swarm optimization and K-means clustering algorithms, ISMP SO-AKM). 在简化均值粒子群算法中加入邻域最优粒子, 此外, 提出了一种惯性权重的动态调整策略, 由余弦函数和对数函数对惯性权重进行改进, 避免了算法在后期由于多样性缺失而陷入早熟收敛.

2 相关工作

2.1 粒子群优化算法 (PSO)

在 D 维搜索空间中, 设种群总数为 m , 每个粒子可视为优化问题的一个可行解, 由一个事先设定的适应度函数来判断粒子的好坏. 用 $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})$ 、 $V_i = (v_{i1}, v_{i2}, \dots, v_{iD})$ 来表示粒子 i 的当前位置和当前速度. 在每次迭代中, $p_i = (p_{i1}, p_{i2}, \dots, p_{iD})$ 表示粒子迄今为止经历的最优位置, $p_g = (p_{g1}, p_{g2}, \dots, p_{gD})$ 表示整个种群迄今为止经历的最优位置. 粒子的速度和位置的更新公式如下:

$$v_{id}^{t+1} = \omega v_{id}^t + c_1 r_1 (p_{id} - x_{id}) + c_2 r_2 (p_{gd} - x_{id}) \quad (1)$$

$$x_{id}^{t+1} = x_{id}^t + v_{id}^{t+1} \quad (2)$$

式中, 下标 $i = 1, 2, \dots, m$; 下标 $d = 1, 2, \dots, D$; t 是当前迭代次数; ω 是惯性权重, c_1 和 c_2 是学习因子, $r_1, r_2 \sim U[0, 1]$ 的随机数. 为了使粒子在搜索空间内充分搜索, 在迭代过程中限定粒子速度向量上下限为 $[v_{\min}, v_{\max}]$, 位置向量上下限为 $[x_{\min}, x_{\max}]$.

2.2 K-means 聚类算法

K-means 算法的基本思想是: 从含有 n 数据对象的样本数据集 $X = \{x_i | i = 1, 2, \dots, n\}$ 中随机选取 K 个初始中心点. 计算剩余样本与每个聚类中心之间的距离, 并将其分配给最近的聚类, 重新计算每个簇中所有点的均值作为新的聚类中心. 循环上述过程直到聚类中心不再改变或者达到预定的最大迭代次数为止. 下面介绍 K-means 算法过程中出现的几个概念.

(1) 两个数据对象 x_i 和 x_j 之间的欧氏距离为 $d(x_i, x_j)$:

$$d(x_i, x_j) = \sqrt{(x_i - x_j)^T (x_i - x_j)} \quad (3)$$

(2) $c_j (j = 1, 2, \dots, K)$ 表示簇 $C_j (j = 1, 2, \dots, K)$ 的质心, 也称聚类中心.

$$c_j = \frac{1}{n_j} \sum_{x_i \in C_j} x_i \quad (4)$$

式中, n_j 是簇 j 的样本个数, x_i 是属于类 C_j 的样本.

(3) 误差平方和 SSE (sum of squared error)

$$E = \sum_{j=1}^K \sum_{x_i \in C_j} |x_i - c_j|^2 \quad (5)$$

通常将误差平方和作为聚类准则函数, 用来评价聚类结果的好坏. 可知, E 代表一个全局误差, E 的值越小, 表示生成的每个簇内越紧凑, 簇间越独立, 得到的聚类结果越好.

3 改进的简化均值粒子群算法与 AKM 聚类算法

3.1 编码方式

PSO 算法采用实数编码, 用二维矩阵 $X[k][d]$ 表示每个粒子, 其中, $k = 1, 2, \dots, K$, K 为聚类个数; $d = 1, 2, \dots, D$, D 为样本向量维数. 如下所示:

$$X_{11} X_{12} \dots X_{1D} X_{21} X_{22} \dots X_{2D} X_{K1} X_{K2} \dots X_{KD}$$

粒子适应度根据式 (6) 计算:

$$f(x) = c / \sum_{j=1}^K \sum_{x_i \in C_j} d(x_i, c_j) \quad (6)$$

式中, c 是常数; $d(x_i, c_j)$ 表示样本 x_i 与对应簇中心 c_j 之间的欧式距离. 可知, 总的类间离散度和越小, $f(x)$ 的值越大, 最终得到的聚类效果越好.

3.2 改进的简化均值粒子群算法(ISMPSO)

3.2.1 邻域最优粒子

邻域最优粒子^[9],用 n_{best} 表示.一般满足两个条件:一个是它接近正在更新的粒子,另一个是它具有较高的适应度值.适应度距离比(Fitness-Distance-Ratio, FDR)定义为邻域粒子适应度值的差与一维距离的比.通常, n_{best} 选取具有最佳 FDR 参数的粒子.对于求最大值问题,选择使 FDR 最大的粒子作为邻域最优粒子.适应度距离比的定义如下:

$$\frac{\text{Fitness}(P_j) - \text{Fitness}(X_i)}{|P_{jd} - X_{id}|} \quad (7)$$

式中, $|\cdots|$ 表示取绝对值; i 表示第 i 个粒子; d 表示维度; P_j 表示邻域粒子经历过的最好位置.对于求最小值问题,将式(7)中的分子项替换为 $(\text{Cost}(P_j) - \text{Cost}(X_i))$.

3.2.2 简化粒子群优化算法(PSO)

文献[10]去掉了速度项,提出一种简化粒子群优化算法.位置更新公式如下:

$$x_{id}^{t+1} = \omega x_{id}^t + c_1 r_1 (p_{id} - x_{id}) + c_2 r_2 (p_{gd} - x_{id}) \quad (8)$$

可以看出,进化方程由之前的二阶降为一阶,使算法更加的简化和高效.

3.2.3 ISMPSO 的位置更新策略

为了进一步提高算法的收敛速度和精度,本文结合邻域最优粒子与均值粒子群算法^[11]的思想,对式(8)位置公式进行改进,新的位置公式如下:

$$x_{id}^{t+1} = \omega x_{id}^t + c_1 r_1 \left(\frac{p_{id} + p_{gd} + p_{nd}}{3} - x_{id} \right) + c_2 r_2 \left(\frac{p_{id} - p_{gd}}{2} - x_{id} \right) + c_3 r_3 \left(\frac{p_{nd} - p_{gd}}{2} - x_{id} \right) \quad (9)$$

式中, c_3 是学习因子; r_3 是 $[0,1]$ 之间的随机数.

式(9)中位置公式扩展为四项,加入邻域最优粒子,有利于粒子获取更为丰富的信息来更新自身位置:第一项表示粒子先前位置对粒子位置更新的影响,用惯性权重 ω 调节影响程度;第二项为“认知”项,表示粒子个体最优位置、全局最优位置和邻域最优位置的平均位置对粒子位置更新的影响.第三项为“社会”项,表示粒子个体最优位置与全局最优位置负方向的平均位置对粒子位置更新的影响.第四部分为“超社会”项,表示邻域最优位置与全局最优位置负方向的平均位置对粒子位置更新的影响.

3.2.4 动态惯性权重

近年来对惯性权重的改进大体可以分为三类:一是常数和随机惯性权重;二是基于时间或迭代次

数变化的惯性权重,可以是线性的,也可以是非线性的;三是自适应惯性权重.在上述方法的启发下,通过引入余弦函数和对数函数,提出了惯性权重的动态调整策略,具体公式如下:

$$\omega = \omega_{\min} \times \left(-\ln \frac{t}{T_{\max}} \right) + (\omega_{\max} - \omega_{\min}) \times \cos \left(\frac{\pi t}{2T_{\max}} \right) + \sigma \times (-\ln \text{rand}())^{1/2} \cos(\pi \text{rand}()) \quad (10)$$

式中, t 是当前迭代次数; $T_{\max}$ 是最大迭代次数; $\text{rand}() \sim U(0,1)$ 的随机数; $\omega_{\min}$ 为最小惯性权重值; $\omega_{\max}$ 为最大惯性权重值.通过调整对数函数与余弦函数,使表达式中前两项 ω 的取值始终在一定的合理范围内,第三项为随机调整策略,目的是调整 ω 整体取值,用调整因子 σ 来控制 ω 的偏离程度.改进后的惯性权重呈非线性递减的趋势,同时加入随机调整策略,有效增强了算法在初期全局搜索的能力及后期局部搜索的能力.

3.3 AKM 聚类算法

为了获得稳定的聚类结果,本文提出了一种 KM 优化算法(improved K-means optimization algorithm, AKM),用以确定 K 值并动态获取初始中心.

3.3.1 K 值的选取

在 2.2 节提到过计算误差平方和(SSE)的公式,可知,随着聚类个数的增加,每一类包含的样本数会减少,SSE 会降低.当 SSE 明显减少时相应的 K 值被定义为最佳簇个数.但是在实际情况中,很难找到一个 SSE 明显下降的点.聚类数通常为整数,范围一般在 $2 \sim \lfloor \sqrt{n} \rfloor$ (n 为样本个数),本文通过遍历这一范围聚类数的聚类指标 SE 来确定 K 值,SE 的公式定义如下:

$$\text{SE} = \sqrt{K \times E} = \sqrt{K \times \sum_{j=1}^K \sum_{x_i \in C_j} |x_i - c_j|^2} \quad (11)$$

式中, $K = 2, 3, \dots, \lfloor \sqrt{n} \rfloor$ (n 为样本个数).记录 K 取不同值时 SE 的大小,找到 SE 取最小值时对应的点,即为所求的最优聚类个数.

3.3.2 初始聚类中心的改进

为了找到一组合适的初始聚类中心点,本文将样本数据集 X 的分布空间均匀划分为 $\lfloor \sqrt{n} \rfloor$ 个子空间,并统计子空间包含的点数目.定义 ϵ 为子空间内含有的最小点数目,记点数目不少于 ϵ 的子空间为

核心子空间,将满足条件的子空间存放在核心子空间集合 S 中,在集合 S 中依次选取距离最远的 K 个子空间的中心点作为初始聚类中心.

AKM 算法具体步骤如下:

输入:数据集 X 数据对象为 n 个,阈值 ϵ

输出: K 值,初始聚类中心

(1)确定聚类个数 K

①聚类数目的取值范围为 $[2, [\sqrt{n}]]$;

②根据式(11)计算每一个 K 值所对应的 SE 的值;

③记录并找到最小的 SE,对应的 K 值即为所要求的最佳簇个数.

(2)初始聚类中心的选取

①将样本数据集的分布空间均匀划分为 $\sqrt{n}$ 个子空间,如果核心子空间数量小于分类数,则调整阈值 ϵ ;

②统计每个子空间 s 中包含的样本点数目,并将点数目不小于 ϵ 的子空间记为核心子空间,存放集合 S 中;

③从集合 S 中找到具有最大点数目的核心子空间,计算该子空间内样本点的平均距离,设为第一个初始聚类中心 c_1 ,将 c_1 加入到初始中心集合,从 S 中删除该子空间;

④计算 c_1 与其他核心子空间中心距离和,找到离 c_1 最远的核心子空间,并计算得到的聚类中心 c_2 ,将 c_2 加入到初始中心集合,从 S 中删除该子空间;

⑤重复 ③④,直到找到第 K 个初始聚类中心点 c_k ,得到初始聚类中心 $c_1, c_2, \dots, c_k$.

3.4 ISMPSO-AKM 算法流程

输入:数据集 X

输出:划分的 K 个簇中心解集 c_j

(1)初始化种群及参数.根据 3.3 节中的 AKM 算法来确定 K 值,得到初始簇中心 $c_1, c_2, \dots, c_k$ 并作为粒子的初始位置编码.

(2)计算每个粒子的适应度值 $f(x)$,比较更新个体最优位置 p_{id} 、全局最优位置 p_{gd} ,根据式(7)更新邻域最优位置 p_{nd} .

(3)由式(10)计算惯性权重的值,并根据式(9)更新粒子的位置.

(4)对于每一个样本数据,按照聚类中心最邻近法则,确定对应聚类划分;

(5)重新计算得到每个类新的聚类中心,并根据式(6)重新计算粒子的适应度值.

(6)判断是否达到结束条件,以聚类中心不再变

化或者达到预定的 $T_{\max}$ 为算法停止准则,若满足,停止迭代并输出全局最优解,否则返回步骤(2),继续执行,直到满足结束条件.

4 仿真实验及分析

4.1 测试数据集及参数设置

在仿真实验中采用 UCI 数据库中的经典数据集 Iris、Wine、Breast 作为三大测试数数据集,验证 KM、PSO-KM、ADPSO-IKM 及 ISMPSO-AKM 算法的性能.测试数据集描述如表 1 所示.

表 1 测试数据集属性

测试数据集	数据总量	聚类个数	数据维度
Iris	150	3	4
Wine	178	3	13
Breast	106	6	10

在实验中,取粒子种群数 $m=30$,最大迭代次数 $T_{\max}=100$,对于 PSO 算法,取 $c_1=c_2=2.0$, $\omega=0.8$;对于 ADPSO 算法,取 $c_{\max}=2.5$, $c_{\min}=0.5$, $\omega_{\max}=0.85$, $\omega_{\min}=0.4$;对于 ISMPSO-AKM 算法,取 $\omega_{\max}=0.9$, $\omega_{\min}=0.4$, $c_1=c_2=c_3=1.49$.

4.2 仿真实验结果

4.2.1 动态获取 K 值及初始中心实验

为了验证 AKM 算法是否可以获取准确的 K 值,使用标准数据集 Iris 作为测试数据,并且确定聚类个数 K 的范围在 2~15(一般不超过 10)之间,绘制出不同聚类数对应的 SE,以便更直观的选择 K 值.如图 1 所示.

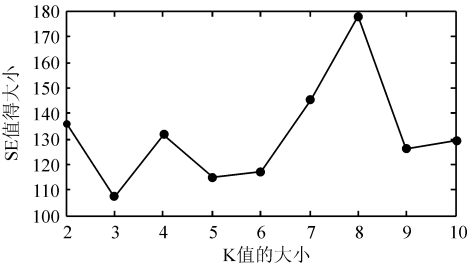


图 1 Iris 数据集上不同聚类数对应 SE

通过观察图 1, Iris 数据集不同分类数对应的 SE 折线图,可以看出,在 $K=3$ 时,SE 显著下降,且此时 SE 具有最小值.而 Iris 数据集实际分类个数也是 3,与计算得到的 K 值结果一致.

同样采用 Iris 数据来验证初始中心选取是否有效,计算 AKM 算法所得到的初始化结果与 IRIS 数据实际聚类中心位置的误差平方和,结果见表 2.

表 2 Iris 数据聚类中心误差对比

Iris 数据集 原始中心	AKM 算法 初始化结果	误差 平方和
5.00,3.42, 1.46,0.24	0.063 1,5.017,3.335, 1.468,0.239	0.063 1
5.93,2.77, 4.26,1.32	5.889,2.751, 4.112,1.406	
6.58,2.97, 5.55,2.02	6.676,3.006, 5.448,1.964	

由表 2 可以看出,由 AKM 算法选取出的聚类中心近似接近 Iris 数据集原始中心,且计算得到的标准误差平方和为 0.063 1,数值较小,说明 AKM 算法可以获得有效的初始聚类中心点.

4.2.2 聚类算法收敛性实验

对表 1 中的三大数据集进行仿真实验,分别测试四种算法的收敛性与稳定性.记录各算法收敛到固定值时适应度值的最大值 $f_{\max}$ 、最小值 $f_{\min}$ 及平均值 f_{ave} .所有的值都取多次实验的平均值作为实验结果,四种算法比较见表 3.

表 3 算法适应度值

数据集	$f(x)$	KM	PSO-KM	ADPSO-IKM	ISMPSO-AKM
Iris	$f_{\min}$	8.301 9	10.306 9	10.388 9	10.476 2
	$f_{\max}$	10.401 0	10.463 2	10.501 2	10.592 7
	f_{ave}	9.691 2	10.398 4	10.473 8	10.535 6
	$f_{\min}$	5.331 4	5.901 7	5.937 6	6.018 9
Wine	$f_{\max}$	5.902 0	6.076 2	6.024 0	6.124 7
	f_{ave}	5.631 2	5.931 7	6.011 9	6.089 4
	$f_{\min}$	6.472 8	6.942 1	7.067 4	7.198 2
	$f_{\max}$	6.715 1	7.015 3	7.119 5	7.235 9
Breast	f_{ave}	6.658 3	6.981 0	7.086 9	7.203 1

由表 3 可以看出,分析三个数据集的聚类结果,相比 KM 和 PSO-KM 算法,ADPSO-IKM 与 ISMPSO-AKM 算法适应度值的最大值、最小值及平均值均有所增加,表明这两种算法提高了聚类结果的准确性.从适应度值的平均值来看,ISMPSO-AKM 算法适应值的平均值都大于 ADPSO-IKM 算法的平均值,表示改进的算法收敛速度更快,收敛精度更高,得到的聚类效果最好.从适应度值最大值与最小值差值来看,与前两种算法差值较大、聚类结果不稳定相比,本文改进的算法差值最小,说明 ADPSO-IKM 算法得到的聚类效果更为稳定.同时,为了方便对比 KM、PSO-KM、ADPSO-IKM 及 ISMPSO-AKM 算法之间收敛性的差异,图 2~4 给

出了四种算法在三个数据集上的收敛性曲线图.

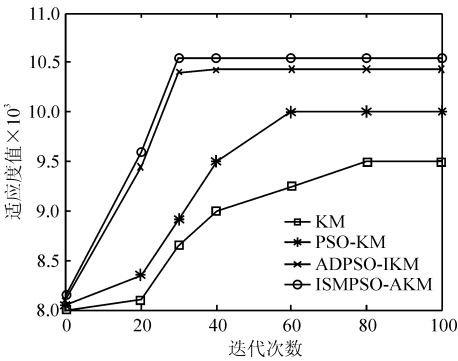


图 2 四种算法在 Iris 上的收敛性

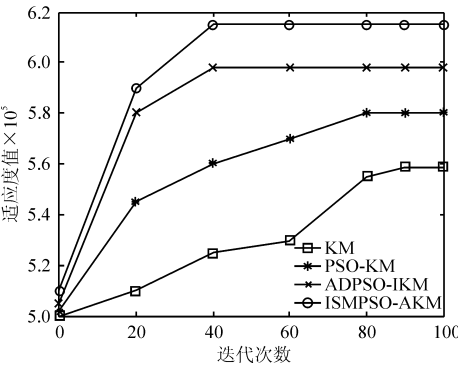


图 3 四种算法在 Wine 上的收敛性

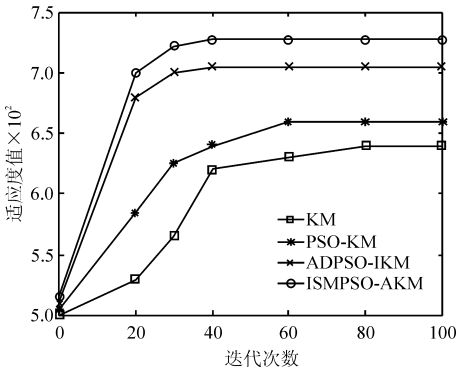


图 4 四种算法在 Breast 上的收敛性

可以看出,在三大数据集上,ISMPSO 算法的引入,不仅扩大了有效搜索区域,同时保证了种群的多样性,增强了算法期全局搜索的性能,在迭代开始时,改进算法的初始适应值高于其他三个算法.此外,引入 AKM 聚类算法,大幅度提高了算法收敛速度.在数据集 Iris 上,KM 很难找到最优解;PSO-KM 经过 60 次左右的迭代收敛到最终聚类效果;ADPSO-IKM 算法迭代到 30 次左右时,开始趋于稳定收敛;而 ADPSO-IKM 算法在迭代不到 30 次时已经得到了较为稳定的聚类结果.在数据集 Wine

上,当面临高维数据时,前两种算法运算速度较慢,而 ISMPSO-AKM 与 ADPSO-IKM 算法的收敛趋势大致相同,但在 20~30 次这一迭代阶段,本文提出的算法收敛性能明显提升,说明 ADPSO-IKM 算法可以更快找到全局最优解.在数据集 Breast 上,PSO-KM 受 KM 算法影响较大,呈平稳缓慢的收敛趋势,ADPSO-IKM 算法倾向于在 35 次迭代时稳定收敛,相比之下,ISMPSO-AKM 算法进一步改进聚类效果,聚类性能更好.

5 结束语

本文提出了一种改进的简化均值粒子群 K-means 优化聚类算法 (ISMPSO-AKM).将原先简化粒子群的位置项扩展为由个体最优位置、全局最优位置及邻域最优位置线性组合而来的四项.此外,构造了一种基于余弦函数和对数函数的惯性权重.引入 AKM 聚类算法克服了 KM 算法受聚类数及初始聚类中心的选取影响较大的缺点,使聚类结果更准确.对 UCI 数据库中的经典数据集 Iris、Wine、Breast 进行仿真实验,结果表明,与其他改进算法相比,本文提出的 ISMPSO-AKM 算法具体更好的性能.但 AKM 算法的引入,增加了算法整体的运算时间,下一步将对这方面进行研究改进.

参考文献:

[1] 王金永. 基于粒子群优化算法的聚类分析研究[D]. 青岛:青岛理工大学,2015.

[2] 王宏杰,师彦文. 结合初始中心优化和特征加权的 K-Means 聚类算法[J]. 计算机科学,2017,44(11A):467-469.

[3] 李武,赵娇燕,严太山. 基于平均差异度优选初始聚类

中心的改进 K-均值聚类算法[J]. 控制与决策,2017,32(4):759-762.

[4] 陈朝威,常冬霞. 基于密度差分的自动聚类算法[J]. 软件学报,2018,29(4):43-52.

[5] KENNEDY J, EBERHART R. Particle swarm optimization[C]//Proceedings of ICNN'95-International Conference on Neural Networks. Perth, WA, Australia, IEEE, 2002.

[6] 李斌,王劲松,黄玮. 一种大数据环境下的新聚类算法[J]. 计算机科学,2015,42(12):247-250.

[7] 许明杰,蔚承建,沈航. 基于 Spark 的并行 K-means 算法研究[J]. 微电子学与计算机,2018,35(5):101-105.

[8] 陈小玉,李晓静,马海英. 一种面向大数据的快速自动聚类算法[J]. 计算机应用研究,2017,34(9):2651-2654.

[9] PERAM T, VEERAMACHANENI K, MOHAN C K. Fitness-distance-ratio based particle swarm optimization[C]// Swarm Intelligence Symposium. IEEE, Saint Louis,USA,2012.

[10] 胡旺,李志蜀. 一种更简化而高效的粒子群优化算法[J]. 软件学报,2007,18(4):861-868.

[11] DEEP K, BANSAL J C. Mean particle swarm optimisation for function optimisation[J]. International Journal of Computational Intelligence Studies, 2009, 1(1):72.

作者简介:

靳雁霞 女,(1973-),博士,副教授.研究方向为智能优化算法、虚拟现实技术.

齐欣(通讯作者) 女,(1994-),硕士研究生.研究方向为智能优化算法. E-mail:448463832@qq.com.

张晋瑞 男,(1994-),硕士研究生.研究方向为虚拟仿真.

程琦甫 男,(1994-),硕士研究生.研究方向为虚拟仿真.

(上接第 68 页)

[7] ZHI L, JIN P, XUAN S, et al. CCF-LRU: a new buffer replacement algorithm for flash memory[J]. IEEE Transactions on Consumer Electronics, 2009, 55(3):1351-1359.

[8] 何简繁. 基于 NAND 闪存的固态硬盘缓存优化策略研究[D]. 杭州:杭州电子科技大学,2018

[9] JOHNSON T, SHASHA D, PROPOSED G W. Q: A Low Overhead High Performance Buffer Management Replacement Algorithm[C]// International Conference on Very Large Data Bases. Morgan Kaufmann

Publishers Inc. Athens, Greece, 1994.

[10] YAN H, YAO Q. An efficient file-aware garbage collection algorithm for NAND flash-based consumer electronics[J]. IEEE Transactions on Consumer Electronics, 2014, 60(4):623-627.

作者简介:

李中单 女,(1995-),硕士研究生.研究方向为模式识别与智能系统. E-mail:2412592897@qq.com.

严华 男,(1971-),博士,教授.研究方向为模式识别与智能系统.