

一种改进鱼群聚类算法在结构面分组中的应用

王述红¹, 任艺鹏¹, 陈俊智², 张紫杉¹

(1. 东北大学 资源与土木工程学院, 辽宁 沈阳 110819; 2. 昆明理工大学 国土资源学院, 云南 昆明 650093)

摘 要: 针对结构面产状常规分类方法存在的不足, 提出一种新型的结构面分类算法. 基于 K-Means 算法的结构面分类, 将人工鱼群算法(artificial fish swarm algorithm, AFSA)与 K-Means 算法相结合, 建立了 AFSA-RSK 结构面分类算法. 利用鱼群算法强大的寻优能力, 代替 K-Means 算法对结构面产状聚心集进行搜寻, 并通过 K-Means 算法进行聚类. 聚类完成后, 选择相应参数指标对聚类效果进行评价. 针对存在的问题, 对鱼群算法的步长和视野进行修正, 提高寻找聚心集的精度, 动态地调整了聚类过程. 将改进后的 AFSA-RSK 算法与其他算法进行比较, 结果表明在迭代速度、聚类精度以及内存占比上, 改进后的 AFSA-RSK 算法都要更优, 更适合在结构面分组方面的应用.

关 键 词: 人工鱼群算法; 岩体结构面; 岩体; 聚类; 边坡

中图分类号: TU 45 **文献标志码:** A **文章编号:** 1005-3026(2019)03-0420-05

An Improved Fish Swarm Clustering Algorithm for Structural Grouping

WANG Shu-hong¹, REN Yi-peng¹, CHEN Jun-zhi², ZHANG Zi-shan¹

(1. School of Resources & Civil Engineering, Northeastern University, Shenyang 110819, China; 2. Faculty of Land and Resources and Engineering, Kunming University of Science and Technology, Kunming 650093, China.

Corresponding author: WANG Shu-hong, E-mail: wangshuhong@mail.neu.edu.cn)

Abstract: Aiming at the shortcomings of the conventional classification method of structural plane production, a new structural plane classification algorithm was proposed. Based on the structural plane classification of K-Means algorithm, the AFSA-RSK structural surface classification algorithm is established by combining the artificial fish swarm algorithm (AFSA) with the K-Means algorithm. The powerful optimization ability of the fish swarm algorithm is used to replace the K-Means algorithm to search for the structural surface set, and clustering by K-Means algorithm. After the clustering is completed, the corresponding parameters are selected to evaluate the clustering effect. According to the existing problems, the step size and visual field of the fish swarm algorithm are modified, the accuracy of finding the cluster is improved, and the clustering process is dynamically adjusted. Comparing the improved AFSA-RSK algorithm with other algorithms, it can be obtained that the improved AFSA-RSK algorithm is better in iterative speed, clustering precision and memory ratio, and it is more suitable for application in structural plane grouping.

Key words: artificial fish swarm algorithm (AFSA); rock joint plane; rock mass; clustering; slope

岩体是包含结构面和岩石的组合, 结构面的稳定一定程度上决定了岩体的失稳与否. 因此对结构面进行统计和分组, 得到其分布规律, 对于研

究大型边坡、地下洞室、隧道的稳定具有很大的意义.

由于表征结构面参数具有较大的随机性和变

异性,因此,大多采用概率统计生成和分组的方式.王述红等^[1]采用双正态分布、Fisher 分布等分布方式,研究了随机结构面对块体搜索结果的影响.宋腾蛟等^[2]考虑几何参数和力学参数,利用人工蜂群算法的优势对结构面产状进行了分类. Shanley 等^[3-4]运用聚类算法对结构面数据进行分析. Hammah 等^[5]首次将模糊 C 均值聚类算法应用于结构面的分组.张奇等^[6]基于层次聚类分析理论,提出凝聚层次聚类结构面分组方法.黄润秋等^[7]利用粒子群算法较强的寻优能力,优化了 K-Means 方法,并结合 K-Means 方法对结构面进行分组.

本文将人工鱼群算法 (AFSA) 与 K-Means 算法相结合,通过修正鱼群算法的步长和视野,建立了改进的 AFSA-RSK 结构面分类算法.将改进后的 AFSA-RSK 算法与其他算法进行比较,可以得到,在迭代速度、聚类精度以及内存占比上,改进后的 AFSA-RSK 算法都要更优,更适合在结构面分组方面的应用.

1 结构面鱼群分类算法 (AFSA-RSK)

1.1 算法简介

鱼群算法是李晓磊博士提出的一种智能寻优算法^[8-9],通过构建人工鱼群并模拟鱼群寻找食物的过程来进行聚类.人工鱼群主要有三个行为:追尾、聚群、觅食.三种行为的最终目的是寻找到最优的个体.整个算法运行全过程控制参数主要有结构面样本集 n 、人工鱼移动步长 Step、视野范围 V 、迭代次数 t .

1.2 鱼群适应度函数

本文采用结构面坐标到结构面聚心距离平方和的倒数作为适应度函数.适应度计算公式为

$$\text{Fit} = \frac{1}{\sum_{i=1}^k \sum_{j=1}^n \|d(c_i - x_j)\|^2}. \quad (1)$$

式中: k 为聚心个数; n 为结构面样本总数; $\|d(c_i - x_j)\|^2$ 为分组结构面中样本 x 离聚心 c 的距离.距离计算公式^[7]为

$$d(c_i - x_j) = \sqrt{1 - (n_c \cdot n_j^T)^2}. \quad (2)$$

式中, n_c, n_j 分别为聚心结构面的单位法向量和 j 结构面的单位法向量.

1.3 移动策略

觅食行为 在假定初始聚心的情况下,人工鱼在结构面样本集视野内随机搜寻,搜寻公式为

$$X_j = X_i + X_i \cdot \text{Step} \cdot \text{Rand}(). \quad (3)$$

式中:Step 为移动步长;Rand 函数输出 (0 ~ 1) 中随机一个数.按照式 (3) 随机移动后,到达新的位置 X_j 处,可以得到 X_j 的坐标值,按照 1.2 节式 (1),式 (2) 计算新位置适应度,若 X_j 处适应度更大,则向前一步,得到前进公式为

$$X_i^{t+1} = X_i^t + \frac{X_j - X_i^t}{\|X_j - X_i^t\|} \cdot \text{Step} \cdot \text{Rand}(). \quad (4)$$

式中, X_i^{t+1} 为多次执行觅食行为后新聚心的位置.

聚群行为 人工鱼 X_i 搜索视野内的伙伴数量,搜索完毕后,计算当前视野内鱼群数目的均值坐标,即为中心坐标值.若中心位置适应度较大,输出中心坐标值,同时按式 (5) 进行移动,得到新的聚心坐标,并与中心坐标值进行比较,较优的记录下来,中心位置计算方法如式 (6) 所示:

$$X_i^{t+1} = X_i^t + \frac{X_{\text{cen}} - X_i^t}{\|X_{\text{cen}} - X_i^t\|} \cdot \text{Step} \cdot \text{Rand}(); \quad (5)$$

$$X_{\text{cen}} = \sum_{i=1}^g X_j / g. \quad (6)$$

式中: X_j 为视野内某结构面的坐标; g 为伙伴数目; X_{cen} 为鱼群经过聚群得到的结构面新聚心.

追尾行为 人工鱼 X_i 搜索当前视野中适应度最大的伙伴 X_{max} ,寻找到之后,在视野范围 V 内,人工鱼进行移动,移动方式见式 (7),按照式 (7) 可以得到新的聚类中心 X_{next} 的坐标值,将 X_{next} 和 X_{max} 中的较优者记录下来.

$$X_{\text{next}}^{t+1} = X_i^t + \frac{X_{\text{max}} - X_i^t}{\|X_{\text{max}} - X_i^t\|} \cdot \text{Step} \cdot \text{Rand}(). \quad (7)$$

式中, X_{max} 为视野中最大适应度的个体.

1.4 算法流程

1) 初始化鱼群行为参数,并随机假定 k 个聚心,根据结构面样本集 (倾向、倾角) 的极大值与极小值范围,由聚心集生成的初始鱼群按照式 (3),式 (4),式 (6),式 (7) 分别执行觅食、聚群、追尾三种行为.每执行完一种行为,将该行为中最优结构面坐标进行记录,三种行为执行完毕,比较三种行为中较优的坐标,选择最优的结构面聚心坐标.

2) 按照新得到的结构面聚心坐标执行 K-Means 算法.

3) 采用 Xie-Beni 指标 S_{XB} 、分类熵指标 H_m 两项指标检验聚类效果,计算单一 K-Means 算法,并将结果记录,与使用 AFSA-RSK 迭代后的聚类结果进行对比,若 AFSA-RSK 两项指标误差均小于 K-Means 方法,将结果记录,并继续迭代,超过最大迭代次数输出聚类结果,若不满足,

执行步骤 4)。

4) 对步长和视野进行修正,步长修正公式、改进的觅食行为策略和视野修正公式如式(8)~式(10)所示。

步长修正公式:

Step = [Step - ((Step_max - Step_mean) * t) / T_max] . (8)

Step_max 选为 0.16,则相应的改进觅食行为策略为

X_i^{t+1} = X_i^t + (X_j - X_i^t) / |X_j - X_i^t| * [Step_max - ((Step_max - Step_mean) * t) / T_max] . (9)

迭代次数对步长的影响如图 1 所示。

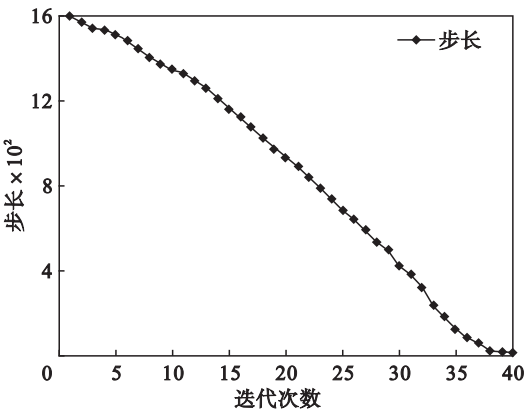


图 1 步长变化图
Fig. 1 Step change chart

视野修正公式为

V(t) = ((V_max - V) / 2) + (V * t) / T_max . (10)

式中:V_max 选为 1;T_max 为最大迭代次数。

对鱼群视野进行修正,将鱼群视野范围与迭代次数建立函数关系,随着迭代次数的增加,鱼群每次行为之前视野的选取都与前一次迭代的视野相关,建立视野随迭代次数动态的变化过程的函数,有利于寻优精度的提升和运算速度的加快。

隶属度计算公式如式(11)所示:

U_{ik} = 1 / sum_{i=1}^k ((d(x_j, c_i) / d(x_i, c_i))^{2(m-1)}) . (11)

式中:m 为权值分配系数,一般取 1 到 2,本文取 1.5;d(x_j, c_i) 是第 j 个结构面样本到 i 个聚类中心的距离;c_i 为聚心坐标。

5) 修正视野与步长后,改进后的 AFSA - RSK 算法每执行一步与前步结果进行比较,并将两者间误差较小的值记录下来,达到最大迭代次

数后,输出记录的最优结果,并与 K - Means 算法计算结果比较.若优于,则输出结果;否则,再次修改步长,重新迭代,超过最大迭代次数后输出结果。

作者根据前述原理以及设计流程,基于 Matlab 开发平台编写了该分析程序。

1.5 聚类有效性评价

本文采取 Xie - Beni 指标 S_XB、分类熵指标 H_m [10] 两项指标进行聚类效果评价,计算式为

S_{XB} = (sum_{i=1}^k sum_{j=1}^n U_{ik}^2 d^2(X_j, c_i)) / (n * (min_{i != k} d^2(c_i, c_k))) , (12)

H_m(U, k) = - sum_{k=1}^n sum_{i=1}^k U_{ik} log_a(U_{ik}) . (13)

式中:k 是聚心个数;n 为样本总数;U_{ik} 为计算所得隶属度;d^2(X_j, c_i) 为某结构面样本 X_j 到聚类中心 c_i 距离的平方;min_{i != k} d^2(c_i, c_k) 为两个距离最近的聚类中心距离的平方.分类指标 H_m 和指标 S_XB 越小,代表聚类效果越良好;反之,聚类效果越差.利用以上两组评价指标,对聚类结果进行评价。

2 工程实例分析

2.1 工程简介

本文采用云南省新平县鲁奎山铁矿岩质高边坡作为算例.该矿山边坡主要有碎石土、砾岩、泥灰岩等构成。

2.2 结构面优势分组

为验证改进 AFSA - RSK 算法在结构面聚类分组的有效性,选取鲁奎山铁矿某断面 127 组结构面,采用 Matlab 和 DIPS 软件经过赤平投影分析可以得到密度云图,如图 2 所示。

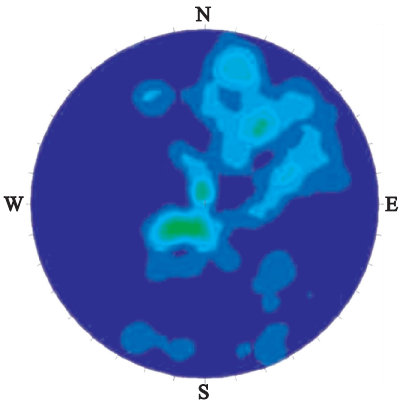


图 2 结构面密度云图
Fig. 2 Structural surface density cloud image

根据上文所述方法,要对聚类中心数目 k 进

行假定,根据结构面产状玫瑰花图,并结合密度等值线图观察可以得到,聚类中心数目 k 为 2~4. 例如,当选取 $k=2$ 时,随机生成 2 条人工鱼对应随机的结构面聚心,分别采用前述步骤,执行相关行为,每进行一步,对适应度进行计算,若超过最

大迭代次数,终止算法,并输出结果. 限于文章篇幅,选取 $k=3$ 时改进 AFSA - RSK 算法与 K - Means 两种算法的聚类结果,如图 3 所示.

利用 S_{XB} , H_m 两种指标对 $k=2, k=3, k=4$ 时的聚类结果进行评判,评价结果如表 1 所示.

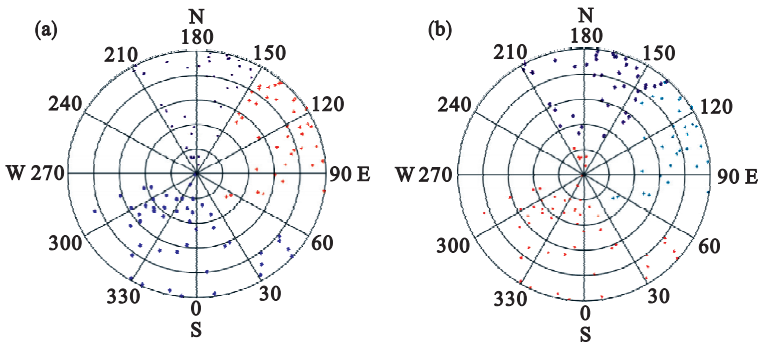


图 3 AFSA - RSK 和 K - Means 聚类结果 ($k=3$)
Fig. 3 The clustering results of AFSA-RSK and K-Means ($k=3$)
(a)—AFSA - RSK; (b)—K - Means.

通过表 1 可以看出,经过修正后的 AFSA - RSK 算法的两项指标的评价值优于 K - Means 算法,且当聚类中心为 3 时,无论是改进 AFSA - RSK 还是 K - Means 算法,两项指标均小于聚类数为 2 和聚类数为 4 时的情况. 这也说明从聚类效果来看,聚类效果 3 是较为合理的,两项指标均更优.

通过上述分析可以得到,基于改进鱼群算法优化过后的 K - Means 算法,适用于对结构面产状的聚类分析,改进 AFSA - RSK 算法动态的优化机制,使得鱼群在搜索结构面聚心的过程中,不断搜索到适应度高的个体,即结构面样本离聚心距离更小的个体,最终达到优化聚类过程的目的.

由前文可以得到,当 $k=3$ 时,AFSA - RSK

聚类结果较优. 选取 $k=3$ 时,将 AFSA - RSK 结果与文献[7]、文献[11]计算结果进行比较,三种算法都是通过优化迭代对聚心进行求解,聚类结果见图 4. 采用参数 S_{XB} 对计算结果进行评价,参数变化图见图 5.

表 1 聚类有效性评价
Table 1 Cluster validity evaluation

方法	聚类数	S_{XB}	H_m
K - Means	$k=2$	0. 245 3	0. 312 3
	$k=3$	0. 199 7	0. 239 0
	$k=4$	0. 277 4	0. 261 9
AFSA - RSK	$k=2$	0. 220 4	0. 298 5
	$k=3$	0. 173 2	0. 217 4
	$k=4$	0. 263 9	0. 243 8

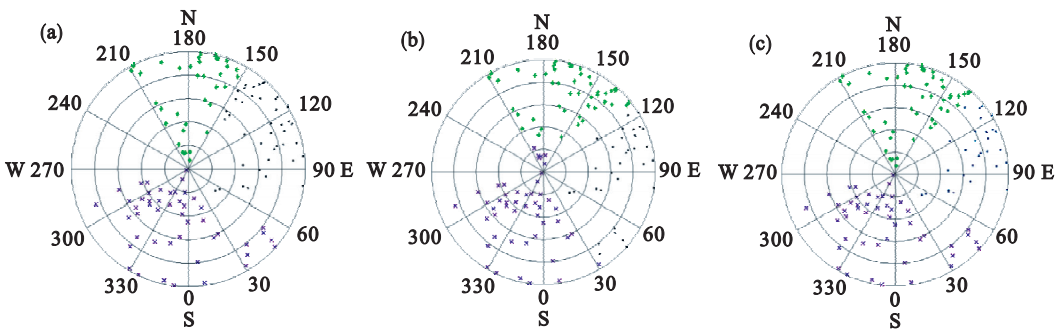


图 4 聚类结果对比
Fig. 4 Clustering results comparison
(a)—AFSA - RSK; (b)—PSO; (c)—Firefly Algorithm.

由图 5 可看出,鱼群算法通过修改迭代步长和视野,迭代精度有所提高,且随着迭代次数的增

加,4 种算法的搜索精度都在提高,视野与步长经过改进后的 AFSA - RSK 算法,算法误差更小.

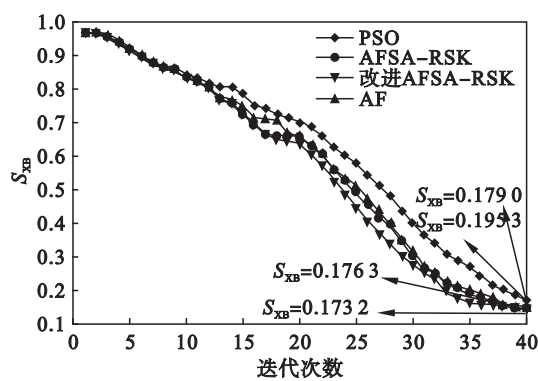


图 5 误差变化
Fig. 5 Error change graph

3 结 论

- 1) 改进 AFSA – RSK 算法动态的优化机制,使得鱼群在搜索结构面聚心的过程中,可以不断搜索到适应度高的个体,最终达到优化聚类过程的目的。
- 2) 将鱼群算法中步长和视野与迭代次数建立联系,使得人工鱼在搜索过程中动态地调整步长和视野,优化鱼群搜索聚心的过程,完成聚类。
- 3) 与其他算法进行比较,在迭代精度、迭代速度和内存占比上,经改进后的 AFSA – RSK 算法都要优于其他算法,证明了本文提出算法的高效性和合理性。

参考文献:

[1] Wang S H, Ni P P, Yang H, et al. Modeling on spatial block topological identification and the improgressive failure analysis of slope and cavern rock mass[J]. *Procedia Engineering*, 2011, 10(7): 1509 – 1514.

[2] 宋腾蛟,陈剑平,张文,等. 基于人工蜂群算法的岩体结构面多参数优势分组研究[J]. *岩土力学*, 2015, 36(3): 861 – 868.

(Song Teng-jiao, Chen Jian-ping, Zhang Wen, et al. A method for multivariate parameter dominant partitioning of discontinuities of rock mass based on artificial bee colony algorithm[J]. *Rock and Soil Mechanics*, 2015, 36(3): 861 – 868.)

[3] Shanley R J, Mahtab M A. Delineation and analysis of clusters in orientation data[J]. *Journal of the International Association for Mathematical Geology*, 1976, 8(1): 9 – 23.

[4] Mahtab M A, Yegualp T M. A rejection criterion for definition of clusters in orientation data[C]// *Proceedings of the 22nd Symposium on Rock Mechanics*. New York: American Institute of Mining Metallurgy and Petroleum Engineers, 1982: 116 – 123.

[5] Hammah R E, Curran J H. Fuzzy cluster algorithm for the automatic identification of joint sets[J]. *International Journal of Rock Mechanics and Mining Sciences*, 1998, 35(7): 889 – 890.

[6] 张 奇,王 清,阙金声,等. 基于凝聚层次聚类分析法的岩体随机结构面产状优势分组[J]. *岩土工程学报*, 2014, 36(8): 1432 – 1437.

(Zhang Qi, Wang Qing, Que Jin-sheng, et al. Dominant partitioning of discontinuities of rock masses based on AGNES[J]. *Chinese Journal of Geotechnical Engineering*, 2014, 36(8): 1432 – 1437.)

[7] 宋金龙,黄润秋,裴向军. 基于粒子群算法的岩体结构面产状模糊 C 均值聚类分析[J]. *工程地质学报*, 2012, 20(4): 591 – 598.

(Song Jin-long, Huang Run-qiu, Pei Xiang-jun. Particle swarm optimization algorithm based fuzzy C-means cluster analysis for discontinuities occurrence in rock mass[J]. *Journal of Engineering Geology*, 2012, 20(4): 591 – 598.)

[8] 李晓磊,邵之江,钱积新. 一种基于动物自治体的寻优模式: 鱼群算法[J]. *系统工程理论与实践*, 2002, 22(11): 32 – 38.

(Li Xiao-lei, Shao Zhi-jiang, Qian Ji-xin. An optimizing method based on autonomous animats: fish-swarm algorithm[J]. *Systems Engineering—Theory & Practice*, 2002, 22(11): 32 – 38.)

[9] 李晓磊. 一种新型的智能优化算法——人工鱼群算法[D]. 杭州: 浙江大学, 2003.

(Li Xiao-lei. A new intelligent optimization method—artificial fish school algorithm[D]. Hangzhou: Zhejiang University, 2003.)

[10] Bezdek J C. Pattern recognition with fuzzy objective function algorithms[M]. New York: Plenum Press, 1981: 203 – 239.

[11] 宋腾蛟,陈剑平,张文,等. 基于萤火虫算法的岩体结构面优势产状聚类分析[J]. *东北大学学报(自然科学版)*, 2015, 36(2): 284 – 287.

(Song Teng-jiao, Chen Jian-ping, Zhang Wen, et al. Clustering analysis of dominative attitudes of rock mass structural plane based on firefly algorithm[J]. *Journal of Northeastern University (Natural Science)*, 2015, 36(2): 284 – 287.)