

基于 t 分布变异的改进麻雀搜索算法*

吴超略, 韦文山, 郭 羿, 邬贵昌

(广西民族大学 电子信息学院, 广西 南宁 530006)

摘 要: 针对麻雀搜索算法收敛速度慢、难以跳出局部最优等问题, 提出一种基于 t 分布变异的改进麻雀搜索算法。在更新麻雀种群加入者位置后, 引入自适应 t 分布变异, 对加入者位置产生扰动, 避免陷入局部最优, 增强算法性能; 通过比较灰狼优化算法、飞蛾火焰优化算法和原始麻雀搜索算法, 在 6 个基准函数上进行仿真实验。实验结果和 Wilcoxon 符号秩检验结果都表明所提出的改进麻雀搜索算法的收敛精度与速度均优于其他算法, 达到提高算法收敛速度, 增强算法跳出局部极值能力的效果。

关键词: 麻雀搜索算法; 智能优化算法; Sin 映射; t 分布

中图分类号: TP18

文献标识码: A

DOI: 10.20044/j.csdg.2097-1788.2022.02.012

引用格式: 吴超略, 韦文山, 郭羿, 等. 基于 t 分布变异的改进麻雀搜索算法[J]. 网络安全与数据治理, 2022, 41(2): 74-78.

Improved sparrow search algorithm based on t -distribution variation

Wu Chaolue, Wei Wenshan, Guo Yi, Wu Guichang

(College of Electronic Information, Guangxi University for Nationalities, Nanning 530006, China)

Abstract: Aiming at the problems of slow convergence speed and difficult to jump out of local optimization of sparrow search algorithm, an improved sparrow search algorithm based on t -distribution mutation is proposed. After updating the position of the follower, the adaptive t -distribution mutation is introduced to disturb the position of the follower, avoid falling into local optimization and enhance the performance of the algorithm. By comparing gray wolf optimizer, moth-flame optimization and original sparrow search algorithm, simulation experiments are carried out on six benchmark functions. The experimental results and Wilcoxon's signed rank test results both show that the convergence accuracy and speed of the proposed improved sparrow search algorithm are better than other algorithms, which can improve the convergence speed of the algorithm and enhance the ability of the algorithm to jump out of local extremum.

Key words: sparrow search algorithm; intelligent optimization algorithm; Sin map; t -distribution

0 引言

群智能优化算法是一种仿生算法, 旨在模拟自然界中某些生物的行为或某些物理现象, 因其寻优能力强、操作简单等特点, 被许多科研人员研究。常见的群智能优化算法有: 粒子群优化算法(Particle Swarm Optimization, PSO)^[1]、萤火虫算法(Firefly Algorithm, FA)^[2]、灰狼优化算法(Grey Wolf Optimizer, GWO)^[3]、乌鸦搜索算法(Crow Search Algorithm, CSA)^[4]和飞蛾火焰优化算法(Moth-Flame Optimization, MFO)^[5]。受自然界中麻雀种群觅食行为的启发, 薛建凯^[6]等

人于 2020 年提出了麻雀搜索算法(Sparrow Search Algorithm, SSA)。相比之下, 该算法拥有更优的收敛率、更高的精度和易于实现等特点。但是, 在算法运行的末期, SSA 算法也不能避免收敛速度下降、易陷入局部最优的问题。

为了改善麻雀搜索算法跳出局部最优难的问题, 加强算法运行效率, 许多学者提出了有效的改进策略: 吕鑫等^[7]通过 Tent 混沌序列初始化麻雀种群, 同时引入高斯变异和 Tent 混沌扰动对个体进行变异和扰动, 使算法不易陷入局部最优点; 柳长安等^[8]融合反向学习策略和自适应 t 分布变异, 引入精英粒子, 扩大了算法搜索范围, 增强算法后期

* 基金项目: 广西民族大学研究生教育创新计划项目(gxun-chxps202083)

局部搜索能力;付华等^[9]在加入者位置更新时加入鸡群算法的随机跟随策略,保证多样性的同时又提高了搜索性能;Zhang等^[10]利用 Logistic 混沌映射对种群位置进行初始化提高初始解的质量,为了加快 SSA 算法的收敛速度和效率,采用两个自适应参数更新发现者位置和预警麻雀数量。

以上改进措施虽然能在一定程度上提高算法跳出局部最优的能力,然而,问题仍然存在,如搜索精度不够、收敛速度不快等。基于此,本文提出一种基于 t 分布变异的改进麻雀搜索算法(t -SSA),该算法在加入者位置更新后,引入自适应 t 分布变异,对加入者位置进行扰动变异,避免陷入局部最优,提高算法的寻优精度。通过在 6 个基准函数上进行仿真实验,结果表明,与 SSA 算法相比, t -SSA 算法具有更好的优化精度和更快的收敛速度。

1 麻雀搜索算法的基本原理

麻雀搜索算法(SSA)受自然界中麻雀捕食行为和反捕食行为影响,将麻雀种群抽象为三类:发现者、加入者和预警者。其中,发现者具有较高的适应度值,负责为麻雀种群寻找食物区域,并把相关区域和方向提供给加入者;为了得到更好的食物,提高自身适应度值,加入者会一直跟随具有较高适应度值的发现者,不断监视发现者进而争夺食物;当预警者发现捕食者后会立即向种群中的麻雀发出警报,此时处于种群边沿的个体会快速地朝安全区域飞去,即反捕食行为。此外,在 SSA 算法中,麻雀种群的发现者和加入者的身份并不是一成不变的,假如能寻觅到更优的食物区域,每只麻雀都能是发现者,不过二者在种群总数量中的比重是固定的。

在 SSA 算法中,麻雀种群的表示形式为:

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,d} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,d} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,d} \end{bmatrix} \quad (1)$$

其中, n 为麻雀的数量; d 为待优化变量的维度。可用下述形式来代表麻雀种群的适应度值:

$$F_x = \begin{bmatrix} f(X_1) \\ f(X_2) \\ \vdots \\ f(X_n) \end{bmatrix} \quad (2)$$

其中, $f(X_n)$ 为麻雀个体的适应度值。

对于发现者,用下式表示其位置变化:

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \cdot \exp\left(-\frac{i}{\alpha \cdot \text{iter}_{\max}}\right), & R_2 < \text{ST} \\ X_{i,j}^t + Q \cdot L, & R_2 \geq \text{ST} \end{cases} \quad (3)$$

式中, $X_{i,j}^t$ 代表第 i 个发现者在第 j 维中的位置; t 代表当前算法迭代次数; $\text{iter}_{\max}$ 是一个表示算法最大迭代次数的常数; α ($\alpha \in (0, 1]$) 表示一个随机数; Q 是一个随机数,服从正态分布; L 表示一个 $1 \times d$ 的矩阵,其中每个元素的值均为 1; R_2 ($R_2 \in [0, 1]$) 为预警值,ST ($\text{ST} \in [0.5, 1]$) 为安全值。当 $R_2 < \text{ST}$ 时,表示周围没有捕食者,发现者可以进行广泛搜索;如果 $R_2 \geq \text{ST}$,则表示一些麻雀已经发现了天敌,所有麻雀都需要尽快飞到其他安全区域。

对于加入者,其位置更新公式如下:

$$X_{i,j}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{X_{\text{worst}}^t - X_{i,j}^t}{t^2}\right), & i > \frac{n}{2} \\ X_p^t + |X_{i,j}^t - X_p^t| \cdot A^+ \cdot L, & i \leq \frac{n}{2} \end{cases} \quad (4)$$

式中, i 代表第 i 个加入者; X_{worst}^t 为此时全局最劣位置; X_p^t 表示发现者当前能搜索到的最佳位置; Q 、 L 同式(3); A 为内部元素随机赋值为 -1 或 1 的 $1 \times d$ 矩阵,并且 $A^+ = (AA^T)^{-1}$ 。当 $i > n/2$ 时,表示第 i 个加入者具有较差的适应度值,需要离开当前位置;当 $i \leq n/2$ 时,表示第 i 个加入者在发现者所处的最佳位置附近觅食。

对于预警者,数量通常占麻雀种群数量的 10%~20%,且位置随机确定,其位置更新公式如下:

$$X_{i,j}^{t+1} = \begin{cases} X_{\text{best}}^t + \beta \cdot |X_{i,j}^t - X_{\text{best}}^t|, & f_i > f_g \\ X_{i,j}^t + K \cdot \left(\frac{|X_{i,j}^t - X_{\text{worst}}^t|}{(f_i - f_w) + \varepsilon} \right), & f_i = f_g \end{cases} \quad (5)$$

式中, X_{best}^t 表示当前全局最优位置; β 为步长参数,是随机数,且服从标准正态分布; K ($K \in [-1, 1]$) 是一个表示麻雀移动方向的随机数; ε 表示一个值很小的常数,为了避免分母为 0; f_i 表示当前预警者的适应度值, f_g 表示目前全局最优的适应度值,而 f_w 则为最劣的。当 $f_i > f_g$ 时,说明这时的预警者正在种群的边际,被捕食者捕食的几率大幅增加,应向着 X_{best}^t 全局最优位置移动;当 $f_i = f_g$ 时,表示种群中的麻雀觉察到捕食者的存在,为避免捕食者可能的攻击,必须向别的麻雀飞去。

2 改进的麻雀搜索算法(t-SSA)

2.1 t分布变异

t分布又称学生t分布(Student's t-Distribution),它的概率密度函数曲线与其自由度 n 有关, n 的值越大,其曲线中间越高。当 $n=1$ 时,有 $t(n=1) \rightarrow C(0,1)$;当 $n=\infty$ 时,有 $t(n=\infty) \rightarrow N(0,1)$ 。其中, $C(0,1)$ 为柯西分布, $N(0,1)$ 为高斯分布。即柯西分布和高斯分布为t分布自由度 n 分别为1和 ∞ 时的两个特例。t分布的函数分布如图1所示。

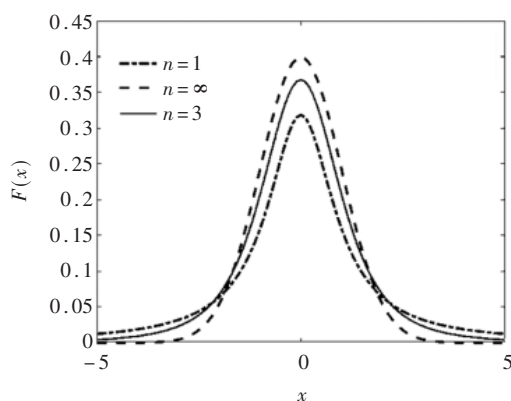


图1 t分布函数分布图

t分布既有柯西分布的特点,也有高斯分布的特点,本文将t分布的自由度参数用算法迭代次数代替,使t分布在算法迭代初期趋近于柯西分布,此时算法的全局寻优能力得到提升;随着迭代次数的增加,t分布又趋近于高斯分布,提高了算法的局部搜索水平,进而提高算法的寻优精度。

本文利用当前迭代次数中发现者所能搜索到的最优位置,结合t分布变异对发现者的位置进行扰动,其公式如下:

$$X_{\text{new}} = X_p^{\text{iter}} + X_p^{\text{iter}} \cdot t(\text{iter}) \quad (6)$$

$$X_{i,j}^{t+1} = \begin{cases} X_{\text{new}}, & f_{\text{new}} < f_i \\ X_{i,j}^{t+1}, & \text{其他} \end{cases} \quad (7)$$

其中, X_{new} 表示经t分布变异扰动后的新位置; X_p^{iter} 表示当前发现者所搜索到的最优位置; $t(\text{iter})$ 表示以当前迭代次数 iter 为自由度的t分布; $X_{i,j}^{t+1}$ 表示当前由式(4)计算得到的第 i 个加入者位置; f_{new} 表示新位置的适应度值, f_i 表示第 i 个加入者的适应度值。当 $f_{\text{new}} < f_i$ 时,表示经过t分布变异扰动之后得到的新位置优于当前加入者位置,并把该位置更新给加入者。

2.2 改进的麻雀搜索算法流程

本文提出的基于t分布变异的改进麻雀搜索算法(t-SSA)步骤如下:

- (1)进行初始化,并设置参数,如种群数量 n 、发现者数量PD、预警者数量SD、最大迭代次数 iter_{max} 、安全值ST等。
- (2)计算麻雀个体适应度值,记录当前全局最优值和最劣值,及其对应位置。
- (3)根据设置好的发现者数量PD,挑选适应度值较好的麻雀数量当作发现者,并按照式(3)更新发现者位置。
- (4)剩余的麻雀为加入者,按照式(4)更新加入者位置。
- (5)应用式(6)计算出t分布变异扰动后的新位置,并按照式(7)对加入者位置进行相应的替换。
- (6)根据设置好的预警者数量SD,随机选取部分麻雀作为预警者,并按照式(5)更新预警者位置。
- (7)更新麻雀种群个体最优位置和全局最优位置。
- (8)判断当前迭代次数是否满足算法最大迭代次数,若是,则循环结束,输出结果;否则返回步骤(2)。

3 仿真实验与结果分析

为验证本文所提出的t-SSA算法的可行性和寻优能力,本文基于Intel® Core™ i5-7200U CPU@2.50 GHz,8 GB内存,Windows 10操作系统和仿真软件MATLAB R2019b进行仿真实验。

3.1 对比实验设计和参数设置

本文将t-SSA算法与3种基本算法:灰狼优化算法(GWO)^[31]、飞蛾火焰优化算法(MFO)^[15]和麻雀搜索算法(SSA)^[6],分别测试6个基准函数(其中 $f_1 \sim f_3$ 为单峰函数, $f_4 \sim f_6$ 为多峰函数),进行对比。为确保实验的公平性,将全部算法统一设置种群数量100,迭代次数200,其余算法参数见表1。6个基准测试函数具体信息见表2。选取各算法在基准测试函数独立运行30次的结果的平均值和标准差作为实验结果,具体实验结果如表3所示,基准测试函数收敛情况如图2所示。

3.2 实验结果分析

从表3和图2能看出,在测试单峰函数 $f_1 \sim f_3$ 时,t-SSA算法均取得了其理论最优值,且具有更快收敛速度和更强的鲁棒性。在测试多峰函数 $f_4 \sim f_6$ 时,对于 f_4 ,t-SSA虽然收敛精度稍差于MFO,但与其余算法相比,其在前期的收敛速度较快;对于 f_5 ,t-SSA

表 1 算法参数设置

算法	参数设置
GWO	$a_{\max}=2 ; a_{\min}=0 ; r_1, r_2 \in[0,1]$
MFO	$b=1$
SSA	PD=20%,SD=10%,ST=0.8
t-SSA	PD=20%,SD=10%,ST=0.8

表 2 基准函数信息

函数	函数名称	维度	范围	最优值
f_1	Sphere Model	30	$[-100,100]$	0
f_2	Schwefel's problem 2.22	30	$[-10,10]$	0
f_3	Schwefel's problem 2.21	30	$[-100,100]$	0
f_4	Generalized Schwefel's problem 2.26	30	$[-500,500]$	-12 569.5
f_5	Ackley's Function	30	$[-32,32]$	0
f_6	Generalized Penalized Function	30	$[-50,50]$	0

和 SSA 收敛结果相同,并且比 GWO、MFO 优,但从图 2 可以看出,t-SSA 收敛速度明显更快;对于 f_6 ,t-SSA 具有较好的寻优精度和较强的鲁棒性。

综合来看,本文所提出的 t-SSA 算法较其他算法在收敛精度、收敛速度和鲁棒性方面都有很大的提升,这表明本文所提出的改进策略是有效的,既提高了算法的寻优精度,又降低了算法陷入局部最

表 3 算法优化结果对比

函数	名称	GWO	MFO	SSA	t-SSA
f_1	平均值	3.26e-15	2 115.91	0	0
	标准差	3.46e-15	4 261.14	0	0
f_2	平均值	1.49e-09	28.05	6.07e-124	0
	标准差	7.63e-10	17.29	3.32e-123	0
f_3	平均值	4.18e-03	52.95	1.27e-158	0
	标准差	3.04e-03	8.05	6.88e-158	0
f_4	平均值	-6 320.42	-8 578.95	-7 778.13	-8 032.92
	标准差	1 615.89	863.08	486.14	1 146.03
f_5	平均值	1.28e-08	11.83	8.88e-16	8.88e-16
	标准差	6.81e-09	5.17	0	0
f_6	平均值	1.18e-02	18 046.12	2.44e-08	9.80e-09
	标准差	7.42e-03	34 674.91	4.76e-08	2.75e-08

优的概率,且拥有更好的性能。

3.3 Wilcoxon 秩和检验

为进一步评估 t-SSA 算法的优化性能,仍需要对算法进行统计检验。因此,本文选择 Wilcoxon 秩和检验验证 t-SSA 算法每轮的优化结果是否与所对比算法存在明显区别。用 P 表示检验结果,当 $P<5\%$ 时,说明两种算法之间存在显著差异;当 $P>5\%$ 时,说明两种算法之间差距不明显,性能相当。在本文

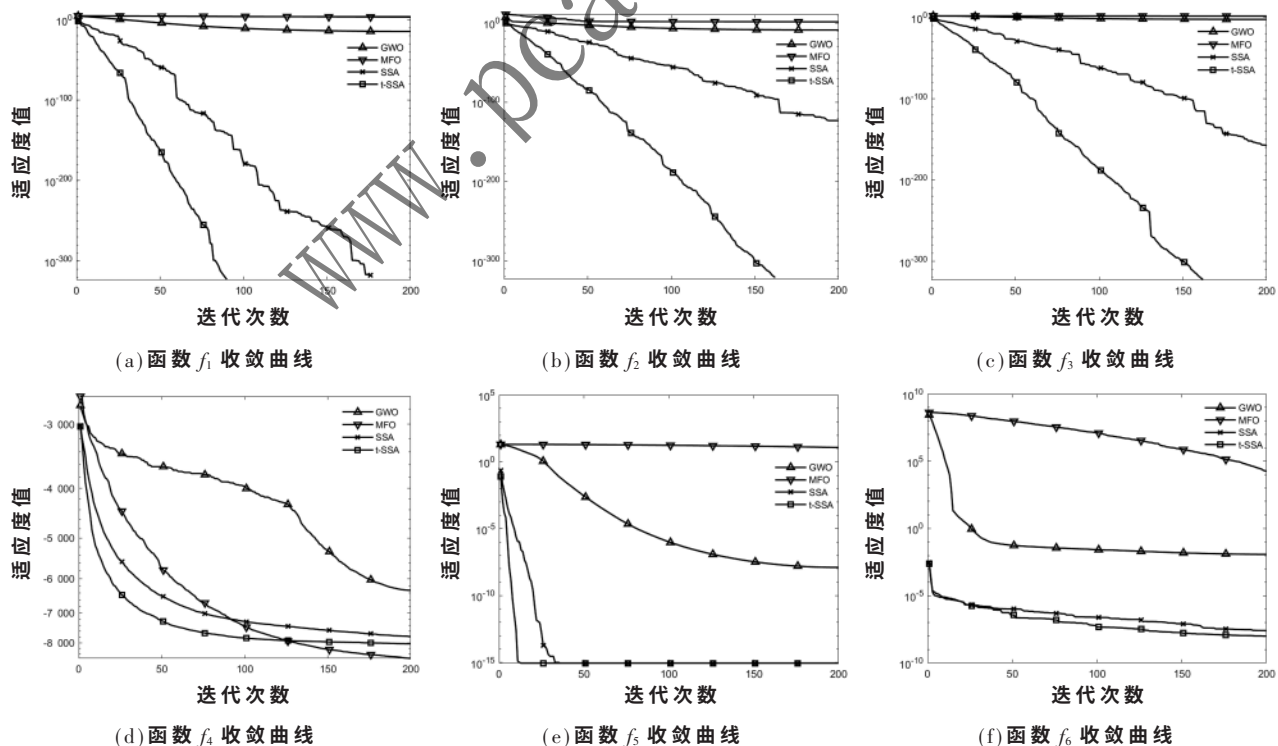


图 2 基准函数收敛曲线

3.1 节同样的实验条件下, t -SSA 分别与 GWO、MFO、SSA 的详细检验结果 P 如表 4 所示, 其中 NaN 代表两个对比算法性能接近, 无法比较。

表 4 Wilcoxon 秩和检验结果

函数	GWO	MFO	SSA
f_1	1.21e-12	1.21e-12	4.19e-02
f_2	1.21e-12	1.21e-12	3.13e-04
f_3	1.21e-12	1.21e-12	3.13e-04
f_4	5.87e-04	8.35e-08	8.65e-01
f_5	1.21e-12	1.21e-12	NaN
f_6	3.02e-11	3.02e-11	3.95e-01

由表 4 可知, 除 t -SSA 与 SSA 的 f_4, f_6 , 其余大部分的 P 值都小于 5%, 这说明 t -SSA 算法与其他算法在统计上存在显著差异。对于 f_5 , t -SSA 与 SSA 差异不明显, 寻优效果相当。

4 结论

为了改善麻雀搜索算法(SSA)在其运行末期出现种群多样性下降、难以跳出局部最优等问题, 本文提出一种新的改进麻雀搜索算法(t -SSA)。首先, 引入自适应 t 分布变异, 对加入者位置进行扰动变异, 防止跳不出局部最优点, 增强算法性能。然后, 通过在 6 个基准函数上进行仿真实验, 结果证明本文提出的 t -SSA 的收敛精度与速度优于其他算法。综合来看, 本文所提出的改进方法是有效的, 能够明显提升 SSA 算法的寻优性能, 下一步将考虑将 t -SSA 算法应用于实际工程问题的优化中, 验证其在工程问题上的应用价值。

参考文献

[1] KENNEDY J, EBERHART R. Particle swarm optimization[C]//Proceeding of International Conference on Neural Networks. Perth, Australia: IEEE, 1995: 1942-1948.

[2] YANG X S. Nature-inspired metaheuristic algorithm[M]. Beckington, UK: Luniver Press, 2008.

[3] MIRJALILI S, MIRJALILI S M, LEWIS A. Grey wolf optimizer[J]. Advances in Engineering Software, 2014, 69(3): 46-61.

[4] ASKARZADEH A. A novel metaheuristic method for solving constrained engineering optimization problems: crow search algorithm[J]. Computers and Structures, 2016, 169: 1-12.

[5] MIRJALILI S. Moth-flame optimization algorithm: a novel nature-inspired heuristic paradigm[J]. Knowledge-Based Systems, 2015, 89: 228-249.

[6] XUE J K, SHEN B. A novel swarm intelligence optimization approach: sparrow search algorithm[J]. Systems Science & Control Engineering, 2020, 8(1): 22-34.

[7] 吕鑫, 慕晓冬, 张钧, 等. 混沌麻雀搜索优化算法[J]. 北京航空航天大学学报, 2021, 47(8): 1712-1720.

[8] 柳长安, 冯雪雯, 孙长浩, 等. 基于改进麻雀算法的最大二维熵分割方法[J]. 激光技术, 2022, 46(2): 274-282.

[9] 付华, 刘昊. 多策略融合的改进麻雀搜索算法及其应用[J]. 控制与决策, 2022, 37(1): 87-96.

[10] ZHANG C L, DING S F. A stochastic configuration network based on chaotic sparrow search algorithm[J]. Knowledge-Based Systems, 2021, 220: 106924.

(收稿日期: 2022-04-04)

作者简介:

吴超略(1998-), 男, 硕士研究生, 主要研究方向: 群智能算法优化、人工智能。

韦文山(1965-), 通信作者, 男, 教授, 主要研究方向: 计算机应用。E-mail: wws1996@163.com。

郭羿(1996-), 男, 硕士研究生, 主要研究方向: 机器学习、智能算法。



版权声明

凡《网络安全与数据治理》录用的文章，如作者没有关于汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权等版权的特殊声明，即视作该文章署名作者同意将该文章的汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权授予本刊，本刊有权授权本刊合作数据库、合作媒体等合作伙伴使用。同时，本刊支付的稿酬已包含上述使用的费用，特此声明。

《网络安全与数据治理》编辑部

www.pcachina.com