

基于马尔科夫链的灰狼优化算法收敛性研究

张孟健¹, 龙道银², 王 霄¹, 杨 靖¹

(1. 贵州大学电气工程学院, 贵州贵阳 550025; 2. 中国电建集团贵州工程有限公司, 贵州贵阳 550001)

摘 要: 针对灰狼优化算法(Grey Wolf Optimization, GWO)在收敛性研究上的不足, 首先, 通过定义灰狼种群状态转移序列, 建立了 GWO 算法的马尔科夫(Markov)链模型, 通过分析 Markov 链的性质, 证明它是有限齐次 Markov 链; 其次, 通过分析灰狼种群状态序列最终转移状态, 结合随机搜索算法的收敛准则, 验证了 GWO 算法的全局收敛性; 最后, 对典型测试函数、偏移函数及旋转函数进行仿真实验, 并与多种群体智能算法进行对比分析. 实验结果表明, GWO 算法具有全局收敛性强、计算耗时短和寻优精度高等优势.

关键词: 灰狼优化; 灰狼种群状态空间; 马尔科夫链; 状态转移概率; 收敛准则

中图分类号: TP301.6 **文献标识码:** A **文章编号:** 0372-2112(2020)08-1587-09

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2020.08.018

Research on Convergence of Grey Wolf Optimization Algorithm Based on Markov Chain

ZHANG Meng-jian¹, LONG Dao-yin², WANG Xiao¹, YANG Jing¹

(1. Electrical Engineering College, Guizhou University, Guiyang, Guizhou 550025, China;

2. Power China Guizhou Engineering Company Limited, Guiyang, Guizhou 550001, China)

Abstract: The global convergence is one of the important features of an intelligent algorithm. In this paper, we take the initiative to handle the convergence of Grey Wolf Optimization(GWO) with Markov chain. Firstly, a Markov chain model of the GWO algorithm is established through defining the state transition sequence of a gray wolf population. Analyzing the properties of the Markov chain proves that it is homogeneous finite. Secondly, based on the convergence criteria of random search algorithms, the global convergence of the GWO algorithm is verified via analyzing the final state transition sequence of the grey wolf population. Finally, simulation studies on typical testing functions, shifting functions and rotating functions are carried out comparing with a few typical swarm intelligent algorithms. The experimental results show that the GWO algorithm has excellent performance on the global convergence, the computational time and precision of optimization.

Key words: Grey Wolf Optimization; grey wolf population state space; Markov chain; state transition probability; convergence criteria

1 引言

灰狼优化(Grey Wolf Optimization, GWO)^[1]算法是澳大利亚学者 Mirjalili 等在 2014 年提出的一种群智能优化算法. GWO 算法来源于对灰狼群的捕食行为和狩猎的等级制度的研究, 以灰狼的等级制度为特征, 以求解优化问题为背景, 通过灰狼群体搜索、包围和攻击等过程达到优化求解的目的. 研究表明, GWO 算法与蚁群算法(Ant Colony Optimization, ACO)^[2]、遗传算法(Ge-

netic Algorithm, GA)^[3]、粒子群优化(Particle Swarm Optimization, PSO)算法^[4]等相比有自身的优点.

马尔科夫(Markov)链是对随机过程进行分析的重要手段, 由算法生成的解集本质上构成了 Markov 链, 因此 Markov 链为分析自然启发性算法提供一个通用的框架, 在工程技术等领域得到广泛的应用^[5]. 目前 Markov 链已成功用于蚁群算法(ACO)^[6]、混合蛙跳算法(SFLA)^[7]、粒子群优化算法(PSO)^[8,9]、元胞遗传算法(CGA)^[10]、人工蜂群算法(ABC)^[11,12]、鸡群算法

收稿日期:2019-11-13; 修回日期:2020-2-18; 责任编辑:马兰英

基金项目:国家自然科学基金(No. 61861007, No. 61640014); 贵州省科技支撑计划(No. [2019]2152, No. [2017]2520-1); 贵州省研究生创新基金(No. YJSCXJH[2019]005); 贵州省物联网理论与应用案例库(No. KCALK201708); 贵州省学科建设(No. ZDXK[2015]8)

(CSO)^[13]等群智能算法的收敛性分析. Suzuki^[14]运用 Markov 链对遗传算法进行稳定性分析. Dorea 等人^[15]对 GA 建立了多级 Markov 链的数学模型,并证明之收敛. He 等^[16]运用离散时间 Markov 链方法,证明花授粉算法在一定条件下能保证全局收敛.

针对 GWO 算法的改进,胡小平等^[17]提出一种基于指数变化的控制参数的调整策略. 魏政磊等^[18]提出基于余弦函数和二次函数的收敛因子更新方法. 段亚青等^[19]提出一种基于对数递减策略的收敛因子控制策略. 腾志军等^[20]引入 Tent 映射来初始化种群,可提高全局搜索能力. 但是文献[17~20]是通过实验来说明 GWO 算法的收敛性,未在理论上进行证明. 薛俊杰等^[21]提出一种改进狼群智能算法(Smart Wolf Pack Algorithm, SWPA),对其收敛性进行理论分析并与其他智能算法进行了对比研究. Long W 等^[22]提出一种基于折射射学习的新型灰狼优化算法,从极限收敛的角度证明了改进 GWO 算法的收敛性. 不难发现,尽管 GWO 算法的改进与应用已有不少研究成果,但是对 GWO 算法的理论分析较少,特别是关于算法收敛性的分析不多见,在一定程度上制约了该算法的改进和应用.

本文利用 Markov 链对 GWO 算法的收敛性进行分析,通过建立 GWO 算法的 Markov 链模型,研究灰狼群状态的转移行为,结合随机优化算法的收敛准则分析 GWO 算法的收敛性,并通过对测试函数的仿真实验进一步验证 GWO 算法的收敛性.

2 灰狼优化算法基本模型

GWO 算法模拟了狼群的社会等级制度和群体的狩猎行为^[1],将其最优解设为 α 狼;次优解设为 β 狼;第三优解设为 δ 狼;剩余的解归为 ω 狼. 灰狼群狩猎过程包括接近和包围猎物,数学表达式^[1]如下:

$$D = |C \cdot X_p(t) - X(t)| \quad (1)$$

$$X(t+1) = X_p(t) - A \cdot D \quad (2)$$

其中, D 为个体与目标之间的距离, t 为当前的迭代次数, C 和 A 为系数向量, X_p 为目标的位置向量, $X(t)$ 为单只灰狼的位置向量. A 和 C 的计算公式如下:

$$A = 2a \cdot r_1 - a \quad (3)$$

$$C = 2r_2 \quad (4)$$

其中, r_1, r_2 为 $[0, 1]$ 之间的随机数, a 为收敛因子:

$$a = 2 - 2t_i/t_{\max} \quad (5)$$

其中, t_{\max} 表示最大的迭代次数, $i = 1, 2, \dots, N$.

灰狼个体跟踪猎物位置的公式如下:

$$\begin{cases} D_\alpha = |C_1 \cdot X_\alpha - X(t)| \\ D_\beta = |C_2 \cdot X_\beta - X(t)| \\ D_\delta = |C_3 \cdot X_\delta - X(t)| \end{cases} \quad (6)$$

其中, $D_\alpha, D_\beta, D_\delta$ 分别表示 α, β 和 δ 与其他个体之间的

距离, $X_\alpha, X_\beta, X_\delta$ 分别表示 α, β 和 δ 当前的位置, C_1, C_2, C_3 为系数.

在猎捕阶段中,剩余的 ω 狼个体向着 α, β 和 δ 狼前进的步长和方向:

$$\begin{cases} X_1 = X_\alpha - A_1 \cdot D_\alpha \\ X_2 = X_\beta - A_2 \cdot D_\beta \\ X_3 = X_\delta - A_3 \cdot D_\delta \end{cases} \quad (7)$$

$$X(t+1) = (X_1 + X_2 + X_3)/3 \quad (8)$$

3 GWO 算法的 Markov 模型

定义 1 (灰狼状态和灰狼状态空间) 灰狼状态由灰狼捕猎过程中不同等级灰狼的位置构成,记为 X ,其中 $X \in Y$, Y 表示可行解空间. 则灰狼所有可能状态组成的集合构成灰狼的状态空间,记为 $\mathbf{X} = \{X | X \in Y\}$.

定义 2 (灰狼群状态和灰狼群状态空间) 灰狼群中所有灰狼的状态构成灰狼群状态,记为 $\varphi = (X_1, X_2, \dots, X_i)$, ($i = 1, 2, \dots, N_\varphi$), 其中 X_i 为第 i 灰狼的状态. 灰狼群所有可能状态组成的集合构成灰狼群的状态空间,记为 $\Phi = \{\varphi = (X_1, X_2, \dots, X_i) | X_i \in \mathbf{X}, (i = 1, 2, \dots, N_\varphi)\}$.

定义 3 (状态等价) 对于 $\varphi \in \Phi, X \in \varphi$, 记为:

$$\psi(\varphi, X) = \sum_{i=1}^{N_\varphi} \chi_{|X|}(X_i) \quad (9)$$

其中, $\chi_{|X|}$ 表示事件 X 的示性函数, $\psi(\varphi, X)$ 表示灰狼群状态 φ 中包含灰狼状态的数量. 如存在两个灰狼群 $\varphi_1, \varphi_2 \in \Phi$, 对于任意 $X \in \mathbf{X}$, 若有 $\psi(\varphi_1, X) = \psi(\varphi_2, X)$, 则称 φ_1 和 φ_2 等价, 记作 $\varphi_1 \sim \varphi_2$.

定义 4 (状态等价类) 由状态等价在 Φ 上可类比出灰狼群状态等价类, 记作 $Le = \Phi / \sim$, 简称灰狼群等价类. 其存在如下性质:

性质 1 某等价类 Le 内任意灰狼群之间都是等价的, 即 $\varphi_i \sim \varphi_j, \forall \varphi_i, \varphi_j \in Le$.

性质 2 Le 内任意灰狼群状态与 Le 外的任意灰狼群状态不等价, 即 φ_i 与 φ_j 不等价, $\forall \varphi_i \in Le, \varphi_j \notin Le$.

性质 3 任意两个不同等价类没有交集, 即 $Le_1 \cap Le_2 = \emptyset, \forall Le_1 \neq Le_2$.

定义 5 (灰狼状态转移) 对于任意 $X_i \in \varphi, \forall X_j \in \varphi$, GWO 算法迭代过程中, 灰狼状态由 X_i 一步转移到 X_j , 记为 $T_\varphi(X_i) = X_j$.

定理 1 GWO 算法中, 灰狼状态由 X_i 一步转移到 X_j 的转移概率 $P(T_\varphi(X_i) = X_j)$ 的表达式为:

$$P(T_\varphi(X_i) = X_j) = \begin{cases} P_\alpha(T_\varphi(X_i) = X_j), & \text{由 } \alpha \text{ 狼实现;} \\ P_\beta(T_\varphi(X_i) = X_j), & \text{由 } \beta \text{ 狼实现;} \\ P_\delta(T_\varphi(X_i) = X_j), & \text{由 } \delta \text{ 狼实现;} \\ P_\omega(T_\varphi(X_i) = X_j), & \text{由 } \omega \text{ 狼实现;} \end{cases} \quad (10)$$

证明 将灰狼群体看为超空间的一组点集,则灰狼猎物的更新过程即在超空间中进行点集之间的交换.根据定义 4 和 GWO 算法的几何性质,可以得到 α 狼由状态 X_i 一步转移到 X_j 的转移概率.

$$P_{\alpha}(T_{\varphi}(X_i) = X_j) = \begin{cases} \frac{1}{|-A_i \cdot D_{\alpha}|}, & X_j \in [X_{\alpha i}, X_{\alpha i} - A_i \cdot D_{\alpha}] \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

β 狼由状态 X_i 一步转移到 X_j 的转移概率.

$$P_{\beta}(T_{\varphi}(X_i) = X_j) = \begin{cases} \frac{1}{|-A_i \cdot D_{\beta}|}, & X_j \in [X_{\beta i}, X_{\beta i} - A_i \cdot D_{\beta}] \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

δ 狼由状态 X_i 一步转移到 X_j 的转移概率.

$$P_{\delta}(T_{\varphi}(X_i) = X_j) = \begin{cases} \frac{1}{|-A_i \cdot D_{\delta}|}, & X_j \in [X_{\delta i}, X_{\delta i} - A_i \cdot D_{\delta}] \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

由于 GWO 算法是通过不同等级的灰狼之间的交流转换及协作来完成,所以灰狼状态由 X_i 一步转移到 X_j 的转移概率是式(11)~式(13)共同决定的.证毕.

定义 6 (灰狼群状态转移概率) 对于 $\forall \varphi_i \in \Phi$, $\forall \varphi_j \in \Phi$, GWO 算法迭代过程中,灰狼群状态由 φ_i 转移到 φ_j , 记作 $T_{\varphi}(\varphi_i) = \varphi_j$. 灰狼群状态由 φ_i 一步转移到 φ_j 的转移概率为:

$$P(T_{\varphi}(\varphi_i) = \varphi_j) = \prod_{m=1}^{N_{\varphi}} P(T_{\varphi}(X_{im}) = X_{jm}) \quad (14)$$

即灰狼群状态由 φ_i 一步转移到 φ_j 的概率为灰狼群 φ_i 内所有灰狼的状态同时转移成 φ_j 内所有灰狼的状态.

定理 2 在 GWO 算法灰狼群的内部迭代中,灰狼群状态由 φ_i 一步转移到 φ_j 的转移概率为:

$$P(T_{\varphi}(\varphi_i) = \varphi_j) = \begin{cases} P_{m,k} \cdot P(T_{\varphi}(X_{ik}) = X_{jk}) \cdot P(y_i \rightarrow y_j) \\ 0, & \text{otherwise} \end{cases} \quad (15)$$

其中, k 表示灰狼群迭代更新前后位置唯一发生改变的灰狼的位置序号; $P_{m,k}$ 表示选择灰狼群第 k 只灰狼进行位置更新的概率,且 $0 < P_{m,k} < 1$; m 表示灰狼群的数量, $m \leq N_{\varphi}$, $P(T_{\varphi}(X_{ik}) = X_{jk})$ 表示 φ_i 和 φ_j 灰狼群状态中各自第 k 个灰狼状态之间的转移概率.

此外, $P(X_i \rightarrow X_j)$ 表示 φ_i 和 φ_j 灰狼群状态中各自最佳灰狼状态之间的转移概率,其表达式如下:

$$P(y_i \rightarrow y_j) = \begin{cases} 1, & f(y_i) \geq f(y_j) \\ 0, & f(y_i) < f(y_j) \end{cases} \quad (16)$$

证明 灰狼群状态 $X_i = (X_{i1}, X_{i2}, \dots, y_i)_{1 \times m}$ 一步转移至状态 $X_j = (X_{j1}, X_{j2}, \dots, X_{jk}, y_j)_{1 \times m}$, 根据 GWO 算法更新机制,种群内部每一次迭代更新前后只有一头灰

狼位置发生变化,且选择灰狼群中适应度最差的单个个体进行更新.因此,在灰狼群状态内部迭代时,只有一头灰狼状态能完成转移,其余多个灰狼状态发生转移的概率为 0.

假设待转移的灰狼在种群中的位置排序为 k , 被选中的概率记为 $P_{m,k}$. 算法在内部寻优计算时,通过计算每一个灰狼个体的适应度值并进行排序.当 $k \geq m$ 时,且 $0 < P_{m,k} < 1$, 此时的灰狼群状态内部的转移可表示为:

$$P(T_{\varphi}(\varphi_i) = \varphi_j) = P_{m,k} \cdot P(T_{\varphi}(X_{ik}) = X_{jk}) \cdot P(y_i \rightarrow y_j) \quad (17)$$

从上式可知,由于当前灰狼群 m 个灰狼中的适应度值最优灰狼(α 狼)状态不发生转移,则当前的灰狼群最优状态会保留或与更佳的状态替换,即:

$$P(y_i \rightarrow y_j) = \begin{cases} 1, & f(y_i) \geq f(y_j) \\ 0, & f(y_i) < f(y_j) \end{cases}$$

定义 7 (Markov 链) 设随机过程 $\{x_n, n \in T\}$, 参数集 T 为离散的时间序列,即 $T = \{0, 1, 2, \dots\}$, 对应有取值 x_n 的全体组成离散的状态空间 $I = \{i_0, i_1, i_2, \dots\}$. 若对任意的整数 $n \in T$ 和 $\forall i_0, i_1, i_2, \dots, i_{n+1} \in I$ 满足以下条件概率:

$$\begin{aligned} p\{x_{n+1} = i_{n+1} | x_0 = i_0, x_1 = i_1, \dots, x_n = i_n\} \\ = p\{x_{n+1} = i_{n+1} | x_n = i_n\} \end{aligned} \quad (18)$$

则称 $\{x_n, n \in T\}$ 为 Markov 链.

定理 3 GWO 算法中灰狼群状态序列 $\{\varphi(t) : t > 0\}$ 是有限齐次 Markov 链.

证明 任何优化算法的搜索空间都是有限的,故任一灰狼状态中 x_i 都是有限的,因此灰狼的状态空间 X 是有限的.此外,由于一个灰狼群状态空间 $\varphi = (X_1, X_2, \dots, X_i)$, ($i = 1, 2, N_{\varphi}$) 是由 N_{φ} 个灰狼组成, N_{φ} 是有限正整数,故灰狼群状态空间 Φ 亦是有限的.

Markov 性 由定义 6 可知,灰狼群状态序列 $\{\varphi(t) : t > 0\}$ 中,对于任意 $\varphi(t-1) \in \Phi$, $\varphi(t) \in \Phi$, 其转移概率 $P(T_{\varphi}(\varphi(t-1)) = \varphi(t))$ 由灰狼群内所有的灰狼的转移概率 $P(T_{\varphi}(X(t-1)) = X(t))$ 决定.由定理 1 可知,灰狼群内任一灰狼的状态转移概率 $P(T_{\varphi}(X(t-1)) = X(t))$ 仅与 $t-1$ 的状态 $X(t-1)$ 、系数向量 C_i 、个体与目标之间的距离 $D_{\alpha}, D_{\beta}, D_{\delta}$ 相关,所以转移概率 $P(T_{\varphi}(\varphi(t-1)) = \varphi(t))$ 也仅与 $t-1$ 时刻的状态相关.由定义 7 可知,灰狼群状态序列 $\{\varphi(t) : t > 0\}$ 具有 Markov 性.

有限性 因为状态空间为可列集,灰狼群状态空间 Φ 为有限的,所以灰狼群状态序列 $\{\varphi(t) : t > 0\}$ 构成一个有限的 Markov 链.

齐次性 由定理 1 可知, $P(T_{\varphi}(X(t-1)) = X(t))$ 仅与 $t-1$ 时刻的状态 $X(t-1)$ 有关,而与 $t-1$ 时刻无关,因此,状态序列 $\{\varphi(t) : t > 0\}$ 是有限齐次 Markov 链.

4 GWO 算法的收敛性分析

GWO 算法的 Markov 链可表示为:

$$\begin{cases} W_\alpha(t) = (x_\alpha(t-1), X_\alpha(t)) \\ W_\beta(t) = (x_\beta(t-1), X_\beta(t)) \\ W_\delta(t) = (x_\delta(t-1), X_\delta(t)) \end{cases} \quad (19)$$

其中, t 表示算法迭代次数, $x_\alpha(t-1)$ 表示第 $t-1$ 代 α 狼所处的位置, $X_\alpha(t)$ 表示除 α 狼外其他等级灰狼的位置, $x_\beta(t-1)$ 表示第 $t-1$ 代 β 狼所处的位置, $X_\beta(t)$ 表示除 α 狼、 β 狼外其他等级灰狼的位置, $x_\delta(t-1)$ 表示第 $t-1$ 代 δ 狼所处的位置, $X_\delta(t)$ 表示除 α 狼、 β 狼及 δ 狼外其他等级灰狼的位置. 有 $W(t) = (x(t-1), X(t))$.

定义 8 记灰狼优化算法的 Markov 链状态空间为 Φ , 则其状态空间的维数为 $L = |\Phi| = |Dim|^d$.

GWO 算法的本质属于随机搜索算法的范畴, 本文用随机优化算法的收敛性判定准则^[23]来证明灰狼优化算法的收敛行为.

4.1 收敛准则

对于优化问题 $\langle Y, f \rangle$, 有随机优化算法 Z , 第 k 次迭代的结果为 x_k , 则下一次迭代的结果为 $x_{k+1} = D(x_k, \zeta)$; 其中, Y 为可行解空间, f 为适应度函数, ζ 为算法 Z 迭代中曾经搜索过的解. 在 Lebesgue 测度空间, 搜索的下界被定义为:

$$\sigma = \inf \{ t | v(x \in Y | f(x) < t) > 0 \} \quad (20)$$

其中, $v(X)$ 为集合 X 上的 Lebesgue 测度, 则最优解可定义为:

$$R_{\xi, M} = \begin{cases} \{ x \in Y | f(x) < \sigma + \xi \}, & \sigma \text{ 有限} \\ \{ x \in Y | f(x) < -c \}, & \sigma = -\infty \end{cases} \quad (21)$$

其中, ξ 大于零, c 为充分大的正数. 如果算法可以找到 $R_{\xi, M}$ 中的一个点, 则可认为算法找到了可接受的全局最优或近似全局最优点.

Condition 1 $f(Z(x, \zeta)) \leq f(x)$, 且若 $\zeta \in Y$, $f(Z(x, \zeta)) \leq f(\zeta)$.

Condition 2 对 $\forall B \in Y$, s. t. $v(B) > 0$, 有:

$$\prod_{k=0}^{\infty} (1 - v_k(B)) = 0 \quad (22)$$

其中, $v_k(B)$ 为算法 Z 第 k 次迭代搜索解在集合 B 上的概率测度.

定理 4^[23] (全局收敛的充要条件) 设 f 可测, 可测空间 Y 为 R^n 上可测度的子集, 算法 Z 满足 Condition 1 和 Condition 2, $\{x_k\}_{k=0}^{\infty}$ 是算法 Z 产生的解序列, 则有概率测度:

$$\lim_{k \rightarrow \infty} P(x_k \in R_{\xi, M}) = 1 \quad (23)$$

即算法 Z 全局收敛. 式中, $P(x_k \in R_{\xi, M})$ 为算法第 k 步迭代搜索的解 x_k 在 $R_{\xi, M}$ 中的概率测度.

4.2 收敛性证明

引理 1 GWO 算法满足 Condition 1, 即灰狼优化算法的种群狩猎方向是单调的.

定义 9 (灰狼最优状态集 G) 设优化问题 $\langle Y, f \rangle$ 的最优解是 g^* , 定义灰狼群的最优状态集为:

$$G = \{ \varphi^* = (X_1, X_2, \dots, X_n) | f(X) = f(g^*), \varphi \in \Phi \} \quad (24)$$

若 $G = \Phi$, 则在可行解的空间中, 那么每个解不仅是可行解而且是最优解, 此时进行优化则没有意义, 所以下文的讨论都是在 $G \subset \Phi$ 的情形之下.

定理 5 GWO 算法中, 对于灰狼群状态序列 $\{ \varphi(t) : t \geq 0 \}$, 最优灰狼群状态集合 G 是状态空间 Φ 上的一个闭集.

证明 设 $\forall \varphi_i \in G, \forall \varphi_j \in G$, 对于任意转移步长 $l, l \geq 1$, 由 Chapman-Kolmogorov 方程可得:

$$\begin{aligned} P_{\varphi_i, \varphi_j}^l &= \sum_{\varphi_{r1} \in \Phi} \dots \sum_{\varphi_{rl} \in \Phi} P(T_\varphi(\varphi_i) = \varphi_{r1}) \cdot P(T_\varphi(\varphi_{r1}) \\ &= \varphi_{r2}) \dots P(T_\varphi(\varphi_{rl-1}) = \varphi_j) \end{aligned} \quad (25)$$

其中, $P_{\varphi_i, \varphi_j}^l$ 表示灰狼群状态 φ_i 经过 l 步转移到状态 φ_j 的概率, 在式 (25) 展开式的每一项乘积中都存在 $P(T_\varphi(\varphi_{rs-1}) = \varphi_{rs}), (s \in [1, l])$ 项, 满足 $\varphi_{rs-1} \in G$ 且 $\varphi_{rs} \notin G$. 根据定义 6 可得灰狼群的转移概率为:

$$P(T_\varphi(\varphi_{rs-1}) = \varphi_{rs}) = \prod_{k=1}^{N_s} P(T_\varphi(X_{ik}) = X_{jk}) \quad (26)$$

由 $\varphi_{rs-1} \in G$ 和 $\varphi_{rs} \notin G$, 有 $f(X_s) > f(X_{s-1}) = f(g^*) = \inf(f(y)), y \in Y$, 则至少存在 $P(T_\varphi(\varphi_{rs-1}) = \varphi_{rs}) = 0$, 故此时 $P_{\varphi_i, \varphi_j}^l = 0$, 因此 G 是 Φ 上的一个闭集.

定理 6 灰狼群状态空间 Φ , 不存在非空闭集 Q , 使得 $Q \cap G = \emptyset$.

证明 (反证法) 假设状态空间 Φ 中存在一个非空闭集 Q , 且 $Q \cap G = \emptyset$. 设 $\varphi_i = (g^*, g^*, \dots, g^*) \in G, \forall \varphi_j = (x_{j1}, x_{j2}, \dots, x_{jd}) \in Q$, 有 $f(x_{js}) > f(g^*)$, 则经过有限次迭代会满足定理 1 中式 (11) ~ (13), 出现 $P_{\varphi_i, \varphi_j}^l(T_\varphi(X_i) = X_j) > 0$. 所以当步长 l 足够大时, $P_{\varphi_i, \varphi_j}^l$ 展开式中必存在某项乘积表达式, 使得其中的每一项一步转移概率 $P(T_\varphi(\varphi_{rs+i}) = \varphi_{rs+i+1})$ 都满足定理 1 中公式 (10), 即 $P(T_\varphi(\varphi_{rs+i}) = \varphi_{rs+i+1}) > 0$. 由定理 5 中式 (26) 可知 $P_{\varphi_i, \varphi_j}^l > 0$, 即 $P_{\varphi_i, \varphi_j}^l \neq 0$, 可得 Q 不是闭集, 与题设矛盾. 因此灰狼群状态的 Markov 链不可约, 故状态空间 Φ 不含 G 之外的闭集.

定理 7 假定 Markov 链有一个非空集 E , 且不存在另一个非空闭集 O , 使得 $E \cap O = \emptyset$, 则有:

$$\lim_{k \rightarrow \infty} P(X_k = j) = \begin{cases} \pi_j, & j \in E \\ 0, & j \notin E \end{cases} \quad (27)$$

定理 8 当灰狼群内部迭代趋于无穷时, 灰狼群状

态序列必将进入最优状态集 G .

证明 由定义 9 及定理 5、定理 6、定理 7 可得,定理 8 成立. 证毕.

引理 2 GWO 算法的灰狼群状态空间的一般状态转移至最优状态的转移概率为 1, 即 $\lim_{t \rightarrow \infty} P^{(t)}(\zeta_i \rightarrow \zeta_j) = 1$.

文献[9]基于 Markov 链的 PSO 算法收敛性分析时,证明了仅从粒子状态空间的性质并不能保证算法不具有平稳分布. 对于 GWO 算法的 Markov 模型的性质与 PSO 算法不同,由定理 3 可知,灰狼群状态序列 $\{\varphi(t) : t > 0\}$ 是有限齐次 Markov 链.

证明 假设灰狼群状态 $\zeta(j)$ 为最优解,若使得算法收敛,则经过无穷次的状态转移后,灰狼群状态空间从一般状态转移至最优状态的概率应为 1. 由于:

$$\lim_{t \rightarrow \infty} P^{(t)}(\zeta_i \rightarrow \zeta_j) = \sum_{k=1}^{N_g} P^{(t)}(\zeta_{ik} \rightarrow \zeta_{jk}) \quad (28)$$

由定理 2 可知,GWO 中灰狼群状态内部的每次迭代是根据灰狼群的等级特性向最优的状态转移,即只有最差的灰狼群状态位置进行更新. 故:

$$\lim_{t \rightarrow \infty} P^{(t)}(\zeta_i \rightarrow \zeta_j) = 1 \quad (29)$$

引理 3 GWO 算法满足 Condition 2.

定理 9 GWO 算法收敛到全局最优,即 $\lim_{t \rightarrow \infty} P\{X(t) \in G \mid X(0) = \Phi_0\} = 1$.

证明 由于 GWO 算法满足 Condition 1 和 Condition 2,灰狼优化算法每次迭代过程中,个体都会选择更新为最优个体的保留机制. 即当迭代趋于无穷时, $\lim_{k \rightarrow \infty} P(x_k \in R_{\xi, M}) = 1$, $\{x_k\}_{k=0}^{\infty}$ 是 GWO 算法迭代产生的解序列. 通过定理 4 可以得出,GWO 算法是一个全局收敛算法.

假设 在时刻 t 时, $X(t)$ 已进入全局最优状态解集 G ; 由 Condition 1 可得,在 $t+1$ 时刻, $X(t+1)$ 必然落在全局最优状态解集 G 中,即 $P\{X(t+1) \in G \mid X(t) \in G\} = 1$, 则

$$\begin{aligned} P\{X(t+1) \in G\} &= P\{X(t) \notin G\} \cdot P\{X(t+1) \in G \mid X(t) \notin G\} \\ &\quad + P\{X(t) \in G\} \cdot P\{X(t+1) \in G \mid X(t) \in G\} \\ &= [1 - P\{X(t) \in G\}] \cdot P\{X(t+1) \in G \mid X(t) \notin G\} \\ &\quad + P\{X(t) \in G\} \cdot 1 \end{aligned} \quad (30)$$

令 $P\{X(t+1) \in G \mid X(t) \notin G\} \geq v(t) \geq 0$,

$\lim_{t \rightarrow \infty} \prod_{i=1}^n (1 - v(i)) = 0$, 则可得:

$$\begin{aligned} 1 - P\{X(t+1) \in G\} &\leq [1 - v(t)] \cdot [1 - P\{X(k) \in G\}] \\ \Rightarrow 1 - P\{X(t+1) \in G\} &\leq [1 - P\{X(k) \in G\}] \cdot \prod_{i=1}^t [1 - v(i)] \\ \Rightarrow \lim_{t \rightarrow \infty} P\{X(t+1) \in G\} &\geq 1 \end{aligned}$$

由于 $P\{X(t+1) \in G\} \leq 1$, 故 $\lim_{t \rightarrow \infty} P\{X(t+1) \in G\} =$

1 , 即 $\lim_{t \rightarrow \infty} P\{X(t) \in G\} = 1$.

4.3 分析小结

综上所述,在对群体智能算法的改进研究中,不仅需要函数进行详细的仿真实验,同时也应该注重收敛性等理论证明. 因此给出如下建议:对于 GWO 算法的改进研究,可以从算法的种群规模初始化策略,如混沌初始化、量子旋转等;算法控制参数的非线性迭代策略;算法的种群等级制度的选择策略;算法的自适应性等,从而加快算法的收敛速度,减少迭代的次数和寻优时间. 对于 GWO 算法的理论研究,可以结合改进机制对其收敛性进行建模分析和证明^[26]. 通过对测试函数的仿真实验,可以对算法的参数进行调节,验证算法的鲁棒性等性能.

5 实验仿真与分析

鉴于 SFLA^[7]、PSO 算法^[9]、ABC 算法^[11,27] 基于 Markov 模型均给出了明确的收敛证明,因此,本文将从对比标准的 GWO 算法^[1] 与 SFLA、PSO 算法、ABC 算法的收敛特性和有效性. 选择 4 个标准测试函数、2 个偏移测试函数和 2 个旋转测试函数进行实验. 各算法的基本参数设置如表 1 所示.

表 1 群体智能优化算法的参数设置

算法	参数设置
GWO	种群 $N = 30$, 控制参数初值 $a_{first} = 2$, 终值 $a_{final} = 0$
SFLA	种群 $N = 50$, 族群 $m = 10$, 每个族群的青蛙个数 $n = 5$
PSO	种群 $N = 30$, $c_1 = c_2 = 2$, $w = 0.75$
ABC	蜜源数量为 $N/2$, 控制参数 $limit = 100$

文中,所有算法的最大迭代次数 $T_{max} = 1000$, 测试函数维数 $Dim = 30$.

5.1 测试函数

4 种标准测试函数分别为 Sphere (F_1)、Rastrigin (F_2)、Ackley (F_3)、Griewank (F_4), 其中, Sphere 函数是单峰函数, Rastrigin 函数、Ackley 函数和 Griewank 函数为多峰函数; 偏移测试函数^[24] 选择 Shifted Rastrigin (F_5) 和 Shifted Schwefel's 1.2 (F_6); 旋转测试函数^[25] 选择 Rotated Griewank (F_7) 和 Rotated Weierstrass (F_8).

5.2 算法的性能分析

本文仿真实验环境为: Windows 7 操作系统, Intel (R) Core (TM) i5-4210U CPU @ 2.4G, 4G 内存, Matlab 2018a. 每种算法针对测试函数均独立运行 20 次, 对应的收敛曲线如图 1、图 2、图 3 所示.

4 种标准测试函数实验结果如表 2 所示. 其中, 平均值体现算法的收敛精度和寻优能力; 标准差体现算法对抗局部极值的能力; 平均耗时体现算法的快速性; 算法的寻优成功率 (Success Rate), 定义 $\varepsilon < 10^{-8}$ 为寻优成功.

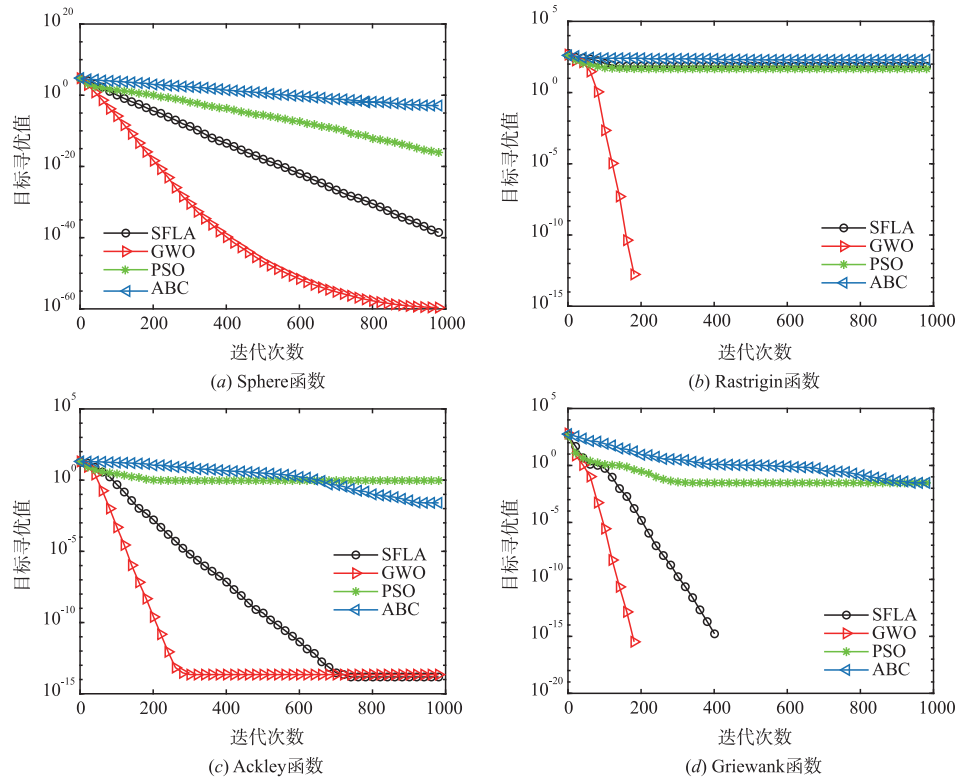


图1 4种标准测试函数收敛曲线

表2 4种标准函数的测试结果

函数		算法			
		F_1	F_2	F_3	F_4
SFLA	Mean	4.29E-41	7.84E+01	9.18E-15	6.27E-03
	Std	8.32E-41	2.00E+01	2.77E-15	8.26E-03
	Time/s	1.37E+01	1.49E+01	1.43E+01	1.63E+01
	Success Rate	100.00%	0.00%	100.00%	0.00%
PSO	Mean	2.08E-14	4.13E+01	1.51E+00	3.21E-02
	Std	6.22E-14	1.16E+01	6.37E-01	2.88E-02
	Time/s	1.55E+00	1.56E+00	1.64E+00	1.77E+00
	Success Rate	100.00%	0.00%	10.00%	0.00%
ABC	Mean	1.31E-03	2.03E+02	1.93E-02	5.66E-02
	Std	5.13E-04	1.10E+01	6.27E-03	4.94E-02
	Time/s	6.51E+00	6.91E+00	7.26E+00	8.20E+00
	Success Rate	0.00%	0.00%	0.00%	0.00%
GWO	Mean	2.78E-58	2.27E-14	1.66E-14	0.00E+00
	Std	3.67E-58	2.94E-14	2.38E-15	0.00E+00
	Time/s	6.40E-01	5.84E-01	6.00E-01	7.01E-01
	Success Rate	100.00%	100.00%	100.00%	100.00%

从表2可知,对于均值,Sphere函数的寻优结果表明SFLA、GWO、PSO算法均能达到100%的成功率,其中GWO算法的寻优效果最佳;Rastrigin函数的寻优结

果表明GWO算法最佳,且能保证寻优成功率;Ackley函数的寻优结果表明SFLA与GWO算法的寻优效果优于PSO算法和ABC算法,能够达到100%的寻优成功

率; Griewank 函数的寻优结果表明 GWO 算法寻优能力最佳. 从寻优耗时可得出, GWO 算法的平均耗时最短, SFLA 的平均耗时最长. 另外, 对于标准差, GWO 算法表

现出最佳抗局部极值的能力.

对于偏移和旋转测试函数, 收敛曲线如图 2、图 3 所示, 仿真实验的统计结果如表 3 所示.

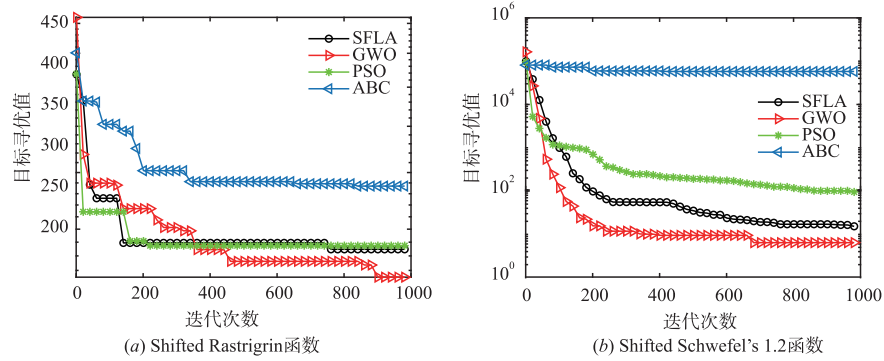


图2 2种偏移测试函数收敛曲线

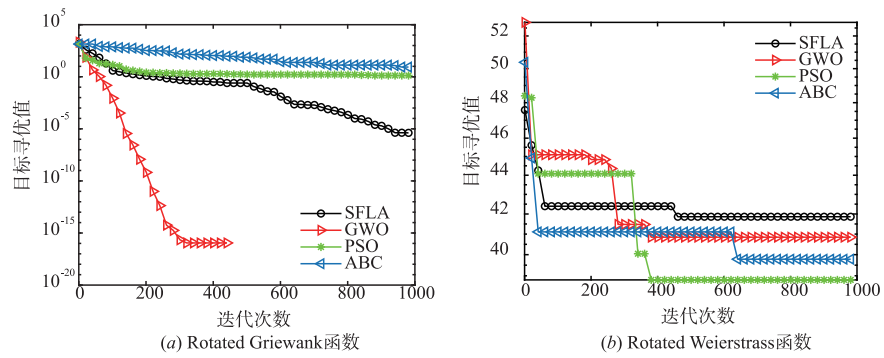


图3 2种旋转测试函数收敛曲线

表 3 偏移和旋转函数的测试结果

函数		F_5	F_6	F_7	F_8
算法	Mean	1.98E+02	1.30E+01	5.63E-03	4.14E+01
	Std	1.18E+01	6.75E+00	9.00E-03	4.34E-01
	Time	1.22E+01	1.84E+01	2.66E+02	1.86E+02
	Success Rate	0.00%	0.00%	0.00%	0.00%
SFLA	Mean	1.86E+02	1.33E+02	2.55E+00	4.19E+01
	Std	1.06E+01	5.77E+01	1.90E+00	1.13E+00
	Time	1.30E+00	2.75E+00	5.71E+01	6.27E+01
	Success Rate	0.00%	0.00%	0.00%	0.00%
PSO	Mean	2.26E+02	4.91E+04	1.75E+01	3.87E+01
	Std	8.20E+00	5.94E+03	3.70E+00	1.95E+00
	Time	6.48E+00	1.58E+01	3.77E+02	4.15E+02
	Success Rate	0.00%	0.00%	0.00%	0.00%
ABC	Mean	1.70E+02	6.76E+00	0.00E+00	4.16E+01
	Std	1.30E+01	1.74E+00	0.00E+00	5.86E-01
	Time	6.25E-01	1.86E+00	5.51E+01	6.09E+01
	Success Rate	0.00%	0.00%	100.00%	0.00%
GWO	Mean	1.98E+02	1.30E+01	5.63E-03	4.14E+01
	Std	1.18E+01	6.75E+00	9.00E-03	4.34E-01
	Time	1.22E+01	1.84E+01	2.66E+02	1.86E+02
	Success Rate	0.00%	0.00%	0.00%	0.00%

由于偏移测试函数的偏移矩阵的数值是随机的,由 rand 函数生成,因此 4 种群体智能算法的寻优结果均没有达到最优值(测试函数值为 0)。从均值和标准差综合来看,GWO 算法的收敛精度和抗局部极值能力优于 SFLA、PSO 和 ABC 算法,且 GWO 算法的耗时最短。

旋转测试函数 Rotated Griewank 函数的寻优计算结果表明 GWO 算法能够达到 100% 的成功率,SFLA 的寻优结果优于 PSO 算法和 ABC 算法,但耗时 SFLA 比 GWO 算法和 PSO 算法长;Rotated Weierstrass 函数寻优计算结果表明 4 种群体智能算法均未能达到理论最优值,其中 ABC 算法的寻优均值最小,SFLA 的标准差与 GWO 算法的标准差处于同一个数量级,但 GWO 算法的寻优耗时最短。

综上,标准的 GWO 算法的抗局部极值的能力要优于 SFLA、PSO 算法和 ABC 算法。且 GWO 算法寻优耗时最短,收敛精度也相对较佳。实验结果表明,GWO 算法有着较好的收敛特性,且抗局部极值能力也优于典型的群体智能算法。

6 结论

本文描述了灰狼的状态空间和灰狼群的状态空间,通过定义灰狼群状态转移序列建立了基于 Markov 链的 GWO 算法数学模型。通过分析该算法的 Markov 链性质,证明它是有限齐次 Markov 链。结合随机搜索算法的收敛准则,证明了 GWO 算法的收敛性。基于测试函数进行实验仿真,实验结果表明 GWO 算法比 SFLA、PSO、ABC 具有更好的收敛精度和寻优能力。本文对 GWO 算法随机过程的理论分析和得出的结论对算法的具体理解和算法的改进有一定的理论指导意义。下一步将对 GWO 算法进行更加深入的理论研究,如利用非动力学的理论、鞅过程理论对灰狼捕猎的运动和稳定性进行分析,另外,将对 GWO 算法进行改进,并将之应用于多目标优化、无线传感器节点部署等问题。

参考文献

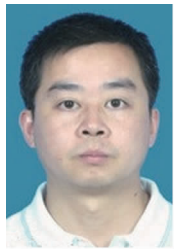
- [1] MIRJALILI S, MIRJALILI S M, LEWIS A. Grey Wolf optimizer [J]. *Advances in Engineering Software*, 2014, 69(1): 46–61.
- [2] DORIGO M, BIRATTARI M, STUTZLE T. Ant colony optimization [J]. *Computational Intelligence Magazine IEEE*, 2007, 1(4): 28–39.
- [3] WHITLEY D. A genetic algorithm tutorial [J]. *Statistics and Computing*, 1994, 4(2): 65–85.
- [4] TRELEA I C. The particle swarm optimization algorithm: convergence analysis and parameter selection [J]. *Information Processing Letters*, 2003, 85(6): 317–325.
- [5] BOOPATHI M, SUJATHA R, KUMAR C S, et al. Quantification of software code coverage using artificial bee colony optimization based on Markov approach [J]. *Arabian Journal for Science and Engineering*, 2017, 42: 3503–3519.
- [6] 朱庆保. 蚁群优化算法的收敛性分析[J]. *控制与决策*, 2006, 21(7): 763–768.
ZHU Qing-bao. Analysis of convergence of ant colony optimization algorithms [J]. *Control and Decision*, 2006, 21(7): 763–768. (in Chinese)
- [7] 骆剑平, 李霞, 陈泯融. 混合蛙跳算法的 Markov 模型及其收敛性分析[J]. *电子学报*, 2010, 38(12): 2875–2880.
LUO Jian-ping, LI Xia, CHEN Min-rong. The markov model of shuffled frog leaping algorithm and its convergence analysis [J]. *Acta Electronica Sinica*, 2010, 38(12): 2875–2880. (in Chinese)
- [8] GANG X, YU G S. Reprint of: On convergence analysis of particle swarm optimization algorithm [J]. *Journal of Computational and Applied Mathematics*, 2018, 340(1): 709–717.
- [9] 任子晖, 王坚, 高岳林. 马尔科夫链的粒子群优化算法全局收敛性分析[J]. *控制理论与应用*, 2011, 28(4): 462–466.
REN Zi-hui, WANG Jian, GAO Yue-lin. The global convergence analysis of particle swarm optimization algorithm based on Markov chain [J]. *Control Theory & Applications*, 2011, 28(4): 462–466. (in Chinese)
- [10] 李军华, 黎明. 元胞遗传算法的收敛性分析和收敛速度估计[J]. *模式识别与人工智能*, 2012, 25(5): 874–879.
LI Jun-hua, LI Ming. Convergence analysis and convergence rate estimate of cellular genetic algorithms [J]. *Pattern Recognition and Artificial Intelligence*, 2012, 25(5): 874–879. (in Chinese)
- [11] 宁爱平, 张雪英. 人工蜂群算法的收敛性分析[J]. *控制与决策*, 2013, 28(10): 1554–1558.
NING Ai-ping, ZHANG Xue-ying. Convergence analysis of artificial bee colony algorithm [J]. *Control and Decision*, 2013, 28(10): 1554–1558. (in Chinese)
- [12] 王继超, 李擎, 崔家瑞, 等. 一种改进的人工蜂群算法——粒子蜂群算法[J]. *工程科学学报*, 2018, 40(7): 871–881.
WANG Ji-chao, LI Qing, CUI Jia-rui, et al. An improved artificial bee colony algorithm: particle bee colony [J]. *Chinese Journal of Engineering*, 2018, 40(7): 871–881. (in Chinese)
- [13] 吴定会, 孔飞, 纪志成. 鸡群算法的收敛性分析[J]. *中南大学学报(自然科学版)*, 2017, 48(8): 2106–2112.
WU Ding-hui, KONG Fei, JI Zhi-cheng. Convergence analysis of chicken swarm optimization algorithm [J]. *Journal of Central South University (Science and Technology)*, 2017, 48(8): 2106–2112. (in Chinese)

- [14] SUZUKI J. A Markov chain analysis on simple genetic algorithms [J]. *IEEE Transactions on Systems, Man & Cybern.* 1995, 25(4): 655 - 659.
- [15] DOREA CH C Y, JUDINOR A G Jr, MORGADO R, et al. Multistage Markov chain modeling of the genetic algorithm and convergence results [J]. *Numerical Functional Analysis and Optimization*, 2010, 31(2): 164 - 171.
- [16] He X S, Yang X S, Karamanoglu M, Zhao Y X. Global convergence analysis of the flower pollination algorithm: a discrete-time Markov chain approach [J]. *Procedia Computer Science*. 2017, 108(1): 1354 - 1363.
- [17] 胡小平, 曹敬. 改进灰狼优化算法在 WSN 节点部署中的应用 [J]. *传感技术学报*, 2018, 31(5): 753 - 757.
HU Xiao-ping, CAO Jing. Improved grey wolf optimization algorithm for WSN node deployment [J]. *Journal of sensors and actuators*, 2018, 31(5): 753 - 757. (in Chinese)
- [18] 魏政磊, 赵辉, 李牧懂, 等. 控制参数非线性调整策略的灰狼优化算法 [J]. *空军工程大学学报(自然科学版)*, 2016, 17(3): 68 - 72.
WEI Zheng-lei, ZHAO Hui, LI Mu-dong, et al. A grey wolf optimization algorithm based on nonlinear adjustment strategy of control parameter [J]. *Journal of Air Force Engineering University: Natural Science Edition*, 2016, 17(3): 68 - 72. (in Chinese)
- [19] 段亚青, 王华倩, 乔学工. 基于测距和灰狼优化的无线传感器网络定位算法 [J]. *传感技术学报*, 2018, 31(12): 1894 - 1899.
DUAN Ya-qing, WANG Hua-qian, QIAO Xue-gong. Sensor node localization based on RSSI ranging and Grey Wolf optimizer algorithm in wireless sensor network [J]. *Journal of sensors and actuators*, 2018, 31(12): 1894 - 1899. (in Chinese)
- [20] 滕志军, 吕金玲, 郭力文, 等. 一种基于 Tent 映射的混合灰狼优化的改进算法 [J]. *哈尔滨工业大学学报*, 2018, 50(11): 40 - 49.
TENG Zhi-jun, LV Jin-ling, et al. An improved hybrid grey wolf optimization algorithm based on Tent mapping [J]. *Journal of Harbin Institute of Technology*, 2018, 50(11): 40 - 49. (in Chinese)
- [21] 薛俊杰, 王瑛, 李浩, 肖吉阳. 一种狼群智能算法及收敛性分析 [J]. *控制与决策*, 2016, 31(12): 2131 - 2139.
XUE Jun-jie, WANG Ying, LI Hao, XIAO Ji-yang. A smart wolf pack algorithm and its convergence analysis [J]. *Control and Decision*, 2016, 31(12): 2131 - 2139. (in Chinese)
- [22] LONG W, WU T B, CAI S H, et al. A novel grey wolf optimizer algorithm with refraction learning [J]. *IEEE Access*, 2019, 7: 78012 - 78028.
- [23] SOLIS F J, WETS J B. Minimization by random search techniques [J]. *Mathematics of Operations Research*, 1981, 6(1): 19 - 30.
- [24] LIANG J J, QIN A K, SUGANTHAN P N, et al. Comprehensive learning particle swarm optimizer for global optimization multimodal functions [J]. *IEEE Transactions on Evolutionary Computation*, 2006, 10(3): 281 - 295.
- [25] QIN A K, HUANG V L, SUGANTHAN P N. Differential evolution algorithm with strategy adaptation for global numerical optimization [J]. *IEEE Transactions on Evolutionary Computation*, 2009, 13(2): 398 - 417.
- [26] YANG X Sh, HE X Sh. *Mathematical Foundations of Nature-Inspired Algorithms* [M]. Switzerland: Springer Nature Switzerland AG, 2019.
- [27] KARABOGA D, BASTURK B. On the performance of artificial bee colony algorithm [J]. *Applied Soft Computing*, 2008, 8(1): 687 - 697.

作者简介



张孟健 男, 1996 年出生于安徽芜湖, 贵州大学电气工程学院硕士研究生, 研究方向为群体智能优化算法、无线传感器网络。
E-mail: 2311082906@qq.com



杨靖(通信作者) 男, 1973 年生于贵州贵阳, 教授, 博士, 硕士生导师, 主要研究方向为智能计算及多目标优化、物联网技术及应用等。
E-mail: jyang7@gzu.edu.cn