

自适应扩散混合变异机制微粒群算法*

吕艳萍¹, 李绍滋¹⁺, 陈水利², 郭文忠³, 周昌乐¹

¹(厦门大学 计算机科学系 智能多媒体技术实验室,福建 厦门 361005)

²(集美大学 理学院,福建 厦门 361021)

³(福州大学 计算机科学系,福建 福州 350002)

Particle Swarm Optimization Based on Adaptive Diffusion and Hybrid Mutation

LÜ Yan-Ping¹, LI Shao-Zi¹⁺, CHEN Shui-Li², GUO Wen-Zhong³, ZHOU Chang-Le¹

¹(Intelligent Multimedia Technology Laboratory, Department of Computer Science, Xiamen University, Xiamen 361005, China)

²(School of Sciences, Jimei University, Xiamen 361021, China)

³(Department of Computer Science, Fuzhou University, Fuzhou 350002, China)

+ Corresponding author: Phn: +86-592-2580080, Fax: +86-592-2580258, E-mail: szlig@xmu.edu.cn, <http://imt.xmu.edu.cn>

Lü YP, Li SZ, Chen SL, Guo WZ, Zhou CL. Particle swarm optimization based on adaptive diffusion and hybrid mutation. *Journal of Software*, 2007,18(11):2740–2751. <http://www.jos.org.cn/1000-9825/18/2740.htm>

Abstract: Conventional algorithms of particle swarm optimization (PSO) are often trapped in local optima in global optimization. In this paper, following an analysis of the main causes of the premature convergence, it proposes a novel PSO algorithm, which is called InformPSO, based on the principles of adaptive diffusion and hybrid mutation. Inspired by the physics of information diffusion, a function is designed to achieve a better particle diversity, by both taking into account their distribution and the number of evolutionary generations and adjusting their “social cognitive” abilities. Based on genetic self-organization and chaos evolution, clonal selection is built into InformPSO to implement the local evolution of the best particle candidate, gBest, and make use of a Logistic sequence to control the random drift of gBest. These techniques greatly contribute to breaking away from local optima. The global convergence of the algorithm is proved using the theorem of Markov chain. Experiments on optimization of unimodal and multimodal benchmark functions show that, comparing with some other PSO variants, InformPSO converges faster, results in better optima, is more robust, and prevents more effectively the premature convergence.

Key words: particle swarm optimization (PSO); premature convergence; information diffusion; clonal selection; Logistic sequence

摘要: 为了避免微粒群算法(particle swarm optimization,简称 PSO)在全局优化中陷入局部极值,分析了标准 PSO

* Supported by the National Natural Science Foundation of China under Grant No.60373080 (国家自然科学基金); the Natural Science Foundation of Fujian Province of China under Grant Nos.A0310009, A0510023 (福建省自然科学基金); the 985 Innovation Project on Information Technique of Xiamen University (2004-2007) of China (厦门大学 985 二期信息技术创新平台项目(2004-2007)); the Key Project of the Ministry of Education of China under Grant No.206073 (国家教育部科学技术研究重点基金项目)

Received 2006-08-10; Accepted 2006-10-10

算法早熟收敛的原因,提出了自适应扩散混合变异机制微粒群算法(InformPSO).结合生物群体信息扩散的习性,设计了一个考虑微粒分布和迭代次数的函数,自适应调整微粒的“社会认知”能力,提高种群的多样性;模拟了基因自组织和混沌进化规律,引入克隆选择使群体最佳微粒 gBest 实现遗传微变、局部增值,具有变异确定性;利用 Logistic 序列指导 gBest 随机漂移,进一步增强逃离局部极值能力.基于种群的随机状态转移过程,证明了新算法具有全局收敛性.与其他几种 PSO 变种相比,复杂基准函数仿真优化结果表明,新算法收敛速度快,求解精度高,稳定性好,能够有效抑制早熟收敛.

关键词: 微粒群算法;早熟收敛;信息扩散;克隆选择;Logistic 序列

中图法分类号: TP18 文献标识码: A

微粒群算法(particle swarm optimization,简称PSO)是Eberhart和Kennedy等人^[1]于1995年提出的一种新的全局进化优化算法,它源于对鸟类捕食行为的模拟.由于PSO算法采用实数编码,概念简单,实现容易,能够有效解决真实世界的复杂优化问题,已被成功应用于模式识别、多目标优化和神经网络训练等领域.

PSO 算法是基于群体与适应度的优化算法.微粒群的个体(称为微粒)代表问题的一个可行解,每个微粒具有位置和速度两个特征.位置坐标对应的目标函数值即是该微粒的适应度值,算法通过适应度来衡量粒子的优劣.算法首先随机初始化一个微粒群,再通过迭代找到最优解.在每一次迭代中,每个微粒跟踪两个“极值”来更新自己,一个是微粒自身所经历的最佳位置,即个体极值 pBest,另一个是整个微粒群经历过的最佳位置,称为全局极值 gBest.

在找到上述两个极值后,微粒通过如下两个公式来更新速度和位置:

$$V_i(t+1)=w \cdot V_i(t)+c1 \cdot r1 \cdot (pBest_i-X_i)+c2 \cdot r2 \cdot (gBest-X_i) \quad (1)$$

$$X_i(t+1)=X_i(t)+V_i(t+1) \quad (2)$$

其中, X_i 和 V_i 分别表示第 i 个微粒的位置和速度, w 是惯性权值, $c1$ 和 $c2$ 为加速因子, $r1$ 和 $r2$ 是在 $[0,1]$ 范围内的两个随机数.公式(1)中的第1部分称为“动量”部分,表示微粒当前自身运动状态对飞行的影响,为微粒提供一个必要的动量;第2部分称为“个体认知”部分,代表了微粒的个人经验,促使微粒飞向自身曾经发现的最佳位置;第3部分是“社会认知”部分,表示微粒间的信息共享和相互合作,它引导微粒飞向群体中的最佳位置.这3个部分之间的互相平衡和制约决定了算法的主要性能.

自PSO算法被提出后,由于简单、高效,引起了众多学者的极大关注,但也存在早熟收敛问题.因此,学者们致力于从PSO的各个方面来提高其性能,产生了许多有趣的变种.它们主要分为3类:第1类是通过调整PSO的参数平衡算法的全局和局部搜索能力,如PSO_w^[2],cf_PSO^[3];第2类是致力于设计不同类型的拓扑来提高算法性能,如PSO_cf_local^[4],UPSO(unified particle swarm optimization)^[5],FIPS(fully informed particle swarm)^[6],FDR_PSO(fitness-distance-ratio particle swarm optimization)^[7];第3类是将PSO和其他搜索技术相结合,形成混合PSO提高算法性能,如引入交叉提高种群多样性CPSO-H(cooperative particle swarm optimization- hybrid)^[8],CLPSO(comprehensive learning particle swarm optimizer)^[9],变异增强跳出局部极值AEPPO(adaptive escape particle swarm optimization)^[10].

本文根据生物群体信息扩散的习性以及基因自组织和混沌的进化规律,提出一种新型的微粒群算法.第1节分析标准 PSO 算法早熟收敛原因,提出一种具有自适应信息扩散混合变异机制的微粒群算法,并从理论上证明其全局收敛性.第2节描述用于测试算法性能的基准函数,给出本文算法的参数设置和实验结果,以及与其他 PSO 变种结果的比较分析.最后给出相应的结论以及进一步的工作方向.

1 自适应扩散机制混合变异 PSO 算法

1.1 早熟收敛原因

目前,虽然出现了标准 PSO 算法的多个变种,对算法性能有一定的改善,然而,早熟收敛仍然是 PSO 算法的一大难题,尤其是在复杂的多峰搜索问题中.在标准 PSO 迭代中,微粒分别从两个“极值”pBest 和 gBest 中获得

“个体认知”和“社会认知”能力,进而影响自身飞行速度,然而,所有微粒从群体最佳位置 $gBest$ 获得一样的“社会认知”能力未免有些武断.进一步说,不论距离 $gBest$ 远近,每个微粒均从 $gBest$ 学习到相同的信息.在这种情形下,所有微粒快速地受其吸引,即群体多样性迅速降低.而在拥有大量局部极值的复杂搜索环境中, $gBest$ 通常是一个局部最优解,因此,PSO 算法容易陷入一个局部解,从而导致早熟收敛.

PSO 算法的另外一个早熟收敛的原因是,微粒的速度更新公式的后两部分对群体当前最佳微粒 $gBest$ 没有任何贡献,在搜索中, $gBest$ 总是沿着前一代速度方向飞行.这种情况下,在拥有大量局部极值的复杂环境中, $gBest$ 容易陷入局部极值,而没有提供任何机制逃离局部最优解.换句话说, $gBest$ 作为当前重要的记忆单元,只是在强势微粒(适应度高于 $gBest$ 的微粒)的作用下被迫地学习,即只拥有强迫学习机制.

1.2 改进算法基本原理

1.2.1 自适应信息扩散函数

在生命系统实际活动中,个体之间的信息扩散是一个时间过程.显然,靠近群体当前最佳位置 $gBest$ 的微粒能够根据 $gBest$ 的引导迅速地改变自身飞行方向和速度扑向它.然而,远离 $gBest$ 的微粒只能比较缓慢地向其移动,甚至脱离 $gBest$ 的掌控.可以假定,在生物群体活动中,信息扩散是在一段不太长的时间内完成的,那么,远离 $gBest$ 的微粒获得的信息比靠近 $gBest$ 的微粒要少.

因此,在本文的 PSO 算法中,微粒的速度更新公式为

$$V_i(t+1)=w \cdot V_i(t)+c1 \cdot r1 \cdot (pBest_i-X_i)+H_i \cdot c2 \cdot r2 \cdot (gBest-X_i) \quad (3)$$

$$H_i = [1 - (d_i + 1) / (\max_{1 \leq j \leq n} d_j + 1)] \cdot [(n + 1) / (n + n' + 1)] \cdot [1 - t / T] \quad (4)$$

其中, H_i 为信息扩散函数, t 为当前进化代数(当前迭代代数), T 是设定的进化总代数.信息扩散函数由3部分组成,即

1) $1 - (d_i + 1) / (\max_{1 \leq j \leq n} d_j + 1)$ 将移动尺度与微粒隶属于 $gBest$ “周围”的程度联系起来.越靠近 $gBest$ 的微粒,受其影响越大,移动尺度相对越大;反之,移动尺度越小.同一代微粒群变尺度向 $gBest$ 靠拢,有利于提高群体多样性.其中, d_i 是第 i 个微粒与 $gBest$ 的距离,距离是用两个微粒的位置差异来衡量的;1 是为了避免分母为 0.

2) $(n + 1) / (n + n' + 1)$ 将移动尺度与微粒的分布联系起来.搜索时,判断微粒进入 $gBest$ 内围的情况,分布于内围的微粒数目多,则移动尺度小,以此避免群体多样性下降得太快. n' 是内围微粒的数目, n 为其余微粒数目,1 是为了避免信息扩散函数为 0.

3) $1 - t / T$ 将移动尺度与进化代数联系起来.在算法进化初期,微粒以较大的速度向 $gBest$ 移动,有利于加快搜索;随着进化代数的增加,移动尺度逐渐变小,有利于进化后期群体多样性的保持.

由公式(3)、公式(4)得知,微粒速度更新公式增加了一个信息扩散函数,利用微粒的分布和迭代代数自适应控制“社会认知”学习的移动尺度,从而能够充分利用先验知识来指导微粒移动,增强局部搜索能力,提高群体多样性,有效增强避免早熟收敛能力.

1.2.2 克隆选择算子

在多峰优化中,作为 PSO 算法的一个重要记忆单元, $gBest$ 经常是一个局部最优解,容易给其余微粒提供错误的导向信息,导致算法早熟收敛.因而,除了因强势微粒而进行强迫学习外, $gBest$ 还必须具备逃离局部极值的能力,即变异能力,实现记忆信息的自学习优化.然而,PSO 变种的变异算子大多是随机变异,具有盲目性,算法不能进行精确的局部搜索;也有一部分算子是确定性变异,最终难以逃脱陷入局部极值的厄运.本文提出了一种混合变异,结合了克隆选择以及非线性时间序列来指导 $gBest$ 的变异,使其自身优化学习具有定向性,又具有随机规律性的特点.

克隆选择^[11]的基因进化规律表明,作为遗传信息的记忆单元的基因,进化中拥有自我复制能力;然而多变的环境促使基因自我复制出错,导致遗传微变;大部分突变基因由于突变容易直接导致死亡,只有很小一部分突变基因能够正常繁殖.小区域内基因的自我复制、遗传突变、变异死亡促使原始基因多向繁殖和广泛繁殖.受基因进化规律启发,新算法设计克隆选择来改变群体最佳微粒 $gBest$ 的飞行方向,逃离局部极值.基本操作如下:

- 1) 群体当前最佳微粒 $gBest$ 无性克隆繁殖成一个规模为 m 的记忆亚群体,亚群体中个体与 $gBest$ 具有完全相同的属性;
- 2) 记忆亚群体通过柯西分布被编辑和变异,实现小范围内随机遗传微变,形成个体各异的新记忆亚群体;
- 3) 从新亚群体中选择适应度最高的个体作为下一代微粒群速度更新的 $gBest$.

通过克隆选择,群体当前最佳微粒 $gBest$ 在局部范围内具有自学习能力,实现局部区域内的优化、增值,变异具有确定性,增强了算法跳出局部极值的能力;并且能够正确引导其他微粒的飞行,大幅度提高了收敛速度.

1.2.3 Logistic 序列变异

克隆选择算子是确定性的遗传微变,对于局部极值点距离全局最优值点不远的函数优化而言,能够有效增强 $gBest$ 跳出局部解的能力;对于局部极值点与全局最优值点距离甚远的函数而言,克隆选择只是加速算法收敛于某一局部极值点.当 $gBest$ 处于某一局部极值点,且其小范围内无更优极值点时,克隆选择算子将不起效用,此时,需要一种随机又有规律的大漂移机制来协助 $gBest$ 逃离局部极值.

非线性时间序列具有随机性、遍历性等特性,已被引入到进化计算中^[12].当克隆选择算子失效时,新算法亦引入Logistic序列扰动指导 $gBest$ 的变异学习.本文的算法中需要的是大漂移,即大的遍历范围,因此,着重考察非线性时间序列控制参数 r 及序列初始值 $x(0)$ 对序列遍历范围的影响来确定参数.

考察典型的非线性 Logistic 映射:

$$x(t+1)=r \cdot x(t) \cdot (1-x(t)), r \in N, x(0) \in [0, 1] \tag{5}$$

$$gBest(t+1)=gBest(t)+[r3 < P_m] \cdot x(t) \cdot r4 - [r3 \geq P_m] \cdot x(t) \cdot r5 \tag{6}$$

其中, $r3, r4, r5$ 为随机数, $[r3 < P_m] = \begin{cases} 1, & \text{if } r3 < P_m \\ 0, & \text{else} \end{cases}$, P_m 为变异概率, 介于 0.05~0.5 之间, 经验取为 $P_m+0.45+$

$0.45 \cdot \frac{\exp(1: Dim) - \exp(1)}{\exp(Dim) - \exp(1)}$, Dim 为所求解优化问题的维数, $Dim \neq 1$. 图 1 显示了参数 r 和 x 对 Logistic 序列的影响.

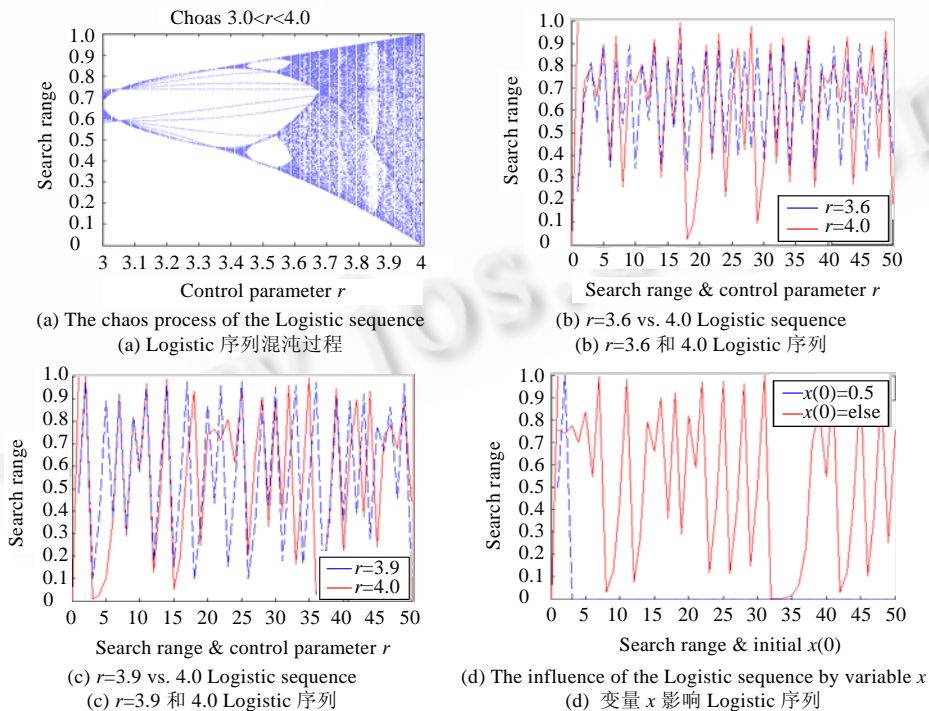


Fig.1 Logistic sequence covering range

图 1 Logistic 序列覆盖范围

如图 1(a)所示,当 $3 \leq r \leq 4$ 时,序列由周期走向非线性复杂状态,大约当 $r=3.56$ 时,序列步入混沌.图 1(b)、图 1(c)考察控制参数 r 对 Logistic 序列搜索范围的影响,当 $r=4$ 时,序列覆盖范围较为广泛,因此,引入本文算法的非线性 Logistic 序列的控制参数为 4.图 1(d)着重考察控制参数 $r=4$ 时,序列初始值 $x(0)$ 对 Logistic 序列的影响,当 $r=4, x(0)=0.5$ 时,迭代后期序列值为 0,当 $x(0)$ 为其他值时,序列呈现混沌状态.故在本文的算法中,非线性 Logistic 序列的控制参数为 4,初始值除了 0.5 以外, $[0,1]$ 之间其他值均可.

1.2.4 InformPSO 算法流程(如图 2 所示)

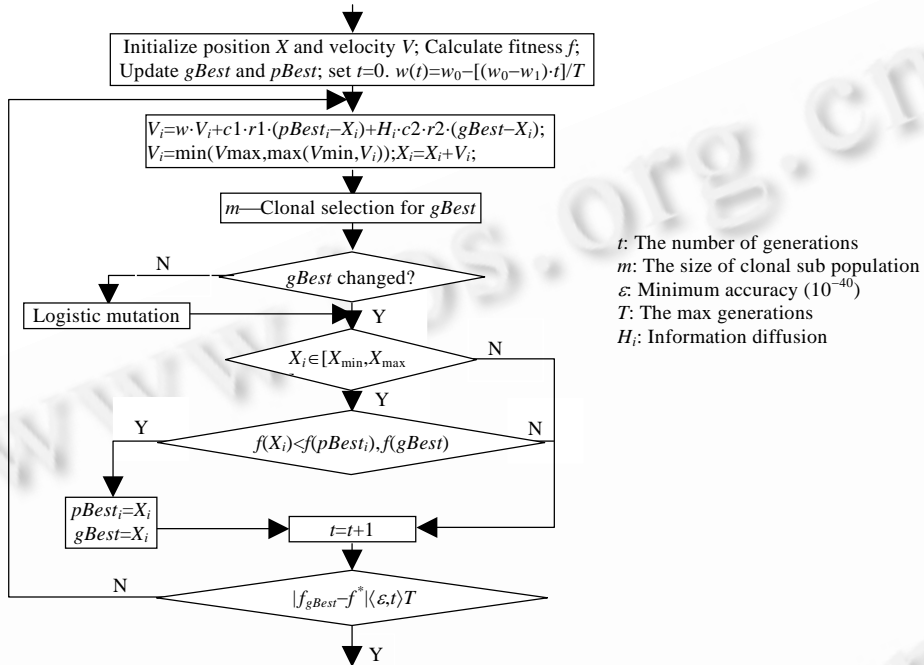


Fig.2 The flow chart of the InformPSO algorithm

图 2 InformPSO 算法的流程图

1.3 全局收敛性

在本文的改进 PSO 算法中,微粒群和记忆单元(pBest 和 gBest)分别构成了不同的随机状态转移过程.微粒群是 PSO 算法的基本种群,强调全局搜索,信息扩散函数控制微粒群飞向群体最佳位置 gBest 的移动速度,进行定义域变尺度的局部搜索,保持种群的多样性.重要的记忆单元 gBest,一方面通过强势微粒强迫学习来记忆搜索过程,加快搜索速度;另一方面利用克隆随机变异选择操作的确定性实现小范围内自身优化增值,增强逃离局部最优能力,正确导向其他微粒的飞行,加速收敛;而当 gBest 处于局部极小而克隆选择又失效时,Logistic 序列扰动 gBest 的飞行,在确定性微变的基础上实现随机逃逸突变,进一步增强跳出局部最优的能力.以上机制的相互协作,正确指引着改进的微粒群算法向全局最优解收敛.

虽然已有 PSO 变种^[10]描述可将文献[13]的 GA 等演化策略收敛性用来证明 PSO 算法的全局收敛性,但未仔细推敲改进前后算法的区别,即以其简单应用来证明 PSO 全局收敛,存在不足.本文在分析传统微粒群收敛性的同时也进行差异比较,如下描述:

假设 1. (i) 问题 P 的定义域 Ω 为 R^n 中的有界闭区域;(ii) 目标函数 $f(x)$ 是区域 Ω 上的连续函数.

由假设 1 易知, f 在 Ω 内的最优解集 $F^* \neq \emptyset$.

定义 1. 定义优化问题的全局最优解集: $F^* = \{F \in S | f(A) = f^* = \min f(F'), F' \in S\}$.

集合 S 称为种群空间,是所有可能出现种群的集合.对于基本种群 $F, F \in S, \delta(F) = |F \cap F^*|$ 表示基本种群包含最

优解的数目, $S_1 = \{F \in S_1 | \delta(F) \geq 1\}$ 为最优种群空间, $S_0 = S \setminus S_1$ 为普通种群空间.

定义 2. 对于任意的随机初始状态 F_0 , 若 $\lim_{t \rightarrow \infty} P\{\delta(F(t)) \geq 1 | F(0) = F_0\} = 1$, 则称算法以概率 1 收敛于全局最优解; 或者随机序列 $\{\xi_n\}$, 若对 $\forall r > 0$, 有 $P\left\{\bigcap_{n=1}^{\infty} \bigcup_{k \geq n} \xi_k - \xi \geq r\right\} = 0$, 则称随机序列 $\{\xi_n\}$ 以概率 1 收敛于随机变量 ξ .

定理 1. 对于界约束最优化问题, $\min f(x), x = (x_1, x_2, \dots, x_n), s.t. a_i \leq x_i \leq b_i, i = 1, 2, \dots, n$, 传统的微粒群优化算法是局部收敛的, 并且以概率 1 收敛于全局最优解.

证明: 与文献[10,13]类似. □

定理 2. 对于界约束最优化问题, 自适应扩散混合变异机制微粒群算法是局部收敛的, 且以概率 1 收敛于全局最优解.

证明: 新算法局部收敛证明类似于文献[10]; 下面证明以概率 1 收敛于全局最优解.

假设 1 可知, 界约束优化问题的最优解集存在. 由算法思想可知, 新算法增加了 3 个操作.

自适应扩散操作: 由公式(3)可知, 前后两代的位移变化量变为

$$\Delta x_i^j = w \cdot v_i + c1 \cdot (pBest_i^j - x_i^j) \cdot r1 + H_i \cdot c1 \cdot (gBest_i^j - x_i^j) \cdot r2, r1, r2 \sim N(0, \sigma_j),$$

所以, $\Delta x_i^j \sim N(0, \sigma_j)$.

克隆选择操作: 由于克隆变异选择操作之后获得的最优解至少比操作之前的要好, 因此, 最优种群仍然不会进化到普通种群空间, 即 $P_{10} = 0$.

Logistic 序列变异: $\forall r > 0$, 记 $B(x^*, r) = \{x \in \Omega | |f(x) - f^*| < r\}$ 为全局最优点 x^* 的邻域. 设随机事件序列, $A_t = \{F(t) \cap B(x^*, r) \neq \emptyset\}$ 表示迭代序列第 t 次迭代进入 $B(x^*, r)$. 对于给定的 r , 由 Logistic 序列的遍历性, 事件 A_1 发生, 必导致事件 A_2 发生, 以此类推, 于是有 $A_1 \subseteq A_2 \subseteq \dots \subseteq A_t \subseteq \dots$, 从而有 $P(A_1) \leq P(A_2) \leq \dots \leq P(A_t) \leq \dots$, 又因 $0 \leq P(A_t) \leq 1$, 所以, $\lim_{t \rightarrow \infty} P(A_t)$ 存在.

再设随机变量序列 $\delta_t = \begin{cases} 1, & \text{第 } t \text{ 次迭代进入 } B(x^*, r) \\ 0, & \text{第 } t \text{ 次迭代不进入 } B(x^*, r) \end{cases}, t = 1, 2, \dots$, 则 $A_t = \{\delta_t = 1\}$.

$$\text{令 } P(A_t) = P\{\delta_t = 1\} = p_t, P\{\delta_t = 0\} = 1 - p_t, S_t = \frac{1}{t} \sum_{i=1}^t \delta_i, t = 1, 2, \dots,$$

则
$$E(S_t) = \frac{1}{t} \sum_{i=1}^t p_i, D(S_t) = \frac{1}{t^2} \sum_{i=1}^t D(\delta_i) = \frac{1}{t^2} \sum_{i=1}^t p_i \cdot (1 - p_i) \leq \frac{1}{4 \cdot t}.$$

其中, $E(S_t), D(S_t)$ 分别为序列 S_t 的数学期望和方差.

由契比雪夫不等式, 有 $P\{|S_t - E(S_t)| \geq r\} \leq \frac{D^2(S_t)}{r^2} = \frac{1}{4 \cdot m \cdot r^2}$, 于是, $\lim_{t \rightarrow \infty} P\{|S_t - E(S_t)| \geq r\} = 0$.

又因 $\delta_t = t \cdot S_t - (t-1) \cdot S_{t-1}$, 从而 $\lim_{t \rightarrow \infty} P\{|\delta_t - E(\delta_t)| \geq r\} = 0$, 于是

$$P\left\{\bigcap_{t=1}^{\infty} \bigcup_{k \geq t} |\delta_k - E(\delta_k)| \geq r\right\} = P\left\{\lim_{t \rightarrow \infty} \bigcup_{k \geq t} |\delta_k - E(\delta_k)| \geq r\right\} = \lim_{t \rightarrow \infty} P\{|\delta_t - E(\delta_t)| \geq r\} = 0.$$

上式说明随机变量序列 $\{\delta_t\}$ 以概率 1 收敛, 因此, 随机事件序列 $\{A_t\}$ 亦以概率 1 收敛, 即 $P\left\{\bigcap_{t=1}^{\infty} \bigcup_{k \geq t} A_k\right\} = 0$. 因此

定理 2 得证, 即自适应扩散混合变异机制微粒群算法以概率 1 收敛至全局最优解. □

2 算法有效性验证

2.1 基准函数问题

为了评价本文提出的自适应扩散混合变异机制微粒群算法(InformPSO)对函数优化时的收敛速度、全局优化能力等求解性能, 引入 10 个基准优化问题^[6,9]进行分析, 将该算法与其他 PSO 算法变种进行了对比. 这 10 个基准函数具有不同的特点, 可以充分考察新型算法对不同类型问题的优化性能.

在表 1 中,高维测试函数形态各异,具有很好的测试性能,其中, $f(x^*)$ 为全局最小值, x^* 为全局最优解点. $f_1 \sim f_5$ 是连续的、单峰函数,通常用于检验算法收敛速度.函数 f_5 (Rosenbrock)是一个经典的复杂优化问题,取值区间内走势平坦,为算法提供少量信息,要收敛到全局最优解点机会微乎其微,用于测试算法收敛速度.函数 $f_6 \sim f_{10}$ 是非线性多峰函数,存在大量局部极值,可有效检验算法的群体多样性、全局搜索性能、逃离局部极值并避免早熟收敛能力.

Table 1 Ten benchmark functions

表 1 10 个基准函数

Name	Function	x^*	$f(x^*)$
Sphere	$f_1 = \sum_{i=1}^D x_i^2$	[0,0,...,0]	0
Hyper-Ellipsoid	$f_2 = \sum_{i=1}^D i \cdot x_i^2$	[0,0,...,0]	0
Sum of different powers	$f_3 = \sum_{i=1}^D x_i ^{i+1}$	[1,1,...,1]	0
Rotated hyper-ellipsoid	$f_4 = \sum_{i=1}^D \left[\sum_{j=1}^i x_j \right]^2$	[0,0,...,0]	0
Rosenbrock	$f_5 = \sum_{i=1}^D [100 \cdot (x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$	[1,1,...,1]	0
Griewank	$f_6 = \frac{1}{4000} \sum_{i=1}^D x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	[0,0,...,0]	0
Ackley	$f_7 = 20 + e - 20 \exp\left(-0.2 \sqrt{\frac{1}{D} \sum_{i=1}^D x_i^2}\right) - \exp\left[\frac{1}{D} \sum_{i=1}^D \cos(2\pi x_i)\right]$	[0,0,...,0]	0
Rastrigin	$f_8 = \sum_{i=1}^D [x_i^2 - 10 \cdot \cos(2\pi x_i) + 10]$	[0,0,...,0]	0
Weierstrass	$f_9 = \sum_{i=1}^D \left(\sum_{k=0}^{k_{\max}} [a^k \cdot \cos(2\pi b^k (x_i + 0.5))] \right) - D \sum_{k=0}^{k_{\max}} [a^k \cdot \cos(2\pi b^k \cdot 0.5)]$ $a=0.5, b=3, k_{\max}=20$	[0,0,...,0]	0
Schwefel	$f_{10} = 418.9829 \times D - \sum_{i=1}^D x_i \sin\left(x_i ^{1/2}\right)$	[420.96,420.96,...,420.96]	0

2.2 参数设置

为了清晰地比较多个PSO算法变种对测试函数的性能,参数设置与文献[6,9]一致,惯性权值 $w=0.95 \sim 0.4$,加速因子 $c_1=c_2=2$,详见表 2.在表 2 中,Initial range为微粒搜索位置范围,Size为群体规模,Dim为问题维数,Max_Gen为最大迭代次数.除了求解精度 $\varepsilon=10^{-40}$ 以外, $f_1 \sim f_5$ 参数设置与文献[6]一致,而 $f_6 \sim f_{10}$ 与文献[9]一致.InformPSO还须设置克隆选择中亚群体的规模 m ,假设事先并不需要知道被优化函数是单峰还是多峰,为了更有效地帮助gBest实现局部增值, m 经验设置为 5.

Table 2 Parameter setting for ten benchmark functions

表 2 10 个基准函数参数设置

Function	Initial range	Max_Gen	Size	Dim
f_1	[-100,100]	3 000	10	20
f_2	[-5.12,5.12]	3 000	10	10
f_3	[-1,1]	3 000	10	10
f_4	[-65,65]	3 000	10	10
f_5	[-2.048,2.048]	3 000	10	10
f_6	[-600,600]	3 000	10	10
f_7	[-32.768,32.768]	3 000	10	10
f_8	[-5.12,5.12]	3 000	10	10
f_9	[-0.5,0.5]	3 000	10	10
f_{10}	[-500,500]	3 000	10	10

2.3 对比实验

本节实现了 PSO_w, FDR_PSO, CLPSO 和 InformPSO 这 4 种 PSO 算法, 对第 2.1 节中的 10 个基准函数进行测试, 4 种算法设置了相同的参数, 见表 3.

Table 3 Experimental results for unimodal functions

表 3 单峰函数实验结果

Function	Best/Worst fitness value			
	PSO_w	FDR_PSO	CLPSO	InformPSO
f_1	7.6987e-012	3.7170e-033	6.5649e-015	9.8482e-041
	2.0264e-008	6.7152e-032	1.1590e-014	8.9789e-041
f_2	9.5991e-041	4.7203e-041	4.1441e-035	7.9497e-041
	9.8230e-038	9.2995e-041	1.7845e-032	8.7664e-041
f_3	8.6953e-041	2.8954e-041	1.1932e-041	7.1815e-041
	8.9061e-041	5.3017e-041	3.3693e-041	8.0803e-041
f_4	9.3630e-020	5.7246e-033	0.008 5	1.0624e-035
	4.7095e-017	4.1311e-030	0.109 2	1.3780e-033
f_5	5.1222e-020	9.1981e-041	7.5996e-011	9.7949e-041
	1.1557e-016	9.7118e-041	3.4984e-006	9.8454e-041

近年来,有不少学者也提出了标准 PSO 的多个变种.虽然本文没有编程实现所有我们感兴趣的 PSO 变种,但是对于基准函数优化,本文设置相同的参数进行测试,并将结果与文献[9]进行比较.多个 PSO 变种描述如下:

- 1) 惯性权值线性递减(PSO_w)^[2]
- 2) 收缩因子局部PSO(PSO_cf_local)^[4]
- 3) 一元化PSO(UPSO)^[5]
- 4) FIPS^[6]
- 5) 基于适应度距离比例PSO(FDR_PSO)^[7]
- 6) 综合学习PSO(CLPSO)^[9]
- 7) 自适应扩散混合变异 PSO(InformPSO)

图 3(a)~图 3(i)图显示了 PSO_w, FDR_PSO, CLPSO 和 InformPSO 在测试函数 $f_1 \sim f_{10}$ 的优化实验结果.从图 3(a)~图 3(e)可以看出,对于单峰函数的优化, InformPSO 能够保持较快的搜索速度,并有效地收敛至全局最小点. PSO_w, CLPSO 对于函数 f_1, f_4, f_5 的优化,最优精度只能达到 10^{-20} 之内;对于 f_3, f_4 , 最优精度能达到 10^{-30} 左右,但以牺牲收敛速度为代价, PSO_w 甚至会出现早熟收敛现象;而 CLPSO 对上述 5 个单峰函数的优化,收敛则非常缓慢.对于 f_2, f_3, f_5 这 3 个单峰函数, InformPSO 在搜索的全部阶段最优适应度值均是快速的下降过程,在 1 000 次迭代以内最优适应度值精度均能达到目标精度 10^{-40} , 迅速收敛于全局最优解,具备良好的稳定性,如图 3(b)、图 3(c)、图 3(e)所示.对于 f_1, f_4 函数的优化, InformPSO 比其他 3 个 PSO 变种收敛精度高,收敛速度快,特别是与 PSO_w, CLPSO 比较,如图 3(a)、图 3(d)所示.从图 3(a)~图 3(e)中也可观察到 FDR_PSO 对单峰函数的优化性能较佳,但收敛速度仍然不如 InformPSO.

在单峰函数 $f_1 \sim f_5$ 上,每种算法被连续运行 30 次.其中适应度值以函数值表示,所有表 3 的数据均来自于本文的实验结果.

表 3 的各种算法得出的最优/最差适应度的实验结果更加清晰地说明了 InformPSO 运行性能优于其他 3 个 PSO 变种. FDR_PSO 收敛速度比 PSO_w 和 CLPSO 快,收敛精度高; PSO_w 在 f_2, f_3, f_4, f_5 函数上比 CLPSO 有较快的收敛速度;而在 f_1 函数上, CLPSO 取得更好的结果. InformPSO 在表 3 的各项数据显示,该算法能够快速地找到搜索方向,在较少迭代次数内达到求解精度,搜索性能和全局收敛速度显著提高;数据差异较小,具有更好的稳定性,性能明显优于其他 3 种算法.上述现象的产生主要是因为,在其他 3 种算法中,群体最佳微粒 gBest 变化缓慢,不能正确指引其他微粒的飞行;而 InformPSO 在每一次迭代中,由于克隆变异更具确定性, gBest 局部增值,算法能在一个较短时间内找到正确的搜索方向,显著提高了收敛速度;而 PSO 的新近变种 CLPSO 对单峰函数优化性能最差,主要是由于该算法通过各个个体 pBest 的交叉学习保持种群多样性,没有 gBest 学习算子而导致收敛较为缓

慢。

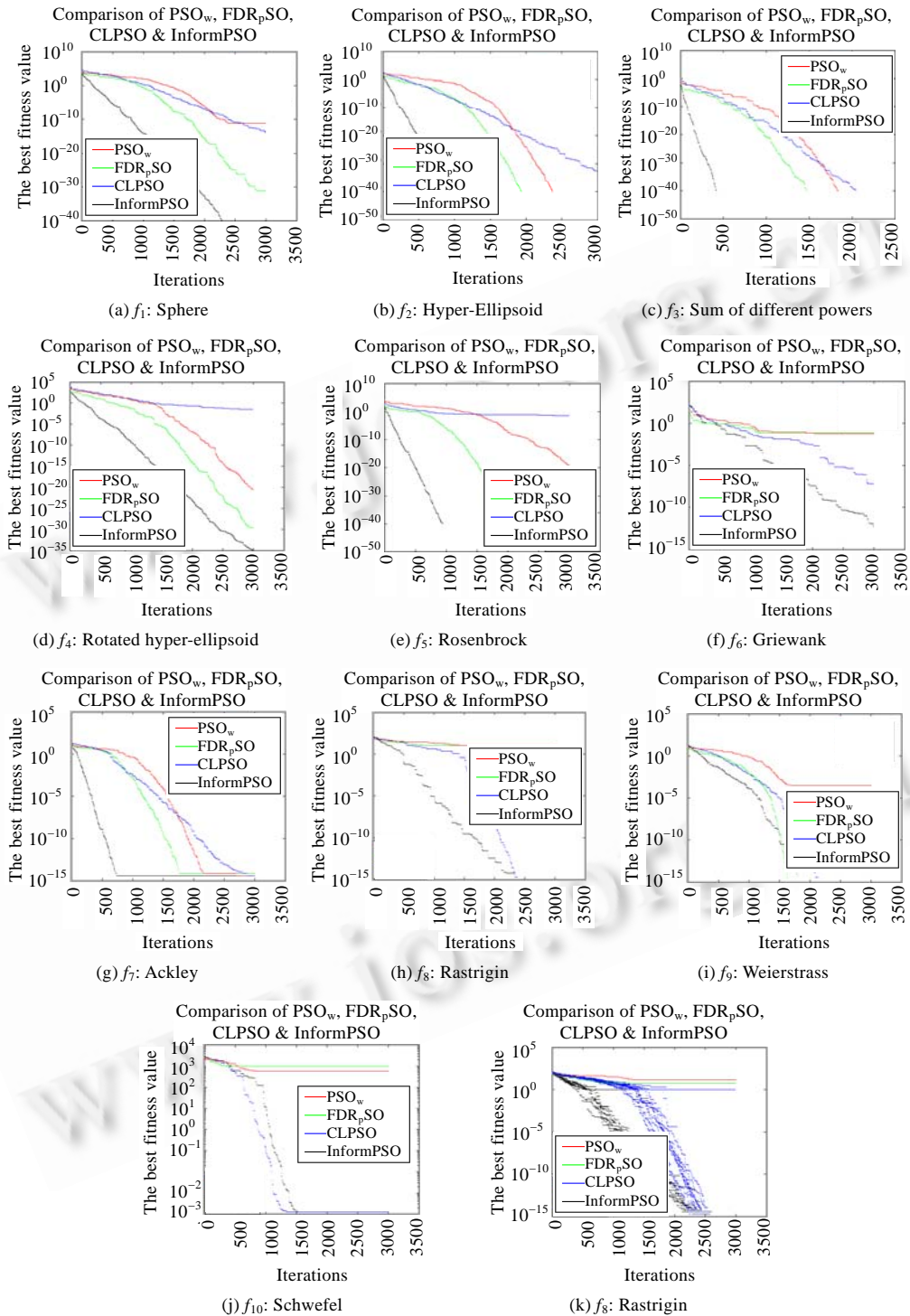


Fig.3 InformPSO vs. PSO_w , FDR_pSO , $CLPSO$ on benchmark functions

图3 InformPSO 与 PSO_w , FDR_pSO , $CLPSO$ 在基准函数上的比较

从图 3(f)~图 3(j)可以观察到,除了优化 f_9 (weierstrass)函数时InformPSO收敛比FDR_PSO稍慢以外,在其他多峰函数优化测试结果均优于PSO_w和FDR_PSO;对于函数 f_6, f_8, f_9 ,它能在最大迭代次数内找到这 3 个多峰函数的全局最小点.观察PSO_w和FDR_PSO算法最优适应度值变化曲线,除了 f_7 (ackley)以外,这两个变种在迭代初期种群微粒运行良好,但迭代后期容易失去种群的多样性而收敛于某一局部极值点,而且也没有能力改进群体最佳适应度值,即gBest不具备跳出局部极值的能力.根据最优适应度变化曲线图(图 3(f)、图 3(h)~图 3(j)),我们发现InformPSO在 f_6, f_8, f_9, f_{10} 函数优化上不会发生类似现象,这主要是由于在新算法的速度更新公式中,信息扩散函数根据微粒的分布情况和迭代次数,控制着所有微粒飞向当前最佳位置gBest的速度分量,使它们变尺度向其靠近,不容易失去群体多样性,有效避免了早熟收敛;gBest的克隆选择操作使得它在每一迭代中均匀变异,即使陷入局部极值,也具有微变能力逃离局部极值,并且克隆选择的变异又具有确定性,使gBest局部增值,因此,在迭代过程中,最优适应度值不断发生变化,正确导向其他微粒的飞行,极大地加快了收敛速度.

观察图 3(f)、图 3(h)~图 3(j)中InformPSO和CLPSO算法最优适应度值变化曲线,InformPSO测试性能与CLPSO相差无几.图 3(k)是它们在函数 f_8 (rastrigin)连续测试 20 次的运行结果,可以看出,InformPSO在该函数的测试上,跳出局部最优能力大为增强,当它陷入一个局部最优值为 0.995 0 的点时,经过一些迭代,能够逃离该局部极值点;而CLPSO只是保持多样性,缺乏变异能力,所以,当算法陷入一个局部最优值点时,迭代后期不具备能力逃离该点,最终收敛于该局部极值点.图 3(g)也显示了InformPSO比CLPSO具备更佳的收敛速度.但对于Ackley和Schwefel函数,新算法和CLPSO只能收敛于一个局部最优值,主要是由于这两个函数局部最优值与全局最优值的距离较远,新算法的克隆选择算子是微突变,长期效用微弱;而纯粹的非线性时间序列指导gBest变异更多的是具有随机性特点,因此对于此类函数优化不足以跳出局部极值.图 3(g)、图 3(j)表明,这些PSO变种都无法到达Ackley和Schwefel函数的全局极小点.

在多峰函数 $f_6 \sim f_{10}$ 上,每种算法被连续运行 20 次.表 4 显示了,在最大迭代次数内,PSO_w,PSO_cf_local,FIPS,FDR_PSO,CLPSO和InformPSO分别在 20 次优化每个多峰函数时所得到的最差适应度值和最优适应度值其中,适应度值以函数值表示,PSO_cf_local,FIPS的数据来自于文献[9],PSO_w,FDR_PSO,CLPSO和InformPSO的数据来自于本文的实验.

Table 4 Experimental results for multimodal functions

表 4 多峰函数实验结果

Function	Best/Worst fitness values achieved					
	PSO_w	PSO_cf_local	FDR_PSO	FIPS	CLPSO	InformPSO
f_6	0.113 2	2.80e-002	0.029 5	9.32e-002	6.7077e-008	0
	0.155 0	6.34e-002	0.253 3	1.31e-001	0.007 4	0.027 0
f_7	3.5527e-015	5.78e-002	7.1054e-015	3.75e-015	7.1054e-015	8.8818e-016
	3.5527e-015	2.58e-001	7.1054e-015	2.13e-014	3.5527e-015	3.5527e-015
f_8	4.974 8	3.48e+000	3.979 8	1.33e+000	0	0
	10.944 5	9.05e+000	8.954 6	2.12e+000	0.995 0	1.1955e-008
f_9	0	5.16e-002	7.1054e-015	2.02e-003	0	0
	0.001 1	7.85e-002	2.4154e-004	6.40e-003	0	0
f_{10}	475.270 5	2.93e+002	710.630 1	7.10e+001	1.2728e-004	1.2728e-004
	592.191 8	8.78e+002	1.1054e+003	1.50e+002	118.438 5	118.438 5

表 4 的数据进一步表明了,在最大迭代次数内,InformPSO对于上述 5 个多峰函数的优化均取得最好的求解精度.在 f_6, f_8, f_9 的函数优化的各自连续 20 次实验中,除了 f_6 两次收敛于局部极值点以外,新算法大多能收敛于全局最优解 0;而Ackley函数能快速收敛于精度较高的局部最优解 8.8818e-016.实验结果表明,在上述高维多峰基准函数测试中,PSO_w对于 f_9 函数、CLPSO对于 f_6, f_8, f_9 函数、InformPSO对于 f_6, f_8, f_9 函数能找到全局最优解,有效抑制了早熟收敛现象.可以看出,InformPSO的求解精度、收敛速度和求解的稳定性明显优于新提出的PSO变种.

3 总结及工作方向

本文借鉴生物群体中信息扩散的习性,分析了微粒速度和个体速度的关系;模拟基因进化和大漂移突变规

律,给出了一种解决数值优化问题的改进 PSO 算法——自适应扩散混合变异机制微粒群算法,并从理论上证明了它以概率 1 收敛至全局最优解.一方面,InformPSO 通过信息扩散函数自适应控制微粒飞向群体最佳微粒 gBest 的速度,有效提高了种群的多样性;另一方面,克隆选择操作使得 gBest 能够有效地微变寻优,而非线性时间序列指导 gBest 随机大漂移亦能增强算法逃离局部极值能力.两种变异机制相互协作引导 gBest 有益变异.实验结果显示,新算法在所述单峰和多峰基准函数优化中,显示出全局搜索能力强、求解精度高、收敛速度快、稳定好等特点,测试性能优于现存 PSO 改进算法.

但是,在 Schwefel 和 Ackley 实验中,我们发现新算法虽然比其他 PSO 性能更佳,但也只能收敛至某一局部极值.这主要是由于这两个函数局部最优点距离全局最优点非常远,新算法的克隆选择算子是微突变,长期效用微弱;而纯粹的非线性时间序列指导 gBest 变异更多的是具有随机性特点,因此对于此类函数优化,不足以跳出局部极值.进一步的工作方向主要有:扩大算法的测试范围,并寻求微粒间相互协作的方式,进一步改善算法收敛性能,并将其应用到大规模的图像恢复问题.

致谢 在此,我们感谢对本文的工作提供有价值评论的匿名审稿人,同时也感谢 Suganthan 教授向我们提供了 CLPSO 的源码.

References:

- [1] Kennedy J, Eberhart RC. Particle swarm optimization. In: Proc. of the IEEE Int'l Conf. on Neural Network. Perth: IEEE Inc., 1995. 1942–1948.
- [2] Shi Y, Eberhart R. A modified particle swarm optimizer. In: Proc. of the CEC'98. Alaska: IEEE Inc., 1998. 69–73.
- [3] Clerc M, Kennedy J. The particle swarm-explosion, stability, and convergence in a multidimensional complex space. IEEE Trans. on Evolutionary Computation, 2002,6(1):58–73.
- [4] Kennedy J, Mendes R. Population structure and particle swarm performance. In: Proc. of the CEC 2002. Honolulu: IEEE Inc., 2002. 1671–1676.
- [5] Parsopoulos KE, Vrahatis MN. UPSO: A unified particle swarm optimization scheme. In: Proc. of the ICCMSE 2004. Attica: VSP Int'l Science Publishers, 2004. 868–873.
- [6] Mendes R, Kennedy J, Neves J. The fully informed particle swarm: Simpler, maybe better. IEEE Trans. on Evolutionary Computation, 2004,8(3):204–210.
- [7] Peram T, Veeramachaneni K. Fitness-Distance-Ratio based particle swarm optimization. In: Proc. of the IEEE Swarm Intelligence Symp. Indianapolis: IEEE Inc., 2003. 174–181.
- [8] Bergh F, Engelbrecht AP. A cooperative approach to particle swarm optimization. IEEE Trans. on Evolutionary Computation, 2004, 8(3):225–239.
- [9] Liang JJ, Qin AK. Comprehensive learning particle swarm optimizer for global optimization of multimodal functions. IEEE Trans. on Evolutionary Computation, 2006,10(3):281–295.
- [10] He R, Wang YJ, Wang Q, Zhou JH, Hu CY. An improved particle swarm optimization based on self-adaptive escape velocity. Journal of Software, 2005,16(12):2036–2044 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/16/2036.htm>
- [11] Castro LN, Zuben FJ. The clonal selection algorithm with engineering applications. In: Whitley LD, Goldberg DE, eds. Proc. of the GECCO 2000. Las Vegas: ACM Press, 2000. 36–37.
- [12] Choi C, Lee J. Chaotic local search algorithm. Artificial Life & Robotics, 1998,2(1):41–47.
- [13] Guo CH, Tang HW. Global convergence properties of evolution strategies. Mathematica Numerica Sinica, 2001,23(1):105–110 (in Chinese with English abstract).

附中文参考文献:

