

粒子群优化方法若干特性分析

潘峰¹ 陈杰¹ 辛斌¹ 张娟¹

摘要 粒子群优化算法 (Particle swarm optimizer, PSO) 是一种基于群体智能的优化方法. 本文提出了标准粒子群优化方法按迭代时间展开的一般性描述公式. 在此基础上分析了标准 PSO 的优化机理, 基于群体社会信息和自身历史经验的情况下, 推导了粒子最大搜索空间的数学描述. 通过将粒子运动的一般性描述图解为历史状态加权求和的形式, 进一步证明了 PSO 参数随迭代周期的积累, 在概率意义上的遗忘特性. 分析表明在经过一定周期的搜索后, 标准 PSO 方法同 Barebones 粒子群方法 (Barebones particle swarm, BBPS) 具有近似的搜索机制. 从信息传递的角度, PSO 的搜索策略是一种在概率意义上具有遗忘特性的历史信息加权求和的结果. 本文的研究结果对标准粒子群算法的一些重要性质 (如: 遗忘特性、标准 PSO 与 BBPS 间的相似性等) 进行了合理解释.

关键词 粒子群优化方法, 最大搜索空间, 遗忘特性, 相似性
中图分类号 TP18

Several Characteristics Analysis of Particle Swarm Optimizer

PAN Feng¹ CHEN Jie¹ XIN Bin¹ ZHANG Juan¹

Abstract Particle swarm optimizer (PSO), a swarm intelligence based optimization technique, is described by a general formula in terms of iterations in the paper. Based on the general formula, its optimization mechanism is analyzed and the general mathematic description of particle's maximum covering space is deduced according to the current social information and personal experience. Furthermore, the general formula is illustrated as the weighted summation of historical position states, so as to prove that in terms of cumulative iterations, parameters of PSO have an inherent forgetting characteristic in probability, moreover the searching mechanisms of canonical PSO and Barebones particle swarm are almost the same. From the perspective of information propagation, the strategy of PSO is a weighted summation of the historical information, which has the forgetting characteristic in probability. Some important properties of canonical PSO, such as forgetting characteristic, similarity between canonical PSO and BBPS, etc, are explained by the results of the research in this paper.

Key words Particle swarm optimizer (PSO), maximum covering space, forgetting characteristic, similarity

自从 1995 年 Kennedy 和 Eberhart^[1] 提出粒子群优化 (Particle swarm optimizer, PSO) 方法以来, 经过十多年的研究表明, PSO 是一种有效的基于群智能和交互智能的并行计算方法. Riccardo^[2] 对 IEEE Xplore 数据库的统计分析表明, 截止至 2007 年在超过 1000 篇关于 PSO 的研究文献中, 有近 2/3 的文献专注于对 PSO 的应用, 在优化计算、生

物医学、通讯网络、控制、模式分类等方面都体现出了良好的性能, 另外 1/3 的文献对 PSO 的改进和分析进行了研究. 尤其是增加了对 PSO 的参数选择、稳定性、收敛性和统计特性等方面的研究.

许多学者分别从不同的角度对 PSO 算法进行了研究, Shi 等^[3] 通过仿真试验的统计结果, 对比了不同参数取值对标准 PSO 优化性能的影响; Bergh 等^[4-5] 证明了尽管标准 PSO 非常高效, 但并不是一种全局优化的方法, 也不能保证一定收敛到某个局部最优点, 并提出了一些有效的改进方法; Clerc 和 Kennedy^[6] 建立了由 5 个参数描述的 PSO 约束模型, 分析了其收敛性和相平面内粒子运动的轨迹特性; Trelea^[7] 运用动态系统理论进行了 PSO 的参数选取条件和稳定分析; Clerc 和 Trelea 都将 PSO 算法作为线性时不变动态系统进行了参数稳定性分析, 但是 PSO 粒子运动的动态方程, 并不是一个线性时不变系统, 因此基于时不变动态系统的分析结果具有一定的局限性; Chen 等^[8] 将 PSO 作为动态时变系统分析了其稳定的充分条件, 将惯性因子 w 值域扩展到 $(-1, 1)$; Kadiramanathan 等^[9] 考虑了 PSO 中参数的随机性, 并基于李亚普诺夫稳定性原理分析了 PSO 稳定的充分条件, 他将陷入停滞后的群体最优粒子描述为一个非线性反馈系统进行分析, 但是由于李亚普诺夫定理的保守性, 该条件对参数的选取非常严格, 使得粒子只能维持很小的局部振动, 从而影响了寻优的性能; Jiang 等^[10] 研究标准 PSO 在不同参数选择条件下对算法随机收敛性的影响; 金欣磊等^[11] 从随机过程的角度出发, 把带随机因子的 PSO 动力学模型在概率意义下转换成线性定常系统加以分析研究, 给出了系统依概率收敛的一个充分条件.

对于计算智能方法而言, 早熟收敛是影响算法性能和寻优效果的主要问题之一, 文献 [12-15] 分析了当群体停滞于解空间中某一局部极值点时算法的搜索性能. Clerc^[12] 分析了在经典 PSO 的参数设置下速度向量分别由“前向力”、“后向力”和随机噪声共同决定, 通过不同的参数设置和概率分布函数, 可以在一定程度上调节“向前”和“向后”的搜索进程. 当群体停滞在局部极值点时, 具有导向性作用的群体最优粒子只能让 PSO 具有有限的“向前”搜索的能力^[13]. 在没有新的信息引入的情况下, 对 gbest 和 pbest 粒子模型的运动特性和能达到的最大搜索空间的分析表明, 典型的参数选取使得群体倾向于搜索已经搜索过的空间.

2003 年, Kennedy^[16] 提出了一种简化的粒子群优化方法: Barebones particle swarm (BBPS). 相对于经典形式的 PSO, BBPS 减少了速度项, 并且采用了基于群体最优 (gbest 或 lbest) 和个体历史最优 (pbest) 信息的高斯采样在解空间中进行搜索. Riccardo 等^[17] 将 N 维连续解空间分解为有限的栅格, 然后分析了 BBPS 分别采用柯西分布、高斯分布以及其他分布时的离散马尔科夫链. 并指出: 采样方法的微小变化可以在很大程度上帮助 BBPS 在解空间中开拓搜索新的空间, 甚至将 BBPS 转化为一种全局优化方法. Omran 将 BBPS 应用于整数规划问题^[18] 和无监督图像分类^[19] 的优化设计中, 取得了良好的效果.

总体而言, 对粒子群优化方法的研究主要以动态系统或者随机过程理论为出发点, 对算法的参数选取、稳定性分析及全局收敛性进行分析. 本文将粒子群优化方法作为离散随机系统展开研究, 在第 2 节推导了单信息最大搜索空间的一般性迭代描述, 在此一般性描述及其随机特性分析的基础上, 在第 3 节和第 4 节分别讨论了标准粒子群优化方法与 Barebones 粒子群算法的相似性和遗忘特性.

收稿日期 2008-08-05 收修稿日期 2009-01-05
Received August 5, 2008; in revised form January 5, 2009

高等学校优秀青年教师教学科研奖励计划 (20010248) 资助
Supported by the Teaching and Research Award Program for Outstanding Young Teachers in Higher Education Institutions of Ministry of Education of China (20010248)

1. 北京理工大学复杂系统智能控制与决策教育部重点实验室 北京 100081
1. Key Laboratory of Complex System Intelligent Control and Decision (Beijing Institute of Technology), Ministry of Education, Beijing 100081

DOI: 10.3724/SP.J.1004.2009.01010

1 标准粒子群优化方法

在任意 $k+1$ 时刻, 第 i 个粒子第 d 维的更新公式定义为

$$\mathbf{v}_{id}^{(k+1)} = w\mathbf{v}_{id}^{(k)} + \varphi_{ld}^{(k+1)}(\mathbf{p}_{ld}^{(k)} - \mathbf{x}_{id}^{(k)}) + \varphi_{id}^{(k+1)}(\mathbf{p}_{id}^{(k)} - \mathbf{x}_{id}^{(k)}) \quad (1)$$

$$\mathbf{x}_{id}^{(k+1)} = \mathbf{x}_{id}^{(k)} + \mathbf{v}_{id}^{(k+1)} \quad (2)$$

其中, $\mathbf{x}_{id}^{(k)}$ 是当前时刻粒子在解空间中的位置向量; 速度分量 $\mathbf{v}_{id}^{(k)}$ 满足约束 $|\mathbf{v}_{id}^{(k)}| \leq V_{\max}$; $\mathbf{p}_{id}^{(k)}$ 是个体历史最优位置向量; $\mathbf{p}_{ld}^{(k)}$ 代表的是邻域结构中的最优位置; w 为惯量因子; $\varphi_{ld}^{(k+1)} = c_1 \text{rand}_{ld}$ 和 $\varphi_{id}^{(k+1)} = c_2 \text{rand}_{id}$ 是符合一定概率分布的随机数, 一般而言, 在标准 PSO 中, rand_{ld} 和 rand_{id} 是 $[0, 1]$ 内满足均匀分布的随机数, 加速度因子 c_1 、 c_2 的取值反映了粒子个体对来自于个体认知信息和群体社会信息的不同侧重。

Shi 等^[3] 对 V_{\max} 在 $[2, X_{\max}]$ 之间单独取值, 且速度惯量 w 在 $[0.1, 1.05]$ 范围内取值时, PSO 系统的优化性能进行了试验分析; Clerc 等^[6] 建议 $(c_1 + c_2)/2$ 的取值为 1.494, 惯量因子 w 取值为 0.729. 群体的邻域结构、惯量因子和加速度因子等可调参数决定着算法的收敛性和优化性能, 以及群体和粒子的运动特性。

2 标准 PSO 单信息最大搜索空间描述

在每一个时刻 k , 粒子在群体信息和自身历史经验的引导下, 在一定范围的空间内进行搜索, 当发现更优的解时, 对引导信息进行更新, 展开新的搜索, 寻优过程不断向前推进. 这样一个递进式搜索过程能够得到不断推进的前提是在粒子运动过程中, 从概率意义上所能覆盖的空间中存在比当前信息更优的解. 否则, 算法将面临停滞于局部的困境. 首先, 引入如下定义.

定义 1. 单信息最大搜索空间是在当前的邻域最优信息和个体最优信息条件下, 在没有新的信息获取时, 粒子从当前运动状态开始, 可以达到的最大搜索空间.

文献 [13] 中推导了在没有新的信息引入的情况下两种特殊情况 (gbest 和 pbest 粒子模型) 的单信息最大搜索空间, 结果表明典型选取的参数使得群体倾向于搜索已经搜索过的空间, 这种重复搜索会降低搜索效率, 产生额外的计算代价. 引理 1 给出了一般意义下, 标准 PSO 任意粒子在任意时刻的运动特性和能达到的最大的搜索空间.

引理 1. 标准 PSO 粒子的单信息最大搜索空间为

$$\begin{aligned} \mathbf{x}_{id}^{(k+n)} = & \left\{ \prod_{m=1}^n \gamma(m) \right\} \mathbf{x}_{id}^{(k)} + \left\{ w^n + \sum_{j=1}^{n-1} \left(w^j \prod_{m=j+1}^n \gamma(m) \right) \right\} \times \\ & \mathbf{v}_{id}^{(k)} + \left\{ \varphi_{ld}^{(k+n)} + \sum_{j=1}^{n-1} \left(\varphi_{ld}^{(k+j)} \prod_{m=j+1}^n \gamma(m) \right) \right\} \mathbf{p}_{ld}^{(k)} + \\ & \left\{ \varphi_{id}^{(k+n)} + \sum_{j=1}^{n-1} \left(\varphi_{id}^{(k+j)} \prod_{m=j+1}^n \gamma(m) \right) \right\} \mathbf{p}_{id}^{(k)} \quad (3) \end{aligned}$$

其中, $\gamma(m) = 1 - \varphi_{ld}^{(k+m)} - \varphi_{id}^{(k+m)}$, $m = 1, 2, \dots, n$. 变量带括号的上标或下标表示迭代周期, 如 $\mathbf{x}_{id}^{(k+n)}$ 表示第 $k+n$ 个周期; 变量不带括号的上标表示幂, 如 w^n 表示 w 的 n 次方.

证明. 将式 (1) 和 (2) 整理为

$$\mathbf{v}_{id}^{(k+1)} = w\mathbf{v}_{id}^{(k)} \quad (4)$$

$$\mathbf{x}_{id}^{(k+1)} = \mathbf{x}_{id}^{(k)} + \varphi_{ld}^{(k+1)}(\mathbf{p}_{ld}^{(k)} - \mathbf{x}_{id}^{(k)}) + \varphi_{id}^{(k+1)}(\mathbf{p}_{id}^{(k)} - \mathbf{x}_{id}^{(k)}) + \mathbf{v}_{id}^{(k+1)} \quad (5)$$

当 $n = 1$ 时,

$$\mathbf{x}_{id}^{(k+1)} = \gamma(1)\mathbf{x}_{id}^{(k)} + w\mathbf{v}_{id}^{(k)} + \varphi_{ld}^{(k+1)}\mathbf{p}_{ld}^{(k)} + \varphi_{id}^{(k+1)}\mathbf{p}_{id}^{(k)}$$

当 $n = 2$ 时,

$$\mathbf{x}_{id}^{(k+2)} = \gamma(2)\mathbf{x}_{id}^{(k+1)} + w^2\mathbf{v}_{id}^{(k)} + \varphi_{ld}^{(k+2)}\mathbf{p}_{ld}^{(k)} + \varphi_{id}^{(k+2)}\mathbf{p}_{id}^{(k)}$$

$$\mathbf{x}_{id}^{(k+2)} = \gamma(2)\gamma(1)\mathbf{x}_{id}^{(k)} + \gamma(2)w\mathbf{v}_{id}^{(k)} + w^2\mathbf{v}_{id}^{(k)} +$$

$$\gamma(2)\varphi_{ld}^{(k+1)}\mathbf{p}_{ld}^{(k)} + \varphi_{ld}^{(k+2)}\mathbf{p}_{ld}^{(k)} +$$

$$\gamma(2)\varphi_{id}^{(k+1)}\mathbf{p}_{id}^{(k)} + \varphi_{id}^{(k+2)}\mathbf{p}_{id}^{(k)}$$

$$\mathbf{x}_{id}^{(k+2)} = \left\{ \prod_{m=1}^2 \gamma(m) \right\} \mathbf{x}_{id}^{(k)} + \left\{ w^2 + \sum_{j=1}^1 \left(w^j \prod_{m=j+1}^2 \gamma(m) \right) \right\} \mathbf{v}_{id}^{(k)} +$$

$$\left\{ \varphi_{ld}^{(k+2)} + \sum_{j=1}^1 \left(\varphi_{ld}^{(k+j)} \prod_{m=j+1}^2 \gamma(m) \right) \right\} \mathbf{p}_{ld}^{(k)} +$$

$$\left\{ \varphi_{id}^{(k+2)} + \sum_{j=1}^1 \left(\varphi_{id}^{(k+j)} \prod_{m=j+1}^2 \gamma(m) \right) \right\} \mathbf{p}_{id}^{(k)}$$

假设当 $n = n_0$ 时, 式 (3) 成立, 则有

$$\mathbf{x}_{id}^{(k+n_0)} = \left\{ \varphi_{ld}^{(k+n_0)} + \sum_{j=1}^{n_0-1} \left(\varphi_{ld}^{(k+j)} \prod_{m=j+1}^{n_0} \gamma(m) \right) \right\} \mathbf{p}_{ld}^{(k)} +$$

$$\left\{ \varphi_{id}^{(k+n_0)} + \sum_{j=1}^{n_0-1} \left(\varphi_{id}^{(k+j)} \prod_{m=j+1}^{n_0} \gamma(m) \right) \right\} \mathbf{p}_{id}^{(k)} +$$

$$\left\{ \prod_{m=1}^{n_0} \gamma(m) \right\} \mathbf{x}_{id}^{(k)} + \left\{ w^{n_0} + \sum_{j=1}^{n_0-1} \left(w^j \prod_{m=j+1}^{n_0} \gamma(m) \right) \right\} \mathbf{v}_{id}^{(k)}$$

则当 $n = n_0 + 1$ 时,

$$\mathbf{x}_{id}^{(k+n_0+1)} = \mathbf{x}_{id}^{(k+n_0)} + \varphi_{ld}^{(k+n_0+1)}(\mathbf{p}_{ld}^{(k)} - \mathbf{x}_{id}^{(k+n_0)}) + \varphi_{id}^{(k+n_0+1)}(\mathbf{p}_{id}^{(k)} - \mathbf{x}_{id}^{(k+n_0)}) + \mathbf{v}_{id}^{(k+n_0+1)}$$

$$\mathbf{x}_{id}^{(k+n_0+1)} = \gamma(n_0+1)\mathbf{x}_{id}^{(k+n_0)} + w^2\mathbf{v}_{id}^{(k)} +$$

$$\varphi_{ld}^{(k+n_0+1)}\mathbf{p}_{ld}^{(k)} + \varphi_{id}^{(k+n_0+1)}\mathbf{p}_{id}^{(k)}$$

上式可整理为式 (6) 的形式, 从而推导得到 $n = n_0 + 1$ 时的描述为

$$\mathbf{x}_{id}^{(k+n_0+1)} = \left\{ \varphi_{ld}^{(k+n_0+1)} + \sum_{j=1}^{n_0} \left(\varphi_{ld}^{(k+j)} \prod_{m=j+1}^{n_0+1} \gamma(m) \right) \right\} \mathbf{p}_{ld}^{(k)} +$$

$$\left\{ \varphi_{id}^{(k+n_0+1)} + \sum_{j=1}^{n_0} \left(\varphi_{id}^{(k+j)} \prod_{m=j+1}^{n_0+1} \gamma(m) \right) \right\} \mathbf{p}_{id}^{(k)} +$$

$$\left\{ \prod_{m=1}^{n_0+1} \gamma(m) \right\} \mathbf{x}_{id}^{(k)} + \left\{ w^{n_0+1} + \sum_{j=1}^{n_0} \left(w^j \prod_{m=j+1}^{n_0+1} \gamma(m) \right) \right\} \times$$

$$\mathbf{v}_{id}^{(k)}$$

当 $n = n_0 + 1$ 时, 式 (3) 成立. \square

$$\begin{aligned} \mathbf{x}_{id}^{(k+n_0+1)} &= \gamma_{(n_0+1)} \left\{ \prod_{m=1}^{n_0} \gamma_{(m)} \right\} \mathbf{p}_{id}^{(k)} + \gamma_{(n_0+1)} \left\{ w^{n_0} + \sum_{j=1}^{n_0-1} \left(w^j \prod_{m=j+1}^{n_0} \gamma_{(m)} \right) \right\} \mathbf{v}_{id}^{(k)} + \\ &\gamma_{(n_0+1)} \left\{ \varphi_{id}^{(k+n_0)} + \sum_{j=1}^{n_0-1} \left(\varphi_{id}^{(k+j)} \prod_{m=j+1}^{n_0} \gamma_{(m)} \right) \right\} \mathbf{p}_{id}^{(k)} + \varphi_{id}^{(k+n_0+1)} \mathbf{p}_{id}^{(k)} + \\ &\gamma_{(n_0+1)} \left\{ \varphi_{id}^{(k+n_0)} + \sum_{j=1}^{n_0-1} \left(\varphi_{id}^{(k+j)} \prod_{m=j+1}^{n_0} \gamma_{(m)} \right) \right\} \mathbf{p}_{id}^{(k)} + \varphi_{id}^{(k+n_0+1)} \mathbf{p}_{id}^{(k)} \end{aligned} \quad (6)$$

3 标准 PSO 与 Barebones PS 相似性分析

3.1 单信息最大搜索空间的描述分析

本节将对信息最大搜索空间的一般性描述 (式 (3)) 进行推导和整理, 从而得到式 (8) 的形式, 再分别对其个系数项进行分析. 为表述方便, 在以下推导中, 首先只对 $\mathbf{p}_{id}^{(k)}$ 和 $\mathbf{p}_{id}^{(k)}$ 的参数进行分析.

$$\hat{\mathbf{x}}_{id}^{(k+n)} = \left\{ \varphi_{id}^{(k+n)} + \sum_{j=1}^{n-1} \left(\varphi_{id}^{(k+j)} \prod_{m=j+1}^n \gamma_{(m)} \right) \right\} \mathbf{p}_{id}^{(k)} \times \left\{ \varphi_{id}^{(k+n)} + \sum_{j=1}^{n-1} \left(\varphi_{id}^{(k+j)} \prod_{m=j+1}^n \gamma_{(m)} \right) \right\} \mathbf{p}_{id}^{(k)} \quad (7)$$

$$\hat{\mathbf{x}}_{id}^{(k+n)} = \left\{ \varphi_{id}^{(k+n)} + \varphi_{id}^{(k+n)} + \sum_{j=1}^{n-1} \left(\prod_{m=j+1}^n \gamma_{(m)} \right) - \sum_{j=1}^{n-1} \left(\gamma_j \prod_{m=j+1}^n \gamma_{(m)} \right) \right\} \mathbf{p}_{id}^{(k)} + \left\{ \varphi_{id}^{(k+n)} + \sum_{j=1}^{n-1} \left(\varphi_{id}^{(k+j)} \prod_{m=j+1}^n \gamma_{(m)} \right) \right\} (\mathbf{p}_{id}^{(k)} - \mathbf{p}_{id}^{(k)})$$

$$\hat{\mathbf{x}}_{id}^{(k+n)} = \left\{ \varphi_{id}^{(k+n)} + \varphi_{id}^{(k+n)} + \sum_{j=1}^{n-1} \left(\prod_{m=j+1}^n \gamma_{(m)} \right) - \sum_{j=1}^{n-1} \left(\prod_{m=j}^n \gamma_{(m)} \right) \right\} \mathbf{p}_{id}^{(k)} + \left\{ \varphi_{id}^{(k+n)} + \sum_{j=1}^{n-1} \left(\varphi_{id}^{(k+j)} \prod_{m=j+1}^n \gamma_{(m)} \right) \right\} (\mathbf{p}_{id}^{(k)} - \mathbf{p}_{id}^{(k)})$$

上式中 $\mathbf{p}_{id}^{(k)}$ 的参数累加项可以按周期 j 展开, 如图 1 所示. 由图可见, 两个累加项相隔一个周期的对应项相互抵消, 最终两个累加项分别保留下 $j = 1$ 和 $j = n - 1$ 时刻对应的一项参数.

图 1 中两参数各项抵消后, $\hat{\mathbf{x}}_{id}^{(k+n)}$ 可以得到下式更直接的描述: 以群体最优向量为中心, 群体最优向量与个体历史最优向量之间的距离为半径的空间内、符合一定概率分布的随机搜索.

根据以上推导, 对式 (3) 中的所有项进行综合表述:

$$\hat{\mathbf{x}}_{id}^{(k+n)} = \mathbf{p}_{id}^{(k)} - \left(\prod_{m=1}^n \gamma_{(m)} \right) \mathbf{p}_{id}^{(k)} + \left\{ \varphi_{id}^{(k+n)} + \sum_{j=1}^{n-1} \left(\varphi_{id}^{(k+j)} \prod_{m=j+1}^n \gamma_{(m)} \right) \right\} (\mathbf{p}_{id}^{(k)} - \mathbf{p}_{id}^{(k)})$$

$\mathbf{x}_{id}^{(k+n)}$ 可以简化为

$$\mathbf{x}_{id}^{(k+n)} = \mathbf{p}_{id}^{(k)} + \phi^{(n)} (\mathbf{p}_{id}^{(k)} - \mathbf{p}_{id}^{(k)}) - Y_{(n)} \mathbf{p}_{id}^{(k)} + Y_{(n)} \mathbf{x}_{id}^{(k)} + \delta^{(n)} \mathbf{v}_{id}^{(k)} \quad (8)$$

其中, $Y_{(n)} = \prod_{m=1}^n \gamma_{(m)}$ 为 $\mathbf{p}_{id}^{(k)}$ 和 $\mathbf{x}_{id}^{(k)}$ 的阶乘衰减因子; $\phi^{(n)} = \{\varphi_{id}^{(k+n)} + \sum_{j=1}^{n-1} (\varphi_{id}^{(k+j)} \prod_{m=j+1}^n \gamma_{(m)})\}$ 为 $\mathbf{p}_{id}^{(k)}$ 和 $\mathbf{p}_{id}^{(k)}$ 之间距离的随机权重; $\delta^{(n)} = \{w^n + \sum_{j=1}^{n-1} (w^j \prod_{m=j+1}^n \gamma_{(m)})\}$ 为初始时刻速度向量的加权参数. 下面将分别对这三个参数进行讨论, 并通过分析表明, 当经过一段时间的搜索后, 即 n 足够大时, $\mathbf{x}_{id}^{(k+n)}$ 等价于 $\hat{\mathbf{x}}_{id}^{(k+n)}$.

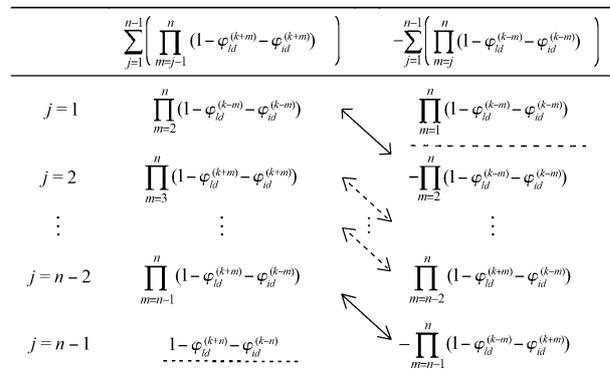


图 1 $\mathbf{p}_{id}^{(k)}$ 参数累加项展开对比
Fig.1 Parameters comparison of $\mathbf{p}_{id}^{(k)}$

3.2 初始位置向量阶乘衰减因子分析

阶乘衰减因子 $Y_{(n)}$ 分别反映了粒子在初始时刻位置信息 $\mathbf{x}_{id}^{(k)}$ 和部分群体最优信息 $\mathbf{p}_{id}^{(k)}$ 随时间变化的特性. 不失一般性, 根据式 (1) 和 (2) 的定义, $\varphi_{id}^{(k)}$ 、 $\varphi_{id}^{(k)}$ 分别为 r_l 、 r_i 间的均匀分布, 即 $\varphi_{id}^{(k)} \sim (0, r_l)$ 、 $\varphi_{id}^{(k)} \sim (0, r_i)$ 且 $\varphi_{id}^{(k)}$ 与 $\varphi_{id}^{(k)}$ 互相独立, 则有

$$E[\gamma_{(m)}] = 1 - E[\varphi_{id}^{(k+m)}] - E[\varphi_{id}^{(k-m)}] = 1 - \frac{r_l + r_i}{2}$$

$$D[\gamma_{(m)}] = D[\varphi_{id}^{(k+m)}] + D[\varphi_{id}^{(k-m)}] = \frac{r_l^2 + r_i^2}{12}$$

$$E[\gamma_{(m)}^2] = D[\gamma_{(m)}] + (E[\gamma_{(m)}])^2 =$$

$$1 + \frac{r_l^2 + r_i^2}{3} + \frac{r_l r_i}{2} - r_l - r_i$$

$\gamma_{(j)}$ 、 $\gamma_{(j+1)}$ 、 \dots 、 $\gamma_{(n)}$ 之间相互统计独立, 则概率密度函数为

$$f(\gamma_{(j)}, \dots, \gamma_{(n)}) = f(\gamma_{(j)}) f(\gamma_{(j+1)}) \dots f(\gamma_{(n)})$$

$$\begin{aligned} E \left[\prod_{m=j}^n \gamma(m) \right] &= E \left[\prod_{m=j}^n \gamma(m) \right] = \prod_{m=j}^n E[\gamma(m)] = \\ &= \left(1 - \frac{r_l + r_i}{2}\right)^{n-j} \end{aligned}$$

$$\begin{aligned} E \left[\left(\prod_{m=j}^n \gamma(m) \right)^2 \right] &= E \left[\prod_{m=j}^n \gamma(m)^2 \right] = \prod_{m=j}^n E[\gamma(m)^2] = \\ &= \left(1 + \frac{r_l^2 + r_i^2}{3} + \frac{r_l r_i}{2} - r_l - r_i\right)^{n-j} \end{aligned}$$

$$\begin{aligned} D \left[\prod_{m=j}^n \gamma(m) \right] &= E \left[\left(\prod_{m=j}^n \gamma(m) \right)^2 \right] - \left(E \left[\prod_{m=j}^n \gamma(m) \right] \right)^2 = \\ &= \left(1 + \frac{r_l^2 + r_i^2}{3} + \frac{r_l r_i}{2} - r_l - r_i\right)^{n-j} - \\ &= \left(1 - \frac{r_l + r_i}{2}\right)^{2(n-j)} \end{aligned} \quad (10)$$

根据对社会认知与个体经验信息的不同侧重, r_g 和 r_i 可以选取不同的数值. 以典型的取值设置为例, $r_g = r_i = r$ 体现了对两种信息的共同关注. 此时, 式 (9) 和 (10) 可以简化为

$$E \left[\prod_{m=j}^n \gamma(m) \right] = (1-r)^{n-j} \quad (11)$$

$$D \left[\prod_{m=j}^n \gamma(m) \right] = \left(1 - 2r + \frac{7r^2}{6}\right)^{n-j} - (1-r)^{2(n-j)} \quad (12)$$

为保证位置更新公式得到的位置向量取值有界, 由式 (11) 和 (12) 描述的 $Y(n)$ 的均值与方差都要求有界. 因此, 上式中的均值需要不大于 1, 即 $0 \leq r \leq 2$; 而且方差也不能大于 1, 即 $|1 - 2r + 7r^2/6| \leq 1$, 所以 $0 \leq r \leq 1.7143$. 在满足以上条件时, 对于任意小的非负 ε , 都存在 N

$$\lim_{n \geq N} E \left[\prod_{m=j}^n \gamma(m) \right] = \lim_{n \geq N} (1-r)^{n-j} < \varepsilon \quad (13)$$

$$\lim_{n \geq N} D \left[\prod_{m=j}^n \gamma(m) \right] = \lim_{n \geq N} \left(1 - 2r + \frac{7r^2}{6}\right)^{n-j} - (1-r)^{2(n-j)} < \varepsilon \quad (14)$$

3.3 初始速度向量的加权参数分析

与阶乘衰减因子 $Y(n)$ 推导过程同理, 根据式 (11) 和 (12) 可以类推得到初始速度向量的加权参数 $\delta^{(n)}$ 的数学期望.

$$E \left[w^j \prod_{m=j+1}^n \gamma(m) \right] = w^j (1-r)^{n-j-1}$$

$$\begin{aligned} E[\delta^{(n)}] &= w^n + \sum_{j=1}^{n-1} E \left[\left(w^j \prod_{m=j+1}^n \gamma(m) \right) \right] = \\ &= w^n + \sum_{j=1}^{n-1} w^j (1-r)^{n-j} = \sum_{j=1}^n w^j (1-r)^{n-j} = \\ &= \frac{w^n - (1-r)^n}{w + r - 1} \end{aligned}$$

对于标准 PSO 算法的典型参数设置, 速度惯量因子 w 取值一般为小于 1 的常数. 因此有 $\forall \varepsilon > 0, \exists N$

$$\lim_{n \geq N} E[\delta^{(n)}] = \lim_{n \geq N} w \frac{w^n - (1-r)^n}{w + r - 1} < \varepsilon \quad (15)$$

该结果表明, 初始速度向量 $\mathbf{v}_{id}^{(k)}$ 的加权参数随着迭代周期的增加, 将逐渐收敛到均值为 0 的一个区域.

3.4 标准 PSO 与 Barebones PS 的相似性讨论

根据式 (13)~(15) 的分析结果, 当粒子个体基于 $\mathbf{p}_{ld}^{(k)}$ 和 $\mathbf{x}_{id}^{(k)}$ 进行了一定周期的搜索后, 由式 (8) 描述的 PSO 更新公式可以简化为

$$\mathbf{x}_{id}^{(k+n)} = \mathbf{p}_{ld}^{(k)} + \phi^{(n)}(\mathbf{x}_{id}^{(k)} - \mathbf{p}_{ld}^{(k)}) \quad (16)$$

根据中心极限定理, 在相当一般的条件下, 当独立随机变量的个数增加时, 其和的分布趋于正态分布. 当 n 足够大时, $\mathbf{p}_{ld}^{(k)}$ 和 $\mathbf{x}_{id}^{(k)}$ 之间距离的随机权重 $\phi^{(n)}$ 趋于正态分布.

Barebones particle swarm (BBPS)^[16] 是基于均值为 $(\mathbf{p}_{ld}^{(k)} + \mathbf{x}_{id}^{(k)})/2$, 方差为 $|\mathbf{p}_{ld}^{(k)} - \mathbf{x}_{id}^{(k)}|$ 的正态分布的随机采样搜索, 如式 (17) 定义.

$$\mathbf{x}_{id}^{(k+1)} = N \left(\frac{\mathbf{p}_{ld}^{(k)} + \mathbf{x}_{id}^{(k)}}{2}, |\mathbf{p}_{ld}^{(k)} - \mathbf{x}_{id}^{(k)}| \right) \quad (17)$$

BBPS 同样可以由下式描述

$$\mathbf{x}_{id}^{(k+1)} = \mathbf{p}_{ld}^{(k)} + N \left(\frac{\mathbf{x}_{id}^{(k)} - \mathbf{p}_{ld}^{(k)}}{2}, |\mathbf{p}_{ld}^{(k)} - \mathbf{x}_{id}^{(k)}| \right) \quad (18)$$

Kennedy 曾经指出^[16], “BBPS 看起来并不像粒子群. 毕竟我们常常定义的是粒子的飞翔, 而 BBPS 更像是鸟群在叽叽喳喳, 而不是飞翔”. 虽然从数学定义上, BBPS 与经典 PSO 的数学描述不尽相同, 而通过对式 (16) 和 (18) 的对比, 可以看到在一定的条件下, 标准 PSO 也可以简化为以群体最优值为中心, 群体最优粒子与个体最优值之间的距离为半径, 基于正态分布的随机采样搜索策略, 这一点同式 (18) 描述的 BBPS 的机制是相同的. 因此, 二者在搜索机理上是共同的. 二者不同点在于, BBPS 通过对参数的简化, 严格的定义了粒子遵循正态分布的采样原则, 并精确的定义了每个粒子在每一个周期的均值和方差. 这样的简化形式减少了算法在执行的过程中的参数设置, 节省了计算过程中所需要的变量存储空间.

4 参数在概率意义下的遗忘特性

引理 1 中给出了一般意义下粒子的运动规律表达式, 根据式 (3) 的描述, 本节将分析标准 PSO 具有概率意义下的遗忘特性, 并给出了在没有 V_{\max} 约束下的加速度因子取值, 以保证粒子运动稳定. 考虑式 (3), 展开如图 2 (见下页) 所示.

图 2 中每一行分别对应 PSO 算法初始时刻四个变量: $\mathbf{p}_{ld}^{(k)}$ 、 $\mathbf{x}_{id}^{(k)}$ 、 $\mathbf{x}_{id}^{(k)}$ 和 $\mathbf{v}_{id}^{(k)}$ 的求和数列. 尤其值得注意的是按列来分析, 每一列对应一个周期的粒子位置更新公式, 而 $\prod_{m=j}^n \gamma(m)$ 则是对应于该周期 (列) 的加权向量. 如第一列为 $\mathbf{x}_{id}^{(k+1)}$ 的加权表达式为

$$\mathbf{x}_{id}^{(k+1)} \Rightarrow \left\{ \varphi_{ld}^{(k+1)} \mathbf{p}_{ld}^{(k)} + \varphi_{id}^{(k+1)} \mathbf{p}_{id}^{(k)} + (1 - \varphi_{ld}^{(k+1)} - \varphi_{id}^{(k+1)}) \mathbf{x}_{id}^{(k)} + w \mathbf{v}_{id}^{(k)} \right\} \times \left\{ \prod_{m=2}^n \gamma(m) \right\} = \mathbf{x}_{id}^{(k+1)} \left\{ \prod_{m=2}^n \gamma(m) \right\}$$

同理, 第 j 列为 $\mathbf{x}_{id}^{(k+j)}$ 的加权, 其权重为 $\prod_{m=j+1}^n \gamma(m)$. 因此, 在 $k+n$ 时刻的位置向量更新公式可以描述如下

$$\mathbf{x}_{id}^{(k+n)} = \mathbf{x}_{id}^{(k+1)} \left\{ \prod_{m=2}^n \gamma(m) \right\} + \mathbf{x}_{id}^{(k+2)} \left\{ \prod_{m=3}^n \gamma(m) \right\} + s + \mathbf{x}_{id}^{(k+n-1)} \left\{ \gamma(m) \right\} + \left\{ \varphi_{ld}^{(k+n)} \mathbf{p}_{ld}^{(k)} + \varphi_{id}^{(k+n)} \mathbf{p}_{id}^{(k)} + w^n \mathbf{v}_{id}^{(k)} \right\}$$

其中, 速度惯量因子 $|w| < 1$, 因此 $\forall \varepsilon > 0, \exists N$ 使得 $\lim_{n > N} w^n < \varepsilon$, 从而上式可以表述为如下形式

$$\mathbf{x}_{id}^{(k+n)} = \left\{ \varphi_{ld}^{(k+n)} \mathbf{p}_{ld}^{(k)} + \varphi_{id}^{(k+n)} \mathbf{x}_{id}^{(k)} \right\} + \sum_{j=1}^{n-1} \mathbf{x}_{id}^{(k+j)} \left\{ \prod_{m=j+1}^n \gamma(m) \right\} \quad (19)$$

式 (19) 表明, 当前时刻基于 $\mathbf{p}_{ld}^{(k)}$ 和 $\mathbf{p}_{id}^{(k)}$ 的随机向量和, 以及前 $(k+n-1)$ 时刻位置向量的加权求和, 二者共同构成当前时刻的位置向量. 根据加速度因子的参数取值, 由式 (12) 可知, $\prod_{m=j}^n \gamma(m)$ 描述的加权因子具有概率意义下的遗忘特性. 当加速度因子的随机数取值范围符合该条件时, 即使没有最大速度向量 V_{max} 约束, 也能保证粒子在有限空间中进行搜索, 而且随着搜索周期的递增, 加权因子 $\prod_{m=j}^n \gamma(m)$ 也将在概率意义上逐渐衰减, 直到该时刻对应的历史信息 (位置向量) 被彻底遗忘. 从参数选取的角度而言, 本文的结果与文献 [11] 中 Clerc 建议的结果 (在没有 V_{max} 约束下加速度因子的取值为 1.494) 是相符合的, 而且进一步扩展了加速度因子的取值范围.

不同 r 取值的情况下进行仿真统计实验, 对每个 r 取值进行 100 000 次独立实验进行统计对比, 每次实验记录下加权因子 $\prod_{m=j}^n \gamma(m)$ 衰减至 ε ($\varepsilon = 0.01$) 水平以下时, 所经历过的搜索周期 n , 如图 3(a) 和 3(b) 的所示. 由图 3 可见, 随着 r 取值的增大, 虽然加权因子衰减至 ε 以下的时间逐渐增加, 但是其峰值分布变化缓慢, 依然集中在 10 个周期以内. 而具有显著变化的是, 随着 r 的增大, 图中统计结果的方差急剧增大, 如图 3(d) 中, 衰减周期主要集中在 75 个周期以内, 但一部分实验甚至在超过了 350 个周期后依然大于 ε . 这就意味着在式 (19) 中由于过去的 $n-1$ 时刻对应的加权系数始终处于一个较高的数值水平, 则历史时刻位置向量 $\mathbf{x}_{id}^{(k+j)}$ 的加权累加和将远大于 $\mathbf{p}_{ld}^{(k+n)}$ 和 $\mathbf{p}_{id}^{(k+n)}$ 的数量级, 从而导致个体行为 $\mathbf{x}_{id}^{(k+n)}$ 偏离于导向信息 (发散).

5 结论

从搜索机制而言, 粒子在一段时间内基于来自于群体的社会信息和自身的历史经验, 在一定的解空间中进行搜索, 无论是群体或者个体, 当发现了更优解的时候, 将对社会或个体的信息进行更新, 然后再基于更新后的信息进行搜索. 算法的寻优过程就由这样的一个个搜索片段连接而成, 并且循环反复, 不断地推动群体寻优过程的递进. 反映在数学描述上, 标准 PSO 基于上一时刻的自身状态信息和最优参考信息来迭代更新自身的位置向量.

标准 PSO 位置向量的更新是基于上一时刻状态信息的迭代递推, 本文对粒子的一个搜索片段进行分析, 从时间上对粒子的更新公式延展为以迭代周期描述的一般性表达式, 推导了标准 PSO 粒子的单信息最大搜索空间. 通过分析表明, 经过一定的迭代周期后, 粒子在进入该搜索片段时刻的初始位置向量和速度向量都将逐渐失去作用, 而经过简化后, 标准 PSO 同 BBPS 在搜索机理上是相同的, 而不同点在于 BBPS 更严格地定义了随机搜索的策略. 从另一个角度而言, 基于单信息最大搜索空间对粒子运动的描述, 粒子的动态更新公式可以描述为过往时刻各位置向量的加权求和, 而各加权系数在概率意义上是递减的, 这也可以解释为, 粒子的动态更新是在概率意义上具有遗忘特性的历史信息加权的结果.

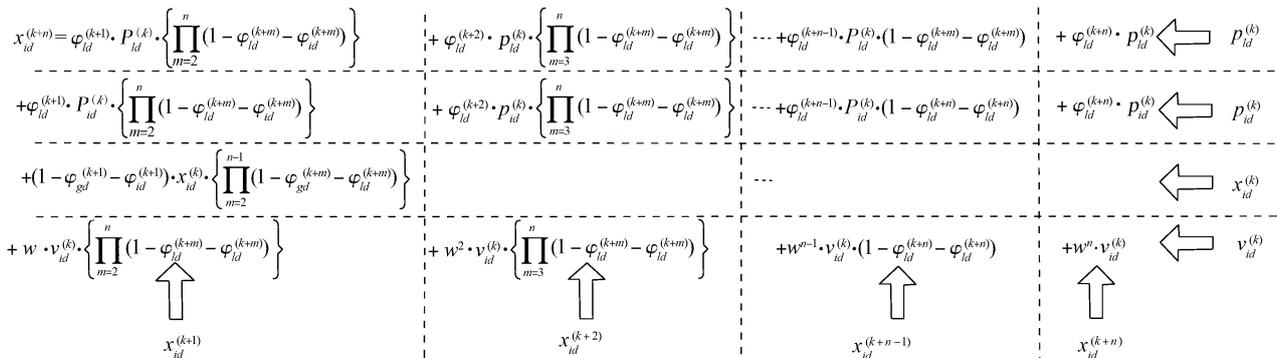


图 2 $\mathbf{x}_{id}^{(k)}$ 一般性描述公式分解示意图

Fig. 2 Decomposition illustration for the general formula of $\mathbf{x}_{id}^{(k)}$

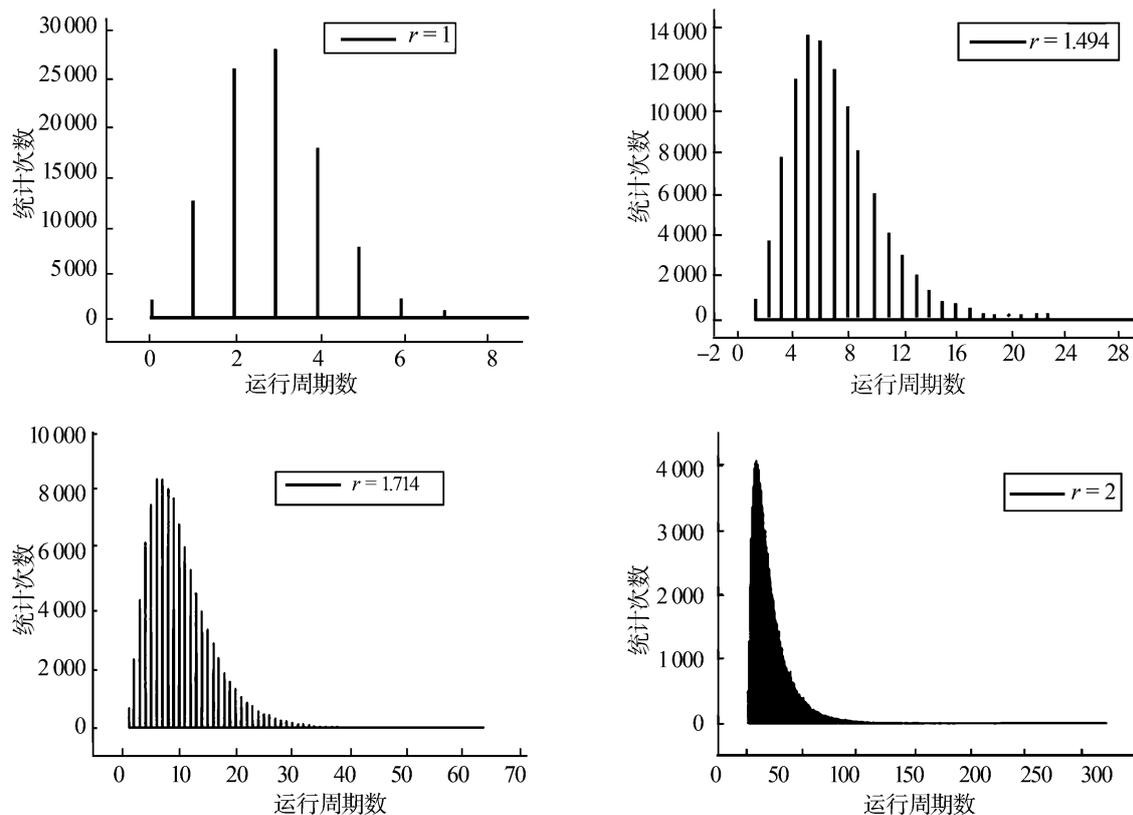


图 3 $\prod_{m=j}^n \gamma(m) < \varepsilon$ 统计特性图
 Fig. 3 The statistical illustrations of $\prod_{m=j}^n \gamma(m) < \varepsilon$

References

- Kennedy J, Eberhart R. Particle swarm optimization. In: Proceedings of IEEE International Conference on Neural Networks. Perth, Australia: IEEE, 1995. 1942–1948
- Riccardo P. *Analysis of the Publications on the Applications of Particle Swarm Optimisation*. Colchester: Hindawi Publishing Corp, 2008. 1–10
- Shi Y, Eberhart R. Parameter selection in particle swarm optimization. In: Proceedings of the 7th International Conference on Evolutionary Programming. London, UK: Springer, 1998. 591–600
- van den Bergh F. An Analysis of Particle Swarm Optimizers [Ph. D. dissertation]. University of Pretoria, South Africa, 2001
- van den Bergh F, Engelbrecht A P. A study of particle swarm optimization particle trajectories. *Information Sciences*, 2006, **176**(8): 937–971
- Clerc M, Kennedy J. The particle swarm: explosion, stability and convergence in a multi-dimensional complex space. *IEEE Transactions on Evolutionary Computation*, 2002, **6**(1): 58–73
- Trelea I C. The particle swarm optimization algorithm: convergence analysis and parameter selection. *Information Processing Letters*, 2003, **85**(6): 317–325
- Chen J, Pan F, Cai T, Tu X Y. The stability analysis of particle swarm optimization without Lipschitz condition constrain. *Journal of Control Theory and Applications*, 2003, **1**(1): 86–90
- Kadirkamanathan V, Selvarajah K, Fleming P J. Stability analysis of the particle dynamics in particle swarm optimizer. *IEEE Transactions on Evolutionary Computation*, 2006, **10**(3): 245–255
- Jiang M, Luo Y P, Yang S Y. Stochastic convergence analysis and parameter selection of the standard particle swarm optimization algorithm. *Information Processing Letters*, 2007, **102**(1): 8–16
- Jin Xin-Lei, Ma Long-Hua, Wu Tie-Jun, Qian Ji-Xin. Convergence analysis of the particle swarm optimization based on stochastic processes. *Acta Automatica Sinica*, 2007, **33**(12): 1263–1268
(金欣磊, 马龙华, 吴铁军, 钱积新. 基于随机过程的 PSO 收敛性分析. *自动化学报*, 2007, **33**(12): 1263–1268)
- Clerc M. Stagnation analysis in particle swarm optimisation or what happens when nothing happens [Online], available: http://clerc.maurice.free.fr/pso/stagnation_analysis, November 20, 2006
- Pan Feng, Chen Jie, Gan Ming-Gang, Cai Tao, Tu Xu-Yan. Model analysis of particle swarm optimizer. *Acta Automatica Sinica*, 2006, **32**(3): 368–377
(潘峰, 陈杰, 甘明刚, 蔡涛, 涂序彦. 粒子群优化算法模型分析. *自动化学报*, 2006, **32**(3): 368–377)
- Jiang M, Luo Y P, Yang S Y. Stagnation analysis in particle swarm optimization. In: Proceedings of the IEEE Swarm Intelligence Symposium. Honolulu, USA: IEEE, 2007. 92–99
- Riccardo P, Broomhead D. Exact analysis of the sampling distribution for the canonical particle swarm optimiser and

- its convergence during stagnation. In: Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation. London, England: ACM, 2007. 134–141
- 16 Kennedy J. Bare bones particle swarms. In: Proceedings of the IEEE Swarm Intelligence Symposium. Indianapolis, USA: IEEE, 2003. 80–87
- 17 Riccardo P, William B L. Markov chain models of bare-bones particle swarm optimizers. In: Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation. London, England: ACM, 2007. 142–149
- 18 Omran M G H, Engelbrecht A, Salman A. Barebones particle swarm for integer programming problems. In: Proceedings of the IEEE Swarm Intelligence Symposium. Honolulu, USA: IEEE, 2007. 170–175
- 19 Omran M, Al-Sharhan S. Barebones particle swarm methods for unsupervised image classification. In: Proceedings of the IEEE Congress on Evolutionary Computation. Swissoel, the Stamford, Singapore: IEEE, 2007. 3247–3252

潘 峰 北京理工大学信息科学技术学院讲师, 博士. 主要研究方向为群智能理论和计算智能. 本文通信作者.

E-mail: andropan@gmail.com

(**PAN Feng** Lecturer, Ph.D. at the School of Information Science and Technology, Beijing Institute of Technology. His research interest covers swarm intelligence theory and computational intelligence. Corresponding author of this paper.)

陈 杰 北京理工大学信息科学技术学院教授, 博士. 主要研究方向为最优控制, 复杂系统和多目标决策与优化.

E-mail: chenjie@bit.edu.cn

(**CHEN Jie** Professor, Ph.D. at the School of Information Science and Technology, Beijing Institute of Technology. His research interest covers optimal control, and multi-objective decision-making and optimization in complex systems.)

辛 斌 北京理工大学信息科学技术学院博士研究生. 主要研究方向为计算智能优、优化理论与方法. E-mail: brucebin@bit.edu.cn

(**XIN Bin** Ph.D. candidate at the School of Information Science and Technology, Beijing Institute of Technology. His research interest covers computational intelligence, optimization theories and methods.)

张 娟 北京理工大学信息科学技术学院讲师, 博士. 主要研究方向为模式识别、运筹学和高层体系结构仿真. E-mail: zhjuan@bit.edu.cn

(**ZHANG Juan** Lecturer, Ph.D. at the School of Information Science and Technology, Beijing Institute of Technology. Her research interest covers pattern recognition, operational research, and high level architecture.)
