DOI:10.13245/j.hust.2009.06.031
第37卷 第6期 华中科技大学学报(自然科学版) Vol.37 No.6
2009年 6月 J. Huazhong Univ. of Sci. & Tech. (Natural Science Edition) Jun. 2009

基于可达状态集扩张的粒子群算法收敛性改进

蔡昭权[1] 黄 翰[2] 郑宗晖[1] 罗 伟[1]
(1 惠州学院网络中心, 广东 惠州 516007;
2 华南理工大学 软件学院, 广东 广州 510006)

摘要: 针对粒子群算法(PSO)改进设计缺乏数学模型和理论依据支持的问题, 研究建立了PSO的吸收态马尔可夫过程模型, 并提出了可达状态集作为收敛性分析的关键指标. 与以往的收敛性分析不同, 研究从可达状态集扩张的角度提出了PSO收敛性对比的理论, 并基于此提出了PSO全局收敛性改进的方法. 最后, 以改进综合学习粒子群算法CLPSO (comprehensive learning particle swarm optimization)为例验证了提出模型与理论的有效性.

关 键 词: 人工智能; 群体智能; 粒子群算法; 收敛性改进; 可达状态集扩张
中图分类号: TP18 文献标识码: A 文章编号: 1671-4512(2009)06-0044-04

Convergence improvement of particle swarm optimization based on the expanding attaining-state set

Cai Zhaoquan[1] *Huang Han*[2] *Zheng Zonghui*[1] *Luo Wei*[1]
(1 Network Center, Huizhou University, Huizhou 516007, China; 2 School of Software Engineering, South China University of Technology, Guangzhou 510006, China)

Abstract: Particle swarm optimization (PSO) is lack of theoretical foundation support for design and improvement. This paper builds up an absorbing Markov process model of PSO, and proposed the attaining-state set as the key factor of convergence analysis. Differently from the prior research, the proposed theoretical results focus on the convergence comparison among the considered PSOs. Later, a convergence improvement method is put forward by the theorem of expanding attaining-state set. Finally, comprehensive learning PSO (CLPSO) is taken as case study and improved to be CLPSO by the proposed theorem. The numerical result proves the presented model and theorem to be valid.

Key words: artificial intelligence; swarm intelligence; particle swarm optimization; convergence improvement; expanding attaining-state set

粒子群算法(PSO)相对遗传算法、进化规划等而言是一项比较新的进化计算工具. 自从文献[1, 2]首次提出粒子群优化算法以来, 越来越多的学者关注PSO的研究. 这些研究工作主要集中在设计PSO的变形以改进求解优化问题的效果. 一种著名的改进是在经典的速度更新式中增加一个惯性权重 w[3], 若 w 设置较大, 则有利于全局搜索; 否则有利于局部搜索. 文献[4, 5]分别给出了线性递减和非线性变化的两种 w 设置方法, 其他的改进形式还有随时间变化的 w 等[6]. 最成功的 w 设置后来由Kennedy[7]所提出, 称为压缩因子权重. 实验数据表明[7]: 带压缩因子的PSO算法比其他惯性权重策略的效果更优. 因此, 这项技术一直延用到目前的PSO算法设计中. 后来, 一些学者做了相关研究[8~11].

虽然已有不少PSO算法的改进研究结果, 但是, 这些结果较少涉及PSO改进的数学模型与理论依据. 一些学者[7, 11, 12]分析了PSO的收敛性, 但并没有给出如何对比不同PSO算法收敛性能的理论和方法. 本研究从可达状态集扩张的角度,

收稿日期: 2008-12-01.
作者简介: 蔡昭权(1970-), 男, 副教授, E-mail: cai@hzu.edu.cn.
基金项目: 国家自然科学基金资助项目(60873078, 60673062, 60803052).

分析了PSO收敛性对比，为PSO算法的改进提供理论依据，并用改进CLPSO[10]算法的效果验证了理论的有效性.

1　PSO随机过程模型与可达状态集

将建立粒子群算法的吸收态马尔可夫过程模型，从理论上研究改进PSO算法的收敛性对比.

一个无约束的实数优化问题可以描述为一个N维最小化问题：

$$\min f(\boldsymbol{x}),\quad \boldsymbol{x}=[x_1,x_2,\cdots,x_N],\tag{1}$$

式中$\boldsymbol{x}$是代表解的N维向量. 在PSO中，一个群体中的成员，称为一个粒子，代表一个解. 记V_i^j和X_i^j为第j维的速度和位置分量，在每一次迭代过程中按下面两个公式进行更新

$$V_i^j=V_i^j+c_1r_{1i}^j(p_i^j-X_i^j)+c_2r_{2i}^j(g^j-X_i^j);\tag{2}$$

$$X_i^j=X_i^j+V_i^j,\tag{3}$$

式中：$p_i=(p_i^1,p_i^2,\cdots,p_i^N)$是第$i$个粒子至今最优适值对应的位置；$g=(g^1,g^2,\cdots,g^N)$是群体中至今最优适值对应的位置；$c_1$和$c_2$是加速常量；$r_{1i}^j$和$r_{2i}^j$是[0, 1]之间的随机数.

经典粒子群算法(简称经典算法)的流程如下：

a. 初始化K个粒子$X_1,X_2,\cdots,X_K$，计算$p_1,p_2,\cdots,p_K$和g，$t=1$；

b. 对于$i=1,2,\cdots,K$和$1,2,\cdots,N$，按式(2)和式(3)进行更新，得到一个X_i后更新p_i和g；

c. $t=t+1$，若$t>n_{\max}$($n_{\max}$为设定的最大迭代次数)，则输出g退出算法，否则，返回第**b**步.

定义1　记粒子群算法a在第t次迭代的(X,V,p,g)为ξ_t^a，$\{\xi_t^a\}_{t=0}^{+\infty}$称为粒子群算法$a$对应的随机过程. 状态空间$Y$满足 $\forall\xi_t^a\in Y$.

定义2　给定来自状态空间Y的随机过程$\{\xi_t\}_{t=0}^{+\infty}$，$Y^*$称为最优状态空间，满足$Y^*\subset Y$，而且$\forall\xi^*\in Y^*$至少有一个解$s^*\in\xi^*$是优化问题的最优解.

定义3　给定一个马尔可夫过程$\{\xi_t\}_{t=0}^{+\infty}$ $(\xi_t\in Y)$，$Y^*\subset Y$为最优状态空间，若$P\{\xi_{t+1}\notin Y^*|\xi_t\in Y^*\}=0$ $(t=0,1,\cdots)$，则称$\{\xi_t\}_{t=0}^{+\infty}$为吸收态马尔可夫过程.

吸收态马尔可夫过程一旦达到最优状态空间，就会吸附在其中永远不会出来. 许多进化算法的设计都采用精英保留策略；这些算法都可以建模为吸收态马尔可夫过程. PSO算法保存了历史最优粒子位置g，自然也是保留了当前最优解. 同样属于吸收态马尔可夫过程.

引理1　满足经典算法基本框架的PSO对应的随机过程为吸收态马尔可夫过程.

证明　记粒子群算法a对应的随机过程为ξ_t^a. 因为经典算法中第**b**步采用式(2)和式(3)进行更新，$\forall Y_1\subseteq Y$和$t=0,1,\cdots$，$P\{\xi_{t+1}^a\in Y_1|\xi_0^a,\cdots,\xi_t^a\}=P\{\xi_{t+1}^a\in Y|\xi_t^a\}$. 因此，$\{\xi_t^a\}_{t=0}^{+\infty}$在$t+1$的状态仅仅由第$t$次迭代的状态所决定，与$\xi_0^a$无关，即$\{\xi_t^a\}_{t=0}^{+\infty}$具有马尔可夫性.

根据第**b**步，用g记录第一次迭代至今最优的粒子位置，以最优粒子记录当前最优解. 一旦在第t次迭代PSO找到最优解$s^*\in\xi^*$，则会用g将其保留到第$t+1$次迭代，即$P\{\xi_{t+1}^a\notin Y^*|\xi_t^a\in Y^*\}=0$. 所以，$\{\xi_t^a\}_{t=0}^{+\infty}$满足吸收态马尔可夫过程性质(定义3). 命题得证.

定义4　粒子群算法a对应的随机过程记为$\{\xi_t^a\}_{t=0}^{+\infty}$，$\Xi_t^a$称为$a$在第$t$次迭代的可达状态集，满足：$\Xi_t^a=\{\xi^x|P\{\xi_t^a\in O(\xi^x,\varepsilon)|\xi_{t-1}^a\notin Y^*\}\geq\delta\wedge\forall\varepsilon>0\wedge\xi^x\in Y\}$ $(t=1,2,\cdots$；$O(\xi^x,\varepsilon)$为ξ^x邻域(勒贝格测度在ε以内))，$\delta>0$且$\prod_{t=0}^{+\infty}(1-\delta)=0$.

2　基于可达状态集的PSO收敛性对比理论

定义5　粒子群算法a对应的随机过程记为$\{\xi_t^a\}_{t=0}^{+\infty}$ $(\forall\xi_t^a\in Y)$，Y^*为最优状态空间，若$\forall\xi_0^a\in Y$，有$\min\limits_{t\to+\infty}P\{\xi_t^a\in Y^*\}=1$，则粒子群算法$a$收敛.

下面定理1和定理2说明了可达状态集与PSO收敛性的紧密关系.

定理1　给定粒子群算法a对应的随机过程$\{\xi_t^a\}_{t=0}^{+\infty}$，若$\exists t'\geq0$，当$t>t'$时，$P\{\xi_t^a\in Y^*|\xi_{t-1}^a\notin Y^*\}\geq\delta$且$\prod_{t=0}^{+\infty}(1-\delta)=0$，则$a$收敛.

证明　根据引理1，$\{\xi_t^a\}_{t=0}^{+\infty}$为吸收态马尔可夫过程. 因$\exists t'\geq0$，当$t>t'$时，$P\{\xi_t^a\in Y^*|\xi_{t-1}^a\notin Y^*\}\geq\delta$，故$P\{\xi_t^a\in Y^*|\xi_{t-1}^a\notin Y^*\}\leq1-\delta$.

若记$P_{\mathrm{not}}^a(t)=\prod_{i=0}^{t-1}P\{\xi_{i+1}^a\notin Y^*|\xi_i^a\notin Y^*\}$，则$\lim\limits_{t\to\infty}P_{\mathrm{not}}^a(t)\leq\prod_{t=0}^{+\infty}(1-\delta)$. 因为$\prod_{t=0}^{+\infty}(1-\delta)=0$，$\lim\limits_{t\to\infty}P_{\mathrm{not}}^a(t)=0$，所以，PSO算法$a$在迭代时间趋于无穷时必能至少一次达到最优状态.

因为a对应的随机过程$\{\xi_t^a\}_{t=0}^{+\infty}$为吸收态马尔可夫过程，所以根据定义3，$P\{\xi_t^a\notin Y^*|\xi_{t-1}^a\in Y^*\}=0$，即$P\{\xi_t^a\notin Y^*|\xi_{t-1}^a\in Y^*\}=1$ $(t=1,2,\cdots)$. 由全概率公式，有

$$P\{\xi_t^a\notin Y^*\}=P\{\xi_t^a\notin Y^*|\xi_{t-1}^a\notin$$

$Y^*\}P\{\xi_{t-1}^a\notin Y^*\}+P\{\xi_t^a\notin Y^*\mid\xi_{t-1}^a\in Y^*\}P\{\xi_{t-1}^a\in Y^*\}=\cdots=P\{\xi_0^a\notin Y^*\}P_{\mathrm{not}}^a(t)$.
因为 $P\{\xi_t^a\in Y^*\}=1-P\{\xi_t^a\notin Y^*\}$, 所以

$$\lim_{t\to\infty}P\{\xi_t^a\in Y^*\}=1-\lim_{t\to\infty}P\{\xi_t^a\notin Y^*\}=1-P\{\xi_0^a\notin Y^*\}\lim_{t\to\infty}P_{\mathrm{not}}^a(t)=1.$$

定理 2 任意给定的粒子群算法 a 和 b, 若它们的可达状态集对 $t=1, 2, \cdots$ 满足 $\Xi_t^a\cap Y^*\subseteq\Xi_t^b\cap Y^*$, 且 a 收敛, b 对应的 $\{\xi_t^b\}_{t=0}^{+\infty}$ 为吸收态马尔可夫过程, 则 b 也收敛.

证明 因为 a 收敛, 根据定义 5, $\forall\xi_0^a\in Y$, 有 $\lim_{t\to+\infty}P\{\xi_t^a\in Y^*\}=1$; 所以 $\forall\varepsilon=1/t>0$, 必定 $\exists t'>0$, 当 $t>t'$ 时, 有 $P\{\xi_t^a\in Y^*\}>1-\varepsilon=1-1/t$. 记 $\delta_t=1-1/t>0$, 则 $\prod_{t=0}^{+\infty}(1-\delta_t)=0$ 成立, 所以当 $t>t'$ 时, $\xi_t^a\in\Xi_t^a$, 即 $\Xi_t^a\cap Y^*\neq\varnothing$. 根据定义 3, 当 $t>t'$ 时, $\Xi_t^b\cap Y^*\supseteq\Xi_t^a\cap Y^*\neq\varnothing$. 根据定义 4, $P\{\Xi_t^b\cap Y^*\neq\varnothing\}\geqslant\delta_t$, 所以当 $t>t'$ 时, $P\{\xi_t^a\in Y^*\mid\xi_{t-1}^b\notin Y^*\}\geqslant\delta_t$, 而且 $\prod_{t=0}^{+\infty}(1-\delta_t)=0$. 根据定理 1, PSO 算法 b 收敛.

定理 2 可得: 可达状态集与最优状态集的 定了 PSO 算法的收敛性. $\Xi_t^a\cap Y^*\subseteq\Xi_t^b\cap$, 2, $\cdots$) 的条件说明了较大的可达状态集可以得到较强的收敛性.

3 基于可达状态集扩张的 CLPSO 改进

改进算法的思想是: 当 CLPSO 算法不能找到更优解时, 随机选择一个较差的粒子, 将该粒子转移到局部收敛区域以外且较优的位置. 将粒子转移时遵守的是最小方差优先原则, 即粒子群在某一维的方差越小时, 转移出局部收敛区的概率就越大. 这是因为粒子群算法与一般的进化算法不同, 粒子下一次迭代的状态除了由当前粒子状态决定外, 还受到粒子当前速度和 g 的影响. 因此, 除了考虑位置上的方差外, 还需要加入粒子的速度和粒子与 g 的距离.

记改进算法为 CLPSO, 其算法步骤如下:

a. 与原 CLPSO[10] 一致.

b. 若 g_i 得到更新, 则 $u=0$; 否则, $u=u+1$. (u 为 g_i 无更新累积迭代数).

c. 若 $u>10$, 则运行 **d**~**f**; 否则, 直接运行 **g**.

d. 计算粒子群各维的位置均值 $(\mu^1, \mu^2, \cdots, \mu^N)$ 和标准差 $(\sigma^1, \sigma^2, \cdots, \sigma^N)$.

e. 随机选择两粒子, 择适值较差的粒子记为 X_r.

f. 对于 $j=1, 2, \cdots, N$, 以 $P_i^j=\max(1-\sigma^j-|v_i^j|-|\mu^j-g^j|, 0)$ 概率令粒子 X_r 在第 j 维跳出局部收敛区: 在区间 $[X_{\min}, \mu^j-\sigma^j]$ 上随机产生一个数 r_1 作为 X_r^j, 计算 X_r 的适值 f_r^1; 在区间 $[\mu^j+\sigma^j, X_{\max}]$ 上随机产生一个数 r_2 作为 X_r^j, 计算 X_r 的适值 f_r^2; 若 f_r^1 优于 f_r^2, 则 $X_r^j=r_1$; 否则 $X_r^j=r_2$.

g. 与 CLPSO[10] 一致.

改进的 CLPSO 比原来的 CLPSO 多了步骤 **b** 和 **d**~**f**, 步骤 **d**~**f** 会增加算法的计算量, 所以只是在 $u>10$ 时, 即 CLPSO 连续 10 次迭代没有改善时才运行. 当 CLPSO 求解效果较好时 ($u>10$ 不成立), 改进的 CLPSO 时间复杂度与 CLPSO 一致. 因此, 当改进的 CLPSO 算法增加计算时间复杂度时, 将获得比 CLPSO 更好的求解效果.

定理 3 假设原算法 CLPSO 和改进的 CLPSO 对应的随机过程分别为 $\{\xi_t^1\}_{t=0}^{+\infty}$ 和 $\{\xi_t^2\}_{t=0}^{+\infty}$, 记 Ξ_t^1 和 Ξ_t^2 为原 CLPSO 和改进的 CLPSO 在第 t 次迭代可达状态集, 且 $Y^*\subseteq Y$ 是最优状态空间; 则 $\Xi_t^1\cap Y^*\subseteq\Xi_t^2\cap Y^*$ $(t=1, 2, \cdots)$.

证明 根据引理 1, $\{\xi_t^1\}_{t=0}^{+\infty}$ 和 $\{\xi_t^2\}_{t=0}^{+\infty}$ 为吸收态马尔可夫过程. 根据改进的 CLPSO 的流程, 原 CLPSO 与改进的 CLPSO 具有相同的初始化设置. 当 $u\leqslant 10$ 时, 附加的变异机制 (改进的 CLPSO 中步骤 **d**~**f**) 并没有执行, 原 CLPSO 与改进的 CLPSO 采用的速度更新算子一致. 此时 $\forall\xi^*\in Y^*$, 若 $\xi^*\in\Xi_t^1$ 则 $\xi^*\in\Xi_t^2$. 当 $u>10$ 时, 附加的变异机制 (改进的 CLPSO 中步骤 **d**~**f**) 将会执行. 根据改进的 CLPSO 中步骤 **e**, 随机选择两粒子, 择适值较差的粒子执行步骤 **f**; $\forall\xi^*\in Y^*$, 当 $\exists\xi^*\in Y^*$ 时, 代表最优解的粒子因为适应值最高不可能在步骤 **e** 中被选中, 所以, 若 $\xi^*\in\Xi_t^1$ 则 $\xi^*\in\Xi_t^2$ $(t=1, 2, \cdots)$. 若最优解是在改进的 CLPSO 中的附加变异机制中找到的, 则 $\exists\xi^*\in\Xi_t^2$ 满足 $\exists\xi^*\notin\Xi_t^1$. 因此, $\Xi_t^1\cap Y^*\subseteq\Xi_t^2\cap Y^*$ 成立.

推论 1 若原 CLPSO 收敛, 则改进的 CLPSO 必收敛; 若改进的 CLPSO 不收敛, 则原 CLPSO 必不收敛.

由定理 3 和推论 1 的结论可得: 改进的 CLPSO 本质上实现了可达状态集的扩张, 从而得到了比原来 CLPSO 算法更强的全局收敛性.

4 实验结果与分析

将给出原 CLPSO 和改进的 CLPSO 求解 13 个标准函数优化问题极小值的部分实验结果, 测

试问题来自于文献[10]. 原 CLPSO 和改进的 CLPSO 都采用 40 个粒子作为种群规模, 最大迭代次数都是 5 000 次, 平均每个问题各计算 30 次, 其求解结果见表 1.

表 1　原 CLPSO 和改进的 CLPSO 算法求解 $f_1 \sim f_{13}$ 函数优化问题的结果比较

函数	最优解		标准差	
	原 CLPSO	改进的 CLPSO	原 CLPSO	改进的 CLPSO
f_1	4.46×10^{-14}	3.46×10^{-70}	1.73×10^{-14}	2.88×10^{-70}
f_2	3.79×10^{-12}	3.46×10^{-42}	2.19×10^{-12}	2.06×10^{-42}
f_4	2.600	0.391	2.400	0.270
f_7	5.780×10^{-3}	3.210×10^{-3}	0.002	0.001
f_8	-9.54×10^{3}	-1.21×10^{4}	215.46	334.59
f_9	4.85×10^{-10}	0	3.63×10^{-10}	0
f_{11}	3.14×10^{-10}	0	4.64×10^{-10}	0
f_{12}	1.12×10^{-11}	6.71×10^{-12}	1.12×10^{-10}	5.78×10^{-12}
f_{13}	1.07×10^{-11}	6.41×10^{-12}	1.70×10^{-27}	5.52×10^{-12}

由表 1 可见, 原 CLPSO 在 f_1, f_2, f_4, f_8, f_9, f_{11}, f_{12}和 f_{13} 求解时都陷入了局部最优解; 而改进的 CLPSO 提高了全局搜索性能, 不仅在效果上优于原CLPSO, 且在 f_8, f_9, f_{11}, f_{12}和 f_{13}都找到全局最优解. 其他函数的求解效果原CLPSO与改进的 CLPSO 持平, 但改进的 CLPSO 标准差较小, 即求解较为稳定. 实验结果说明了本文的改进设计是有效的, 改进的 CLPSO 既能保持原 CLPSO 在多峰函数问题求解上的优势, 又可以在原 CLPSO 陷入局部最优的情况下跳出“早熟”困境, 进而实现全局优化的性能. 实验结果同时也说明了本文提出的基于可达状态集收敛性改进理论(定理 2 和 3 及推论 1)是正确有效的.

参　考　文　献

[1] Kennedy J, Eberhart R C. Particle swarm optimization[C] //Proceeding of IEEE Intentional Conference on Neural Networks. Piscataway: IEEE Service Center, 1995: 1 942-1 948.

[2] Eberhart R C, Kennedy J. A new optimizer using particle swarm theory[C] //Proceedings of the Sixth International Symposium on Micromachine and Human Science. Piscataway: IEEE Service Center, 1995: 39-43.

[3] Shi Y, Eberhart Y C. A modified particle swarm optimizer[C] //Proceedings of the IEEE International Conference on Evolutionary Computation. Piscataway: IEEE Press, 1998: 69-73.

[4] Shi Y, Eberhart R C. Parameter selection in particle swarm optimization[C] //Proceedings of the 7th Annual Conference on Evolutionary Programming. New York: Springer-Verlag, 1998: 591-600.

[5] Shi Y, Eberhart R C. Particle swarm optimization with fuzzy adaptive inertia weight[C] //Proccedings Workshop Particle Swarm Optimization. Indianapolis: IEEE Service Center, 2001: 101-106.

[6] Ratnaweera A, Halgamuge S, Watson H. Self-organizing hierarchical particle swarm optimizer with time varying accelerating coefficients[J]. IEEE Transactions on Evolutionary Computation, 2004, 8(3): 240-255.

[7] Clerc M, Kennedy J. The Particle swarm-explosion, stability, and convergence in a multidimensional complex space[J]. IEEE Transactions on Evolutionary Computation, 2002, 6(1): 58-73.

[8] Krink T, Vesterstroem J S, Riget J. Particle swarm optimization with spatial particle extension[C] //Proceedings of IEEE Congress on Evolutionary Computation. Honolulu: IEEE Inc, 2002: 1 474-1 479.

[9] van den Bergh F, Engelbrecht A P. A cooperative approach to particle swarm optimization[J]. IEEE Transactions on Evolutionary Computation, 2004, 8(3): 225-239.

[10] Liang J J, Qin A K, Suganthan P N, et al. Comprehensive learning particle swarm optimizer for global optimization of multimodal functions[J]. IEEE Transactions on Evolutionary Computation, 2006, 10(3): 281-295.

[11] 吕艳萍, 李绍滋, 陈水利, 等. 自适应扩散混合变异机制微粒群算法[J]. 软件学报, 2007, 18(11): 2 740-2 751.

[12] 黄　翰. 基于关系模型的进化算法收敛性与收敛时间分析及改进研究[D]. 广州: 华南理工大学计算机学院, 2008.