

# 一种种群自适应收敛的快速遗传算法

朱 钰 韩昌佩

(中国科学院上海技术物理研究所 上海 200083)

**摘要** 作为一种全局搜索算法,遗传算法的局部搜索能力较低,后期产生的无效进化与早熟收敛影响优化的速度和精度。已有的改进策略多以算法的时间复杂度为代价提高后期效率,严重限制了遗传算法在工业控制系统中的应用。针对这种情况,提出了一种新型种群自适应收敛的快速遗传算法,即通过提高种群的遗传质量,在严格控制算法复杂度的前提下提高优化性能。仿真结果证明,在不增加时间复杂度的前提下,新算法显著地提升了收敛精度和收敛速度。

**关键词** 遗传算法,早熟收敛,自适应变异算子,工业控制

中图分类号 TP18 文献标识码 A

## Improved Genetic Algorithm with Adaptive Convergence Populations

ZHU Yu HAN Chang-pei

(Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China)

**Abstract** The premature convergence seriously affects the performance of genetic algorithm. At present time, most of the improved algorithms focus on improving the convergence accuracy and speed at the expense of the algorithm time complexity, which limits the applications of genetic algorithm in industrial control system. For this situation, this paper presented a new improved genetic algorithm with adaptive convergence populations. This algorithm optimizes performance through increasing the genetic quality of populations, and at the same time, strictly controls the algorithm complexity. Simulation results show that the new algorithm can significantly improve the accuracy and speed of convergence, without time complexity increasing.

**Keywords** Genetic algorithm, Premature convergence, Adaptive mutation operator, Industrial control

### 1 引言

遗传算法(Genetic Algorithm, GA)是一种典型的随机全局优化算法,其原理是基于自然生物进化过程的隐性抽象<sup>[1]</sup>,能够实现复杂模型的快速寻优。GA在求解过程中不依赖于目标问题的数学模型和最优解邻域,兼具很强的全局搜索能力,且在多变量条件下,其寻优快速性优于模拟退火、粒子群优化等智能优化算法<sup>[2]</sup>,因此广泛应用于函数优化、生产调度、自动控制、图像处理和模式识别等领域。

在部分高精度工业控制系统中,需要在已知控制函数的基础上,对控制模型进行在线分析和状态预测,即时调整控制参数,以提高动态控制性能。图1为一典型的在线优化控制流程。在反馈回路中加入了基于系统仿真模型 $M'$ 的寻优算法,算法以固定的频率接收控制对象的状态向量反馈 $D(t)$ ,以历史状态向量为基础,对数学模型的解空间进行寻优,获得最优参数集 $A(t)$ ,用于实时调整实际系统的控制参数。

在系统的数学模型已知的前提下,这类控制问题可抽象为周期性地搜索多变量多极值的非线性控制函数的最优解向量。在实际应用中,寻优算法曾主要采用数值直接逼近法,但对于高复杂度的系统模型,直接逼近法已无法达到理想的优

化速度,遗传算法逐渐成为主要的发展方向。

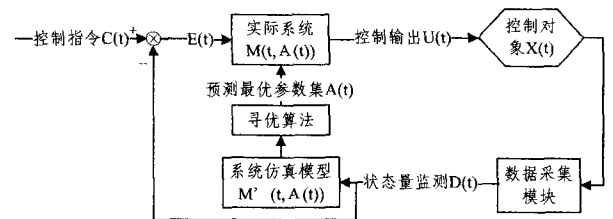


图1 基本遗传算法流程图

由于自然演化的随机性,导致遗传算法的局部搜索能力低。在搜索中后期,种群多样性剧减,群体缺乏有效的等位基因,导致遗传算子作用下不再生成高阶竞争模式,产生了大量的无效进化;同时,种群有可能收敛于解空间内的一个随机非优个体,其称为早熟收敛<sup>[3]</sup>。在工业控制系统的即时仿真中,无效进化将延长最终收敛时间,导致系统在控制周期内无法获得最优解向量;早熟收敛产生错误控制指令和控制参数,导致系统失控。因此,提高遗传算法中后期的搜索效率,抑制早熟收敛,对提高此类控制系统的稳定性和精度具有重大的意义。

影响种群收敛过程的主要因素包括初始种群分布、适应值函数和遗传算子的设计。为了提高遗传算法的中后期收敛

到稿日期:2011-12-03 返修日期:2012-03-17 本文受国家风云四号大气垂直探测仪项目资助。

朱钰(1984-),女,博士生,主要研究方向为电路与高精度高速度运动控制算法, E-mail: cardinaljade@qq.com; 韩昌佩(1979-),男,副研究员,主要研究方向为高可靠性、高精度扫描机构的控制技术。

特性,抑制早熟收敛,学术界提出了多种针对性改进策略,但它们多集中于适应度优化<sup>[3,4]</sup>和自然演化随机过程的补偿<sup>[6,7]</sup>。这类算法均以提高时间复杂度作为抑制早熟收敛的代价。而在工业控制系统中,每次进化的运算时间由控制参数更新周期决定,受控制对象和精度要求的严格限制,现有的大部分改进策略由于时间复杂度过高均无法采用。

文献[5]提出的小区间生成法,通过改进初始种群的生成方式,在不增加进化过程时间复杂度的前提下提高了初始种群的遗传信息量,对高速系统的早熟收敛起到了一定的抑制作用。但是该方法无法改善种群的后期收敛情况,当子代种群丧失遗传优势后,依然会出现长时间的无效进化和早熟收敛现象。

为了获得高速准确的全局优化算法,在已有算法的基础上引入变异算子的自适应校正,设计一种低时间复杂度的种群自适应收敛的快速遗传算法(Genetic Algorithm with Adaptive Convergence Populations, ACPGA),以满足工业实时控制系统的应用需求。

## 2 遗传算法种群收敛的优化策略

### 2.1 遗传算法的原理

遗传算法(GA)的核心概念是基于达尔文进化论思想的自然选择、交叉和变异过程。GA的优化对象称为染色体,染色体携带的关键信息称为遗传信息。GA根据适应值函数对遗传信息进行数学编码,抽象为长度为 $M$ 的解向量,向量中的元素称为等位基因。在优化解空间内的任意染色体 $C_x$ ,其携带的遗传信息都可以用解向量表示为:

$$C_x = \{g_j\}; j=1, 2, 3, \dots, M \quad (1)$$

优化开始时,GA在解空间中生成一个尺度为 $N$ 的随机染色体组作为初始种群:

$$P_0 = \{C_{0i}\}; i=1, 2, 3, \dots, N \quad (2)$$

进化过程中,种群里含有高适应值基因的染色体,比含低适应值基因的染色体有更高的几率进行繁殖。被选为父代的染色体根据适应值的高低进行重组,并通过交叉等位基因、基因突变和选择产生子代种群 $P_k$ 。上述过程不断迭代执行,直至满足优化准则或 $k$ 达到最大迭代次数 $G_{max}$ 为止。

GA的基本流程如图2所示。

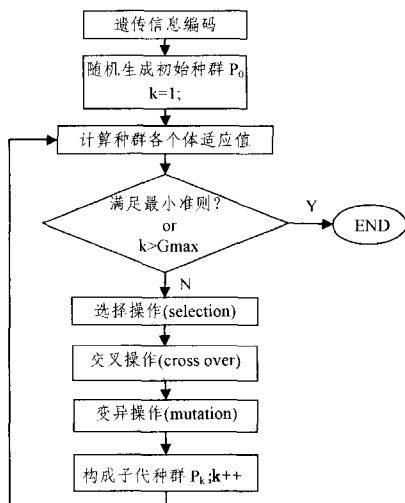


图2 基本遗传算法流程图

由图2可知,适应值函数计算和遗传算子操作在迭代过

程中被循环调用,是影响遗传算法时间复杂度的主要因素。而在工业控制系统中,适应值函数多为多变量的非线性复杂函数,占用大量运算资源,并难以进行优化,与之相比,选择、交叉、变异操作的运算时间可以忽略不计。因此工业上改进遗传算法的关键在于在保证收敛精度的情况下,尽可能降低适应值函数的调用次数。

适应值函数的调用次数与进化代数 $k$ 和种群尺度 $N$ 成正比。对于基本遗传算法,减小种群尺度 $N$ 会降低种群的遗传多样性,导致进化中后期更容易出现无效进化,令 $k$ 增加,同时亦更易陷入早熟收敛。

因此,ACPGA设计的关键问题在于如何在较小种群尺度下抑制无效进化,并避免早熟收敛。为此,应从初始种群生成法和遗传算子的优化方式入手,提高收敛效率。

### 2.2 初始种群的优化策略

初始种群的生成在进化迭代之前执行,对算法的优化速度没有影响。而初始种群的状态会直接影响后续进化迭代的质量。在基础遗传算法中,初始种群在解空间中随机生成,初始染色体可能会在某一局部区域集中分布,其携带的基因信息无法代表整个解空间,导致种群过早丧失遗传优势。

小区间生成法将各等位基因的取值范围平均分成 $N$ 个子区间,在各子区间内部分别随机生成一个初始个体作为 $P_0$ 的元素,以保证初始种群内的染色体均匀分布在解空间上。当解空间连续,且遗传信息在空间内均匀分布时,小区间生成法能够明显改善收敛特性。反之,此算法将引入一部分无效遗传信息,以降低收敛效率。

由于实际工业控制对象的解空间大部分非均匀、非线性,且可能存在不连续区间,因此不适合直接采用理想的小区间生成法。在初始种群生成前,需要分析控制对象,根据遗传信息在解空间中的分布情况决定每个子区间内的采样数,进一步提高初始种群的质量。

### 2.3 遗传算子的优化策略

传统的遗传算子在进化后期会引起种群遗传多样性的丧失,导致无效进化和早熟收敛。大部分遗传算子优化策略多以适应值作为优化阈值,在算法中引入额外的适应值函数运算,在适应值函数复杂度高时,会大幅增加运算时间。其不适用于实时工控系统。为此,需要对遗传算子中影响种群遗传多样性的参数和模块进行考察,以寻找与适应值无关的优化途径。

#### 1) 历史队列

遗传算法是一种迭代算法,在每次迭代后,父代种群被后代所替代。但由于变异和交换算子的随机性,子代种群内的个体并不一定都优于父代个体。传统遗传算法中,父代种群的大部分优秀个体被抛弃,造成运算资源的浪费。为充分利用父代种群,在遗传算子中引入长度为 $Q(Q < N)$ 的先入先出的历史队列。每次迭代,令本次父代的 $V$ 个( $V < Q$ )高适应值个体排入队尾,若队列满,则队首的多余旧个体出队。在选择操作时,历史队列中保存的优秀个体与经过交叉和变异算子作用过的子代种群共同竞争,经过优胜劣汰的选择,生成新一代染色体种群。

历史队列的引入客观上增多了进化的有效个体样本数,同时增加了种群的遗传多样性。

#### 2) 自适应变异算子

变异操作是恢复种群遗传多样性的关键。遗传算法进入

进化末期时,种群中均为近似最优个体。此时绝大多数的变异操作只向种群中引入有害个体,在子代筛选时就被排除,产生大量的无效进化。

为了保证进化后期的局部搜索效率,一个理想的变异算子,其随机变异幅度的平均值应与迭代次数成反比;同时保证携带的遗传信息在取值范围内。

据此,设计以迭代次数为参量的自适应随机变异算子:

$$g_i' = g_i + f(t, g_i) \quad (3)$$

式中,  $g_i$  被选出执行变异操作的遗传信息分量,其取值范围为  $(g_{i\min}, g_{i\max})$ ,  $g_i'$  为变异结果;  $t$  为群体进化代数;  $f(t, y)$  是关于当前进化代数  $t$  和遗传信息分量  $y$  的函数,三者的数学关系可表示为:

$$f(t, y) = \begin{cases} (g_{i\min} - y)(1 - r^{(1 - \frac{t}{T})^2}), & ran = 1 \\ (y - g_{i\min})(1 - r^{1 - \frac{t}{T}})^2), & ran = 0 \end{cases} \quad (4)$$

式中,  $ran$  为控制变异方向的值,是取值 0 或 1 的随机数,  $r$  为控制变异幅值的在  $[0, 1]$  间取值的随机数;  $T$  为最大进化代数。

自适应变异算子的主要作用是通过引入迭代次数  $t$ , 抑制收敛后期全解空间变异的发散特性,同时保证变异操作的一定随机性,从而提高进化种群的遗传信息量,抑制早熟收敛。

#### 2.4 种群自适应收敛的快速遗传算法设计

综合上述优化策略,为了提高运算速度以适应工业实时控制的要求,种群自适应收敛的快速遗传算法(ACPGA)在设计中放弃了需要增加适应值函数调用的优化理论,转而通过初始种群、历史队列和自适应变异算子 3 种优化策略,专注提高遗传算法种群收敛的遗传多样性,以抑制无效进化和早熟收敛。ACPGA 的工作流程如图 3 所示,其中各步简述如下。

##### 1) 遗传信息编码

系统数学模型中,待辨识的不相关的控制参数  $\alpha_i$  即为遗传信息,其集合可表示为时变向量  $A(t)$ :

$$A(t) = \{\alpha_i(t)\}, i \in [1, M]$$

式中,  $t$  为时间变量。

##### 2) 解空间分析

在离线情况下,将控制参数  $\alpha_i$  的值域平均划分成  $W (W > N, N$  为初始种群尺度) 个子域,取每个子域的中点,利用已知的系统数学模型  $M'$ ,固定其他参数,计算遗传适应值,从而得到  $\alpha_i$  的次优、最优解分布趋势。根据分布趋势对子域赋权值,权值的值域为  $[1, V] (N > V)$ 。越接近最优解的子域权值越高。

##### 3) 非均匀区间法

根据 2) 得到控制参数值子域权值,将权值相同的子域看作一个区域,在每一个区域内随机抽取初始样本。区域的权值越高,该区域内抽取的初始样本数越多。由此构成的  $N$  尺度初始种群,能在保证遗传基因多样性的前提下,更集中于最优、次优解附近,有效提高了优化效率。

##### 4) 更新历史队列

每次遗传过程开始时,计算种群个体适应值后,将父代中最佳个体与历史队列中的已有样本进行比较。若当前最佳个体优于队列中的最差个体,则最佳个体入队;若队列满,则最差个体出队。历史队列更新完毕后,执行标准的遗传选择和交叉操作。

##### 5) 自适应变异操作

自适应随机变异算子如式(3)、式(4)所示,在算子作用

下,变异幅度与进化代数成反比。

##### 6) 历史队列参与竞争

变异后的子代与历史队列中的历史最优样本进行选择,择优者进入子代种群。

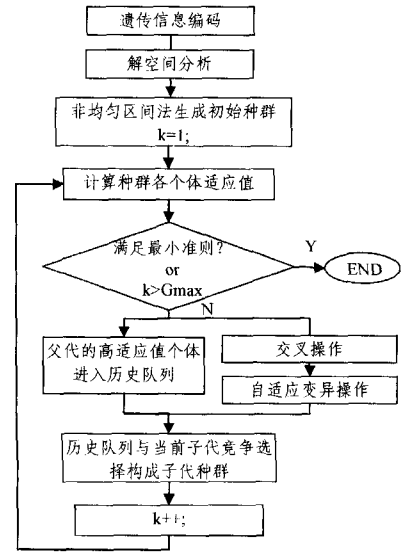


图 3 ACPGA 算法流程图

### 3 仿真实验

为了验证 ACPGA 的全局优化性能,本文采用某多阶多变量运动控制系统的开环传递函数作为对象,对其控制参数进行快速滚动仿真寻优。

运动控制系统的开环控制函数可以写作:

$$M(s) =$$

$$\frac{K(1 + \alpha_1 \frac{1}{s} + \alpha_2 s)(\alpha_3 s + 1)}{(1 + \alpha_4 s)(1 + \alpha_5 s)(1 + \alpha_6 s)(1 + \alpha_7 s + \alpha_8 s^2)(1 + \alpha_9 s + \alpha_{10} s^2 + \alpha_{11} s^3)} \quad (5)$$

式中,  $K$  是开环增益常数,  $\alpha_i (i \in [1, 11])$  是待辨识的控制参数。

系统实际工作时,记录了大量运动状态数据,同时,通过实地调试,获得了所有控制参数的最优取值。使用 matlab 编写 GA 和 ACPGA 两种算法的参数辨识算法。以实测数据为输入,以控制参数实调最优值为参考,在待辨识参数数目分别为 3、6 和 11 时,依次对  $M(s)$  进行参数辨识。每一类辨识进行 20 次仿真,取其结果的平均值。

由表 1 可见,当辨识参数的个数分别为 3、6、11 时,ACPGA 与 GA 的总耗时之比分别为 0.50682, 0.48174 和 0.47418,耗时的减少主要得益于收敛代数的降低;辨识参数的平均相对误差之比分别为 0.50318, 0.44798, 0.38247。无论在优化效率还是优化精度方面,ACPGA 均有显著提高;并且随着辨识参数个数的增加,ACPGA 的性能优势也越明显。

表 1 多个待辨识参数情况下,ACPGA 与 GA 的平均优化性能比较

辨识参数个数	算法	单次进化迭代的平均耗时(sec)	总耗时(sec)	平均相对误差	最大相对误差
3	GA	6.6	1320	0.0471	0.1512
	ACPGA	6.917	669	0.0237	0.0982
6	GA	8.625	1725	0.0471	0.1427
	ACPGA	8.897	831	0.0211	0.0907
11	GA	10.65	2130	0.0468	0.1182
	ACPGA	10.792	1010	0.0179	0.0812

仿真结果证明了 ACPGA 算法的收敛精度和快速性。这主要得益于:

- 1) 非均匀小区间生成法提供的优化初始种群;
- 2) 历史队列对种群有效样本数的提高;
- 3) 自适应变异操作抑制后期收敛的发散特性。

更重要的是, ACPGA 采用的以上 3 种改进策略都避免了适应值函数的额外调用。从表 1 的单个进化迭代的平均耗时可知, ACPGA 算法单个迭代的时间复杂度与 GA 算法相近, 且随着辨识参数个数的增加, 二者的时间消耗差距减小。当辨识参数个数为 11 时, ACPGA 增加的时间开销仅为 GA 的 0.13%。可见, ACPGA 是一种十分简洁的快速算法, 更适合在工业即时仿真中应用。

**结束语** 遗传算法作为一种高效成熟的全局优化算法, 已经逐渐在工业控制领域投入使用。原有的遗传算法研究多关注于算法本身的收敛精度和收敛速度, 忽略了实际系统对于时间复杂度的要求。本文以提高遗传算法的运算速度为切入点, 以已有的小区间生成法和自适应遗传算子理论为基础, 提出了 ACPGA 算法。该算法在不增加算法复杂度的同时, 提高了遗传算法的收敛精度和速度。仿真结果表明, 在复杂函数的多参数辨识过程中, ACPGA 算法的收敛精度、速度远远优于基本 GA 算法, 同时它没有增加算法整体的时间复杂度, 尤其适合作为实时优化算法应用于工业控制系统中。

(上接第 189 页)

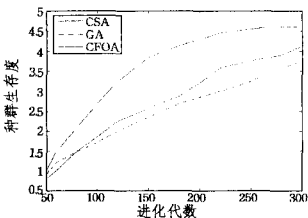


图 6 3 种调度算法种群生存度比较 ( $m=40000, n=120$ )

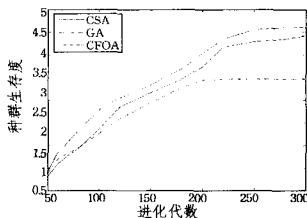


图 7 3 种调度算法种群生存度比较 ( $m=40000, n=100$ )

通过表 1 不难发现: 闭环放大倍数的相对不稳定程度和开环放大倍数的相对不稳定程度的比值大约在 50% 左右波动, 闭环放大倍数变化相对开环放大倍数变化的比值大约在 30% 左右波动。这表明, 在增加反馈机制后, CFOA 算法的稳定性能明显提高。在图 2、图 3 中, CFOA 算法的响应时间是最少的。在图 4、图 5 中, CFOA 算法的执行时间在大部分情况下都是最少的。在图 6、图 7 中, 初始种群经过 CFOA 算法进化后所获得的种群生存度最高, 整体抗体性能最优良。实验结果充分证明了 CFOA 算法的有效性和稳定性。

**结束语** De Castro 博士提出的传统免疫克隆选择算法是一个比较典型的免疫算法, 该算法能有效增强算法的全局搜索能力和多解搜索能力<sup>[9]</sup>, 但其稳定性受抗体亲和度和浓度的影响, 波动较大。本文在该算法的基础上, 提出了一种新的基于反馈机制的克隆反馈优化算法 CFOA, 该算法充分考虑了抗体的浓度和维持抗体多样性两方面因素, 同时融入了进化反馈深度机制和种群生存度概念。最后将 CFOA 算法

## 参考文献

- [1] Potts J C, Giddens T D, Yadav S B. The development and evaluation of an improved genetic algorithm based on migration and artificial selection [J]. *IEEE Transactions on Systems, Man, and Cybernetics*, 1994, 24(1): 73-86
- [2] Zhu Yu. A New Parameter Optimization Algorithm of Penicillin Fermentation Model [C] // 2011 International Conference on Transportation and Mechanical & Electrical Engineering. 2011; 2592-2595
- [3] 邓莉, 鲁瑞华. 一种改进的抑制早熟收敛的模糊遗传算法 [J]. *计算机科学*, 2007, 34(11): 150-153
- [4] 田丰, 姚爱民, 孙小平, 等. 基于个体相似度的双种群遗传算法 [J]. *计算机工程与设计*, 2011, 32(5): 1789-1791, 1848
- [5] 高玮. 改进的快速遗传算法及其性能研究 [J]. *系统工程与电子技术*, 2003, 25(11): 1427-1430
- [6] Patel R, Raghuvanshi M M, Jaiswal A N. Modifying Genetic Algorithm with Species and Sexual Selection by using K-means Algorithm [C] // 2009 IEEE International Advance Computing Conference. 2009: 114-119
- [7] Guo Peng-fei, Wang Xue-zhi, Han Ying-shi. The Enhanced Genetic Algorithms for the Optimization Design [C] // 2010 3rd International Conference on Biomedical Engineering and Informatics. Vol. 7, 2010: 2990-2994
- [8] 浦兴成, 孙凯. 一种改进的自适应蚁群算法及其应用研究 [J]. *重庆邮电大学学报: 自然科学版*, 2011, 23(3): 331-335

应用到网格计算独立任务调度中。实验结果充分证明了该算法的有效性和稳定性。

## 参考文献

- [1] Sakthivel V P, Bhuvaneshwari R, Subramanian S. Artificial immune system for parameter estimation of induction motor [J]. *Expert Systems with Applications*, 2010, 37(8): 6109-6115
- [2] Dong W M, Zhou N, Paul J C. Robust tile-based texture synthesis using artificial immune system [J]. *Neural Computing & Applications*, 2009, 18(3): 223-235
- [3] 刘丽珏, 蔡自兴. 变异协同进化的免疫克隆算法 [J]. *控制与决策*, 2008, 23(6): 613-618
- [4] 李涛. 基于免疫的网络监控模型 [J]. *计算机学报*, 2006, 29(9): 1515-1522
- [5] 凌军, 曹阳, 尹建华, 等. 基于小生境技术的多样性抗体生成算法 [J]. *电子学报*, 2003, 31(8): 1130-1133
- [6] 张著洪. 人工免疫系统中智能优化及免疫网络算法理论与应用研究 [D]. 重庆: 重庆大学, 2004(6)
- [7] Zhu Rong-bo, Qin Ying-ying, Lai Chin-feng. Adaptive Packet Scheduling Scheme to Support Real-time Traffic in WLAN Mesh Networks [J]. *KSII Transactions on Internet and Information Systems*, 2011, 5(9): 1492-1512
- [8] Vijayalakshmi K, Radhakrishnan S. A novel hybrid immune-based GA for dynamic routing to multiple destinations for overlay networks [J]. *Soft Computing*, 2010, 14(11): 1227-1239
- [9] 郑日荣, 毛宗源, 罗欣贤. 改进人工免疫算法的分析研究 [J]. *计算机工程与应用*, 2003, 31(34): 35-37