

文章编号:1006-3080(2010)01-0088-05

基于定向突变的自适应并行免疫算法

翟仲曦, 钱 锋, 杜文莉

(化工过程先进控制和优化技术教育部重点实验室, 上海 200237)

摘要:针对克隆选择算法收敛速度较慢的问题,对算法策略进行研究,提出了一种基于定向突变的自适应并行免疫算法(APIA)。该算法采用自适应并行搜索策略,在记忆库中引入定向突变算子,增强算法的局部定向搜索能力,并改善算法早熟的问题。同时算法还改进了超变异算子,以提高其运行效率。仿真实验结果表明:该算法比克隆选择算法和传统的遗传算法有更好的寻优能力,有效地提高了收敛速度,缩短了搜索时间。

关键词:免疫算法;克隆选择;定向突变;并行搜索

中图分类号:TP18

文献标志码:A

Adaptive Parallel Immune Algorithm with Orthomutation

ZHAI Zhong-xi, QIAN Feng, DU Wen-li

(Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China)

Abstract: Aiming at the slow convergence problem of clonal selection algorithm, this paper proposes an adaptive parallel immune algorithm with orthomutation (APIA). By adopting adaptive parallel search strategy with orthomutation operator in the memory base, the APIA can strength the ability of local directed search and jump out of local optimization. In addition, this proposed algorithm improves hypermutation operator to increase operational efficiency of the algorithm. The simulation experiment results show that the APIA has better optimizing capacity than clonal selection algorithm and traditional generic algorithm, and effectively increases the convergence speed and shortens the search time.

Key words: immune algorithm; clonal selection; orthomutation; parallel search

免疫系统是一个由执行免疫功能的器官、组织、细胞和分子等组成的复杂系统,它能够识别自身异己抗原,并通过免疫应答排除抗原性异物,维持机体的生理平衡。免疫算法作为一种基于生物免疫系统的新兴优化算法,借鉴了生物体独特的免疫操作机制,具有记忆性、多样性、分布性、鲁棒性等特点,有效地提高了优化算法的搜索效率,自问世以来已取

得了较快的发展,成为优化算法领域的研究前沿。其中,de Castro^[1]和 Farmer^[2]分别提出了基于免疫克隆选择机理的克隆选择算法和基于网络模型的免疫算法,为免疫算法的发展拓宽了道路。此后,一些学者^[3]还提出了融合其他智能策略的改进免疫优化算法。但免疫优化算法的研究仍存在许多问题,如对免疫算法搜索策略的自适应性研究不足;免疫优

收稿日期:2009-01-06

基金项目:国家 863 计划课题(2007AA04Z159, 2008AA042902);上海市国际科技基金项目(08160710500);长江学者和创新团队发展计划(IRT0721);上海市重点学科建设项目(B504);国家自然科学基金(20876044)

作者简介:翟仲曦(1984-),男,吉林辽源人,硕士生,主要研究方向:智能计算、智能控制。E-mail:zhaizhongxi@sohu.com

通讯联系人:钱 锋, E-mail: fqian@ecust.edu.cn

势抗体局域搜索的缺乏导向性;现有的免疫搜索算子不够完善等,导致面对复杂的连续多峰值函数时,搜索效率低,收敛速度慢。

本文从克隆选择算法搜索策略的研究出发,引入定向突变算子,建立自适应并行搜索机制,提出了一种二进制编码基于定向突变的自适应并行免疫算法(APIA)。该算法基于定向突变理论和克隆选择算法中的超变异方法,通过定向突变算子和超变异算子的相互配合,将局部搜索和全局搜索较好的结合,有效地提高了优化算法的搜索效率和收敛速度。

1 基于定向突变的并行免疫算法

以 $X = \{X_1, X_2, \dots, X_t\}$ 为变量的连续对象最优化问题,其目标函数为:

$$F = f(X_i) = \min f(X_i), i = 1, 2, \dots, m, \\ X_i \in [X_{\min}, X_{\max}]$$

采用二进制编码方式,对变量 X_i 进行二进制编码为 $A = a_1, a_2, \dots, a_L, A \in I, a_j$ 为等位基因, $j = 1, 2, \dots, L, L$ 为变量 X_i 编码长度,集 I 称为抗体空间, f 为 I 上的正实数函数, f 称为抗体亲和度函数。一般将抗体编码分为 t 段,每段长度为 $l_i, L = \sum_{i=1}^t l_i$ 。对于二进制编码方式, X_i 的译码方式如下:

$$X_i = X_{\min} + \frac{X_{\max} - X_{\min}}{2^{l_i} - 1} \left(\sum_{j=1}^{l_i} a_j 2^{j-1} \right) \quad (1)$$

1.1 克隆选择算法策略

在克隆选择算法(CSA)^[1]中,主要通过超变异操作对抗体 X_i 进行群体进化, $i = 1, 2, \dots, n, n$ 为普通抗体的数量。抗体的超变异率为

$$p(X_i) = \exp \left[-\beta \frac{f(X_i)}{f_{\max}} \right] \quad (2)$$

其中: $p(X_i)$ 为抗体 X_i 的等位基因 a_j 的突变概率; $f(X_i)$ 为抗体 X_i 的亲和度; $f_{\max}$ 为当前普通抗体群的最大亲和度; β 为超变异比例算子,且 $\beta > 0$ 。

超变异算子对低亲和度抗体变异率较大,对高亲和度抗体变异率较小,所以从理论上讲,超变异操作可以兼顾全局搜索和局域搜索。但其针对高亲和度优势抗体的局域搜索能力较弱,且缺乏导向性,导致优良抗体不能及时快速地收敛,降低了搜索速度。因此在抗体搜索的过程中应引入针对高亲和度抗体的局部定向搜索机制,与超变异算子配合进行并行搜索,从而将全局搜索和局部搜索更好的结合,达到全局寻优的最佳效果。

但对于一些复杂的多峰值优化函数,搜索过程很容易陷入局部最优值。当群体最优值没有达到期望值,并且经过一定代数的搜索,群体最优值仍保持不变,便可判定搜索陷入局部最优^[4]。跳出局部最优值的有效方法是增加群体的多样性,提高搜索速度。此时对优势抗体的局部定向搜索不利于群体跳出局部最优解,占用搜索机时,降低搜索效率。因此当搜索陷入局部最优时,应暂时放弃局部定向搜索,充分发挥超变异算子与免疫网络机制的优势,增加抗体的多样性,待跳出局部最优之后,再恢复并行搜索机制。

为此,引入搜索机制阈值 V_s 。抗体最优值在一代搜索中没有变化,则局部最优代数 G 加 1,否则 G 清零。

算法的搜索机制如下:

$$\begin{cases} G \leq V_s, \text{ 群体搜索为并行搜索} \\ \quad \begin{cases} \text{普通抗体全局搜索} \\ \text{优势抗体局部定向搜索} \end{cases} \\ G > V_s, \text{ 判定陷入局部最优值, 只进行普通抗体全局搜索} \end{cases}$$

1.2 定向突变算子

定向突变(又称适应性突变)即生物基因受环境和自身经验直接影响而诱发的具有方向性的突变,是一种结合了确定性和随机性的遗传过程,有利于种群局部特征的进化。虽然至今定向突变理论只是个别实验的支持,并未得到普遍认可^[5],但可借用其局部定向进化的特点,将定向突变过程引入到免疫算法中,增强优势抗体的局部定向搜索能力,加快优势抗体的收敛速度。此外,根据生物免疫系统的记忆机理,在免疫算法中引入记忆库群体,用以存放高亲和度的优势抗体,并为群体更新提供新的抗体,防止免疫最优抗体退化和种群退化。综上所述,可对记忆库中的优势抗体进行定向突变操作。

对记忆库群体 $M = \{X_1, X_2, \dots, X_m\}$ 进行克隆, m 为记忆库抗体数量。对记忆库抗体进行克隆,每个抗体的克隆数为 c ,产生的群体 $M_c = \{X_{11}, X_{12}, \dots, X_{1c}, X_{21}, \dots, X_{mc}\}$,总克隆规模为 $m' = mc$ 。定向突变即对克隆群体 M_c 中每个抗体二进制编码 A 从最低位到第 q 位随机选择一位进行突变,产生群体 M'_c ,其规模为 m', q 的取值依据式(3)进行。

$$q = \text{int}[L \cdot \exp(-(V_s - G)/V_s)] \quad (3)$$

其中 int 为取整函数。对群体 M'_c 与 M 中的相应克隆母体进行比较,择优选取,产生新的记忆库 $M' = \{X'_1, X'_2, \dots, X'_m\}$ 。

定向突变可对记忆库优势抗体的局部可变邻域

进行随机搜索,突变范围与局部最优代数 G 有关。随着 G 增加,抗体对邻域数值的搜索范围也逐渐增大。定向突变既有利于记忆库优势抗体的快速收敛,提高搜索效率;同时,又可在局部最优值出现的初期,通过搜索领域的逐渐扩大,协助搜索跳出局部最优。定向突变算子将导向性、随机性和适应性相结合,对记忆库中的优势抗体进行定向的局部随机搜索,完善了免疫算法的搜索机制,进而提高算法的收敛速度。

1.3 改进的超变异算子

超变异操作是克隆群体抗体增加多样性的主要途径。对于二进制编码的免疫克隆算法,超变异率是针对抗体每一个二进制位的变异概率。每个抗体都有独立的超变异率,且与其自身的亲和度成反比^[7]。抗体 X_i 的超变异率 $p(X_i)$ 如式(2)所示, β 越大, $p(X_i)$ 随 $f(X_i)$ 的增加而减小的幅度越大。理论上,超变异率可以根据抗体的亲和度动态非线性地调整大小,既可搜索到全部峰值,又可以保持群体稳定收敛的状态。对于中低亲和度抗体的 p 不能太大,否则群体的稳定性很差,搜索群体难以长时间稳定;但同时高亲和度抗体的 p 也不能太小,否则操作效果不明显,群体收敛缓慢。

由于 $\exp$ 为非线性函数,并不能兼顾上述的要求,同时超变异算子是增加抗体多样性的主要手段,因此操作过程应首先保证中低亲和度算子有较大的变异率,这样使得高亲和抗体的超变异率过小,不能及时收敛,增加了搜索时间。因此,为保证高亲和度抗体有合适的超变异率,以实现抗体群的局部搜索,在超变异算子中引入超变异率下限值

$$p_{\min} = \frac{\gamma}{L} \quad (4)$$

其中: γ 为限值系数, $0.5 < \gamma < 1$; L 为二进制编码长度。

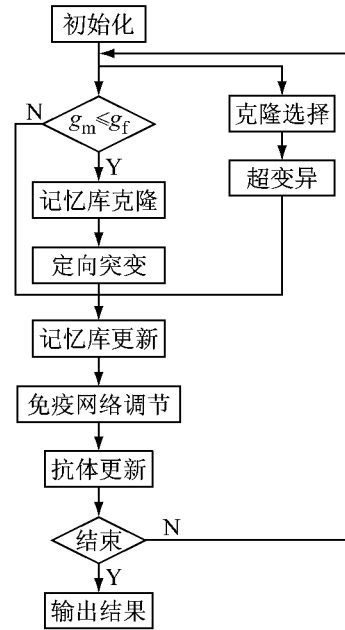
改进后的超变异操作为

$$p(X_i) = \begin{cases} \exp\left[-\beta \frac{f(X_i)}{f_{\max}}\right], & p(X_i) \geq p_{\min} \\ \frac{\gamma}{L}, & p(X_i) < p_{\min} \end{cases} \quad (5)$$

$p_{\min}$ 的引入克服了高亲和度抗体因超变异率过小而使搜索停滞的缺点,保证了群体进行有效的局部搜索,进而提高了算法的收敛速度和搜索效率。

1.4 算法设计

基于定向突变的自适应并行免疫算法(APIA)采用二进制编码,算法流程如下:



算法步骤如下:

(1) 初始化。当前进化代数 $k=0$, 在解域随机产生初始 n 个抗体构成初始种群 $X(0) = \{X_1(0), X_2(0), \dots, X_n(0)\}$, 计算抗体群 $X(0)$ 的亲和度, $\{f(X(0))\} = [f(X_1(0)), f(X_2(0)), \dots, f(X_n(0))]$, 选择 m 个具有高亲和的抗体, 构成初始记忆库 $M(0) = \{X_1(0), X_2(0), \dots, X_m(0)\}$, 设定算法参数。

(2) 判断 $G \leq V_s$, 是则继续, 否则转入步骤(4)。

(3) 定向突变。对当前记忆库优势抗体 $M(k) = \{X_1(k), X_2(k), \dots, X_m(k)\}$ 进行定向突变, 操作过程如前文所述, 构成成熟记忆库 $M'(k) = \{X'_1(k), X'_2(k), \dots, X'_m(k)\}$ 。

(4) 超变异操作。按照克隆数量与亲和度正比的关系对普通抗体群 $X(k) = \{X_1(k), X_2(k), \dots, X_n(k)\}$ 进行克隆复制, 克隆总数为 n , 组成新的抗体群 $X_c(k)$, 依据式(5)对 $X_c(k)$ 进行超变异操作, 形成新的抗体群 $X'(k)$ 。

(5) 更新记忆库。根据抗体亲和度在群体 $M'(k)$ 和 $X'(k)$ 中重新选择 m 个抗体, 构成新的抗体库 $M(k+1)$, 并保留历代最佳抗体 $X_{\text{best}}(k)$ 。

(6) 免疫网络调节。对 $X'(k)$ 中的抗体按激励水平的大小除去 n_d 个低激励水平的抗体, 剩余抗体集合为 $X_s(k)$, 规模为 n_s 。抗体 X_i 的激励水平定义为^[7]

$$SL_i = \frac{1}{2} \left[D_{\text{avg}}(X_i) - k_1 S_{\text{avg}}(X_i) + \frac{f(X_i)}{f_{\max}} \right] \quad (6)$$

其中: D_{avg} 和 S_{avg} 分别为 X_i 与网络中所有抗体相比的基因型平均相似度和相异度; k_1 为抑制因子, $k_1 \in [0, 1]$, 且 k_1 越大, 抗体相似性抑制抗体 X_i 繁

殖的作用就越强。

(7) 新抗体的补充。在记忆库中选择 n_d 个抗体,与 $X_s(k)$ 组成新一代抗体群 $X(k+1)$ 。

(8) 判断终止条件,输出记忆库 M 中的最佳抗体,否则返回到步骤(2)。

2 仿真实验

由于 APIA 是基于对克隆选择算法(CSA)搜索策略的研究而提出的优化算法,因此首先将 APIA 与 CSA 进行仿真比较,同时为了便于分析定向突变和改进超变异算子的有效性,在 CSA 中也引入网络调节机制,以提高算法的群体多样性。

测试选用优化函数 $\min f_1(x) = \sum_{i=1}^3 x_i^2, -10 \leq x_i \leq 10$ 。

f_1 又称 Sphere 函数,为单极值的二次函数,可检验优化算法的收敛速度和搜索能力。该函数只有一个全局最小点(0,0,0),最小值 $f_{\min}=0$ 。

优化 f_1 函数时,算法均采用二进制编码,编码长度 $L=18$ 。在 APIA 中,初始种群规模 $n=40$,记忆库优势抗体克隆数 $c=2$,搜索机制阈值 $V_s=20$,限值系数 $\gamma=0.8$ 。CSA 的种群规模为 40。为了便于比较,两种算法的记忆比例 $\alpha=0.25$,抑制因子、超变异系数均为 $\beta=2.5, k_1=0.5$,最大进化代数 100。仿真结果见图 1。

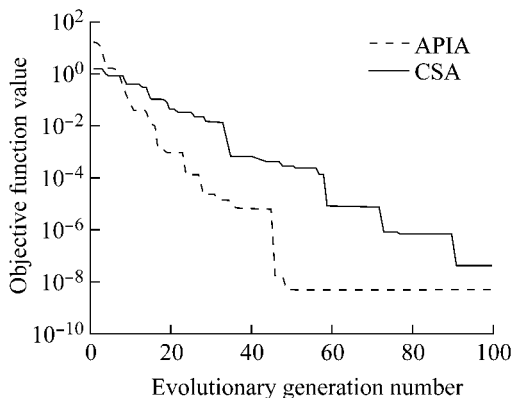


图1 f_1 函数优化结果

Fig.1 Optimization result of the function f_1

从图1的优化结果可以看出,APIA 的收敛速度和搜索精度优于 CSA,进而验证了 APIA 的改进策略增强了免疫算法局部搜索能力。

进一步采用 APIA 优化 3 个典型的多峰值测试函数 Alpine、Griewangk 和 Schaffer,并与标准遗传

算法(SGA)^[6]和克隆选择算法(CSA)进行比较,以验证 APIA 对复杂函数的优化性能。

Alpine 函数:

$$\max f_2(x) = \prod_{i=1}^2 \sin(x_i) \sqrt{\prod_{i=1}^2 (x_i)}, \\ 0 \leq x_i \leq 10$$

Griewangk 函数:

$$\min f_3(x) = \sum_{i=1}^2 \frac{x_i^2}{4000} - \prod_{i=1}^2 \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1, \\ -50 \leq x_i \leq 50$$

Schaffer 函数:

$$\min f_4(x, y) = \frac{[\sin \sqrt{x^2 + y^2}]^2 - 0.5}{[1 + 0.001(x^2 + y^2)]^2} - 0.5, \\ -10 \leq x, y \leq 10$$

优化函数 f_2 在(7.917,7.917)有一个全局最大值 7.885 6; f_3 在(0,0)处有一个全局最小值 0; f_4 是一个复杂的多峰值优化函数,在(0,0)处有全局最小值-1,该函数在距离全局最优值大约 3.14 范围内存在无穷多个取值为-0.990 283 的局部极小值,并且函数强烈振荡,一般很难一次得到全局最优解。

APIA 和 CSA 的参数设置不变,SGA 的初始种群为 40,采用单点交叉,交叉概率为 0.8,突变概率为 0.025。同时设置算法的最大进化代数为 1 000,当搜索的最优函数值与理想极值的误差小于 10^{-3} 时认为算法收敛并停止进化。对 3 个测试函数分别用 3 种算法各运行 30 次,并记录平均每次的搜索时间 AvgT、搜索到最优解所需的平均迭代次数 AvgG 和搜索到最优解的次数 BestN,比较结果见表 1。

从表 1 可以看出,与 APIA 和 CSA 相比,由于 SGA 的操作比较简单,平均搜索时间较短,但是其搜索精度低,不能满足函数寻优要求。APIA 在搜索时间和寻优性能上都要好于 CSA,特别是对 f_3 和 f_4 的多峰值复杂函数效果比较明显。进一步说明了 APIA 中基于定向突变的并行机制和改进的超变异算子有效地提高了免疫算法的收敛速度和寻优能力,在收敛精度和搜索时间中找到了一个较好的平衡点。

表 1 算法收敛性能比较
Table 1 Convergence comparison of algorithms

Function	APIA			CSA			SGA		
	AvgT	AvgG	BestN	AvgT	AvgG	BestN	AvgT	AvgG	BestN
f_2	0.79	19	30	1.37	23	30	3.37	173	21
f_3	4.46	90	30	14.28	209	30	6.27	171	15
f_4	11.90	171	29	24.79	312	27	5.41	271	9

3 结束语

本文首先对克隆选择免疫算法策略进行分析与研究,然后结合生物学定向突变理论,引入定向突变算子,并改进了克隆选择算法中的超变异算子,进而提出了基于定向突变的自适应并行免疫算法。该算法针对不同阶段的情况搜索,动态地调整搜索策略,增强了优势抗体局部搜索的导向性和普通抗体的局部搜索能力,改善了免疫算法易陷入局部最优的缺陷。仿真结果表明:APIA 具有很好的收敛性能,不仅缩短了算法的搜索时间,较好地解决了免疫算法的局部收敛问题;而且通过定向突变算子和超变异算子的相互配合,确保了全局搜索性能和局部搜索性能的动态平衡。但本算法主要通过大量的实验来确定搜索机制阈值,如果能通过免疫疫苗的形式自动确定阈值,则算法的自适应性可进一步加强。

参考文献:

[1] de Castro L N, Zuben F J V. The clonal selection algorithm with engineering applications [C]//Proceedings of Genetic and Evolutionary Computation Conference. Heidelberg, Berlin; Spring-Verlag, 2000; 36-37.

[2] Farmer J, Packard N, Perelson A. The immune system,

adaptation and machine learning [J]. Physical D,1986, **22**, 187-204.

[3] Cairns J. The origins of mutants [J]. Nature,1988,**335**;142-145.

[4] 高波,张忠能. 遗传编程局部最优化问题分析[J]. 计算机应用与软件, 2003, **20**(12); 86-88.

[5] 陈建华. 是否要重新评价拉马克[J]. 自然杂志,1999, **21**(1); 52-57.

[6] Holland J H. Genetic algorithms and classifier systems: Foundations and future directions [C]// Proceedings of the Second International Conference on Genetic Algorithms on Genetic Algorithms and Their Application. Hillsdale NJ, USA;L Erlbaum Associates Inc, 1987;82-89.

[7] 何宏,钱锋. 基于网络理论的动态超变异免疫算法[J]. 华东理工大学学报:自然科学版, 2007, **33**(3); 429-435.

[8] Jang-Sung, Hyun-Kyo Jung, Song-Yop Hahn. A study on comparison of optimization performances between immune algorithm and other heuristic algorithms [J]. Magnetics, IEEE Transaction,1998, **34**(5); 2972-2975.

[9] 何宏,钱锋. 一种新的自适应免疫进化算法[J]. 信息与控制, 2006, **34**(1); 34-38, 46.

[10] Dai Yongshou. Adaptive immune-genetic algorithm for global optimization to multivariable function [J]. Systems Engineering and Electronics,2007, **18**(3); 655-660.

[11] Rabbani M, Panahi H. An efficient hybrid artificial Immune algorithm for clustering [C]//Hybrid Intelligent Systems, 2008. Washington D C; IEEE Computer Society, 2008; 374-379.