

改善收敛早熟的混合遗传算法^①

李险峰, 董绍华

(北京科技大学 机械工程学院, 北京 100083)

摘要: 针对传统遗传算法收敛早熟问题, 在传统包含模拟退火的混合遗传算法的基础上, 设计加入了“包含浓度均衡措施的复制算法”, 通过调整轮盘赌扇区面积, 防止个体适应度的两极分化。从而避免了算法过早收敛于局部最优解; 同时通过一个工程实例计算验证了算法的可行性。

关键词: 收敛早熟; 模拟退火; 混合遗传算法;

Hybrid Genetic Algorithm to Improve Premature Convergence

LI Xian-Feng, DONG Shao-Hua

(Mechanical Engineering College, USTB, Beijing 100083, China)

Abstract: For the premature convergence in traditional Genetic Algorithm, this paper comes up with a new enhanced algorithm which introduced new strategy called “equalization measurement on the concentration of the replication”, together working with the traditional hybrid SA-GA. Through dynamic adjusting on roulette sector to prevent polarization of individual fitness, so that avoiding premature convergence of the algorithm to local optimal solution; Mean while validated the feasibility of algorithm with one engineering example.

Key words: premature convergence; simulated annealing; hybrid genetic algorithm

1 引言

遗传算法 (Genetic Algorithm, 简称 GA) 是上世纪 70 年代由美国 Holland 提出的仿生物进化过程的优化算法, 其主要思想就是算法过程中引入“生物进化论”和“遗传学”中的“物竞天择、适者生存”的原则。由于遗传算法并行度高、问题适应度高, 而且可以获得全局最优解。因此, 几十年来在智能算法领域的发展迅速, 成为研究的热点。

GA 将问题表述为染色体, 将预定义的种群空间 (N) 中一群染色体置于问题的环境中, 根据生物进化论的原则机理, 从中选择出适应环境的染色体进行复制、再通过交叉、同时按照一定的变异控制参数 P_c 和变异概率 P_m 变异生成更能适应环境子代的染色体群 $mutpop_i(k)$, 并通过适应度函数进行检验筛选, 通过选择概率 s_i 确定下一代染色体, 这样一代一代进化下去, 直至求得问题的最优解。遗传算法作为一种通用

的自适应随机搜索算法, 一直被两个主要的问题所困扰: 算法早熟和收敛速度慢。有学者^[1]认为, GA 中的两个因素: “种群多样性”和变异控制参数 P_c 或者“选择压力系数”是促使算法趋于早熟的主要原因。过大的选择压力系数虽然可以加速算法的收敛速度, 却同时加速了种群中适应度值不利于问题求解的个体的迅速“淘汰”, 种群的多样性被破坏, 从而使得算法搜索空间萎缩, 导致算法趋向局部最优; 但是如果降低选择压力系数, 算法在计算效率和速度将会大打折扣。学者徐宗本^[2]等在“遗传算法过早收敛现象特征分析及其预防”中提出: 交叉算子在搜索过程中存在严重的成熟化效应, 在搜索的同时, 不可避免地使种群多样性趋于零, 收敛到局部最优解。因此, 为了兼顾种群多样性和选择压力系数, 必须在种群多样性和选择压力系数之间找到一个均衡点: 即在保证种群多样性的前提之下, 尽可能提高选择的压力系数, 使算法尽快

① 收稿时间:2011-02-20;收到修改稿时间:2011-04-07

收敛到最优解。

2 包含多种算法因子的混合遗传算法

如同遗传算法的思想来自进化论启发,类似地受到物理退火相变过程中能量趋于稳定的启发,1953年Metropolis等提出了模拟退火算法(即用概率接受新状态,所谓Metropolis准则)的基本思路^[3]。具体而言:在温度 t ,由状态变化到状态 j ,两者的能量为 E_i 与 E_j ,若 $E_j < E_i$,则接受该状态;否则,若概率 P_r 大于 $[0,1]$ 之间的随机数则仍然接受新状态 j 为当前状态;若不成立则依旧保持为当前状态。其中:

$$P_r = \exp\left[-\frac{E_j - E_i}{kt}\right] \quad (k \text{ 为波尔兹曼常数});$$

当这种状态多次重复后,系统能量将趋于一个比较平衡的状态(能量较低状态)。为了弥补单一传统遗传算法效率和质量上的局限,越来越多的尝试是将多种不同的智能算法思想引入遗传算法,比如将模拟退火的思路、蚁群算法等其他智能算法融入传统遗传算法结构。本文中作者提出了一种包含模拟退火机制的改进遗传算法。一方面通过修改的选择算子和变异算子来调节选择压力系数,在传统算法的变异(Mutation)阶段和选择阶段(Selection)上影响下一代新个体的产生,提高搜索的效率;同时设计了“含浓度均衡措施的选择复制”因子 β 。在每一代子代染色体的选择轮盘空间上进行选择系数的均匀化,调整轮盘扇区的选择面积,尽量均匀化较好个体在轮盘扇区的分布,避免个体子代的适应度出现传统的两极分化,这是区别于传统遗传算法的一个很重要的不同点之处。通过兼顾子代种群空间和压力系数两个重要的算法指标,实现算法在效率和质量上的平衡。

2.1 混合算法的理论收敛性描述

包含模拟退火(SA)机制的混合遗传算法(HGA)依然属于基于概率机制的优化算法。不同的是传统SA算法是通过赋予在搜索过程中时变而且最终趋于零的概率突变来避免陷入局部最小值并最终趋向于全局收敛;而GA则是通过概率意义的“优胜劣汰”的思想来实现全局优化,两种方法的融合可以实现丰富和提升搜索过程的能力和效率。SA算法是串行的搜索过程而GA算法则是群体并行搜索,两则结合可以赋予SA成为并行的SA,提高搜索性能;同时SA的概率突跳性变异操作,增强了GA的进化功能。

理论上^[3]已经证明SA算法和带有保优操作的GA具有概率1的全局收敛性。为了实现理论上严苛的收敛条件,两者单独而言的优化效率和质量均不是很理想。SA算法需要漫长的退火时间;GA算法需要大量的种群空间和较长的遗传步数来实现。因此,混合算法中的SA的嵌入,是对于GA变异操作的一种有效补充,赋予优化过程中的可控的概率突跳特性,尤其是在高温状态下的突跳性成为跳出局部最优的利器,从另外一个侧面降低了GA算法对于参数过分依赖。混合策略使得传统的SA实现并行的过程搜索,搜索效率得以大幅度提升。

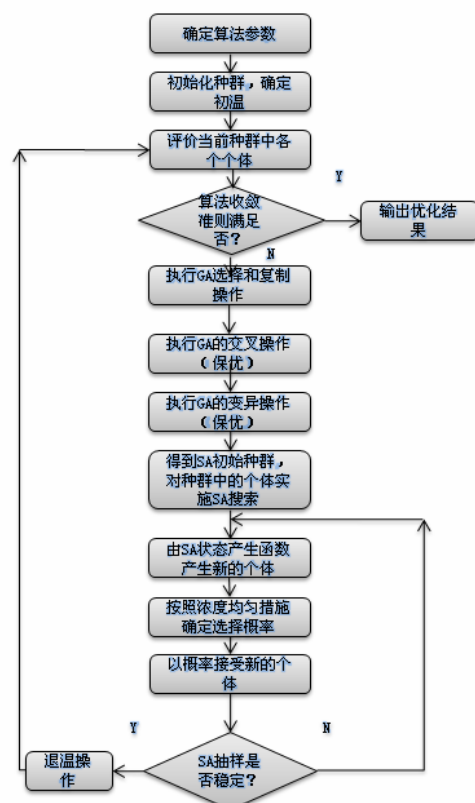


图1 混合算法流程框图

2.2 混合算法流程描述

基于上文的理论思路,混合算法的结构流程框图如图1所示,包含了遗传算法和模拟退火的两部分操作。

① 首先初始化种群,确定最初的系统温度;

② 评价当前的种群中的个体,满足算法收敛准则则输出结果;如果不满足则按照遗传算法进行复制操作;

③ 交叉和变异，按照预设的变异因子概率 P_m ， $P_m \in [0,1]$ 进行交叉变异；

④ 得到 SA 的初始种群，进行按照模拟退火算法思路的筛选。

1) 在第 k 代种群 $pop(k)$ 和当前的温度 t_k 下，访问种群里的每一个个体 $pop_i(k)$ ；

2) 对于每一个个体，随机选择变异位，按照一定的变异控制系数进行变异（也就是在其邻域中随机选择一状态）；其中变异控制系数和染色体配选路径的总数成正比， α 为比例常数； m 是染色体队列位数总数；也就是说变异控制系数控制整体染色体队列中参与变异的位数。

$$P_{d(i)} = \alpha \frac{P_i}{\sum_{j=1}^m P_j} \quad (1)$$

3) 对于新个体 $mutpop_i(k)$ ，计算其适应度值，按照 Metropolis 概率^[4] 准则选择接受或者放弃这个个体，如果接受则更新，否则放弃，循环 N 次；

⑤ 为了避免种群个体适应度的两极分化，保持种群的多样性。具体的步骤如下：

1) 对群体 $pop(k)$ 的每一个个体 $pop_i(k)$ ，结合当前染色体个体的退火温度，设定并计算其适应度函数 f （见式 2）

$$f_i = \frac{\text{cost}_{\max} - \text{cost}_i}{(\text{cost}_{\max} - \text{cost}_{\min})t_k} \quad (2)$$

2) 综合计算轮盘赌扇区面积，并用浓度调整系数 β 校正（可以调整的压力系数），这样每一个子代的优良染色体出现的概率将在整个轮盘赌扇区上最大化均匀出现（见式 3）

$$s_i = \left(\frac{f_i}{\sum_{j=1}^N f_j} \right)^\beta \quad (3)$$

3) 这样，校正后的扇区面积（选择概率 s_i ）为：（见式 4）

$$s_i = \frac{s_i}{\sum_{j=1}^N s_j}, i=1,2,\dots,N \quad (4)$$

4) 按照调整后的选择概率 s_i 从 $pop(k)$ 中选择 M 个个体作为下一代种群；

5) 进行循环；

实验算例

结合一个实际的工程算例，作者分析了不同的优化控制参数对比下的收敛过程。某型号汽车仪表盘安装的操作内容的现场经验，其 16 步主要的操作内容的某一种先后顺序如图 2 所示。

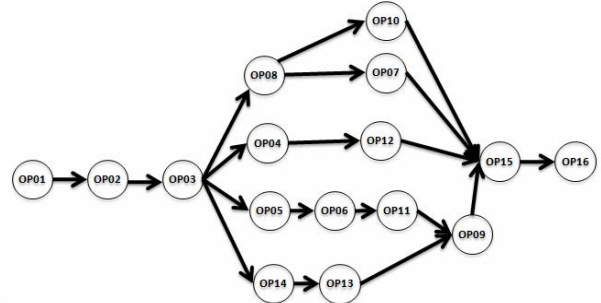


图 2 装配顺序约束描述

作者使用了通用计算分析工具 MATLAB，进行算法的编制和实现。

M	进化迭代次数
N	种群规模，必须取偶数
P_m	变异概率
P_d	变异程度控制参数
K	状态转移次数
t_0	初始温度
α	降温系数

图 3 算法主要参数描述

算法中主要的参数描述如上图（见图 3）：

种群规模和变异概率作为控制算法效率和优化质量的两个重要的参数。经对比分析之后可以发现：

1) 种群数 N 并不是越大越好，种群数目越大，搜索计算时间相对延长； N 取值 30 时搜索的效率最高（见图 4）；

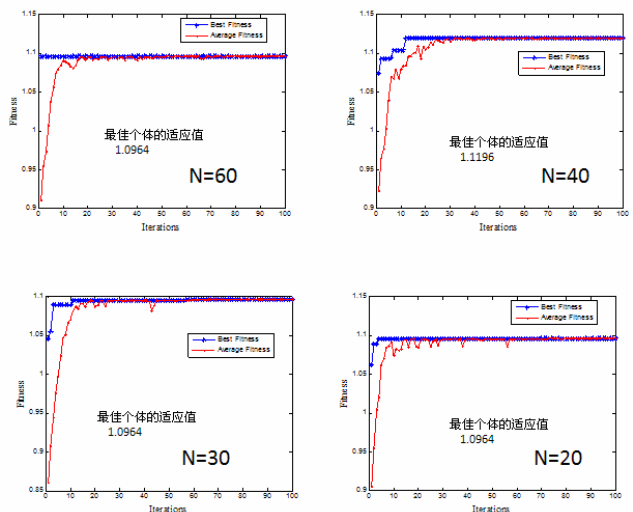


图 4 种群规模参数对于收敛的影响

2) 变异概率 P_m 在算法中是一个决定子代个体是否参与变异的预设值(阈值);当取值为(0,1)的随机数 rand 大于 P_m 时候个体才进行变异操作,进入退火程序。算法实验得到: P_m 取值在 0.5~0.65 时,算法的收敛速度较好(见图 5)。

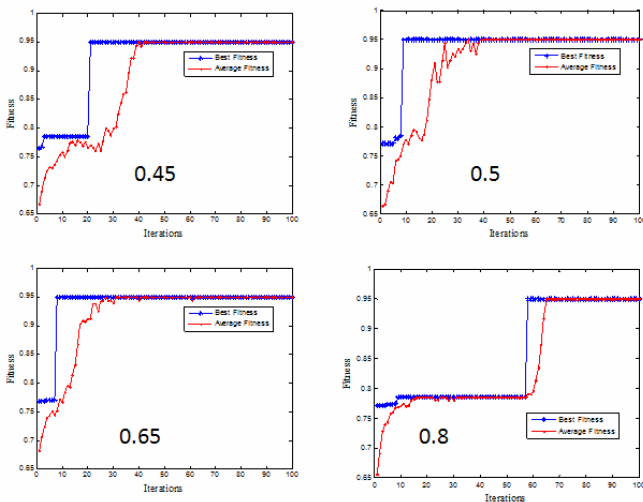


图 5 变异概率对于收敛的影响

3) 选择 $N=30/P_m=0.6$ 组合时候的分析优化结果如下图 6 所示。

3 结论

本文针对遗传算法过程中易出现的早熟现象,提

出了包含模拟退火思路的混和优化算法,并针对早熟的原因设计了均衡浓度因子的选择复制步骤以确保种群空间和选择压力的平衡;通过实验证明了这种算法的正确性和有效性。

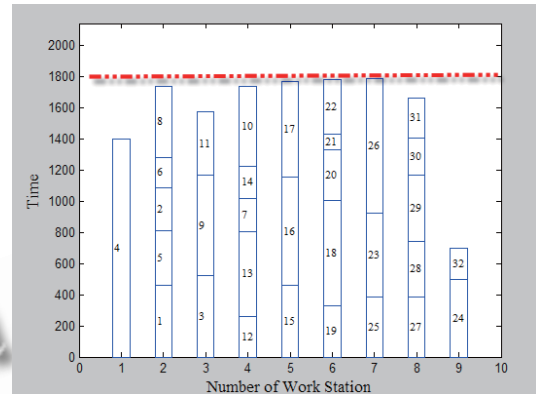


图 6 平衡优化结果

参考文献

- 1 Whitely D. The genetic algorithm and selection pressure; Why rank-based allocation reproduction trials is best. Los Altos: Morgan Kaufmann Publishers, 1989.
- 2 徐宗本,等.遗传算法过早收敛现象的特征分析及其防治.中国科学,1996,26(4):364-375.
- 3 王凌.智能优化算法及其应用.北京:清华大学出版社,2004. 135-136.
- 4 陈永卿,潘刚,李平.基于混合遗传算法的装配线平衡.机电工程,2008,25(4):60-62.

(上接第 238 页)

动、运行和互相通讯的思维是至关重要的。

3 结论

本文通过论述在 B/S 下,API 语句调用 Dll 程序库模块,实现数据在客户端与服务器端的数据传送,将代码写在 Dll 中,再对数据进行加密处理,数据信息不易被盗取,避免了写在 Asp 中的程序代码因为以文本形式显示,而被入侵者通过非法手段获得程序意图,保证了数据传送过程中的安全性,提高了程序的健壮性和稳定性;Asp 脚本语言编程相对于 VB6.0 编程代码的功能相对较弱,这个系统模型则强化了程序代码的功能,使以前 C/S 模型遗留的程序模块、自制的独

立控件、公用程序也可通过上述模型实现复用;实现简单灵活。在我们编写的菌种资源数据库中得到很好的应用。程序的编制可以用 VB6.0 也可以用 C++、JAVA 等多种语言,数据库可以用 Access 也可以用 SQLServer、Oracle 等大型数据库。

参考文献

- 1 蒲鹏,陈慧.ASP 网络编程基础.北京:清华大学出版社.北京交通大学出版社,2009.1-27.
- 2 郭招娣.API 函数在 VB 中的应用.计算机应用与软件,2003,20(9):83-85.
- 3 彭玉忠,王金才,郝荣霞.基于 B/S 的在线考试系统答案更新算法.计算机系统应用,2008,17(4):117-119.