

# 自适应记忆遗传算法研究

赵越<sup>1</sup>, 徐鑫<sup>1</sup>, 赵焱<sup>1</sup>, 初雪宁<sup>2</sup>

(1. 渤海大学 大学计算机教研部, 辽宁 锦州 121013;

2. 东北大学 信息学院, 辽宁 沈阳 110819)

**摘要:**针对遗传算法优化过程中仍然存在许多问题,文中提出了一种新的自适应记忆遗传算法。引入基因库的概念,用以存储重复出现个体的基因编码和对应的适应度值,进而解决重复个体适应度值的重复计算问题;利用 Logistic 曲线方程对遗传算法的交叉概率和变异概率进行自适应调整;以 TSP 为应用背景对文中算法进行实验,结果表明文中算法有效减少了算法的时间复杂度,其加速比能够达到 49.70% 左右。在算法的收敛性方面,改进后的算法收敛速度快于基本遗传算法,其所得解与 TSPLIB 提供的最优解的平均相对误差最大不超过 9.38%。

**关键词:**记忆遗传算法;基因库;自适应;函数优化;旅行商问题

中图分类号:TP301

文献标识码:A

文章编号:1673-629X(2014)02-0063-04

doi:10.3969/j.issn.1673-629X.2014.02.015

## Research on Adaptive Memory Genetic Algorithm

ZHAO Yue<sup>1</sup>, XU Xin<sup>1</sup>, ZHAO Yan<sup>1</sup>, CHU Xue-ning<sup>2</sup>

(1. Teaching and Research Institute of College Computer, Bohai University, Jinzhou 121013, China;

2. College of Information Science & Engineering, Northeastern University, Shenyang 110819, China)

**Abstract:** There are still many problems in the optimization process of genetic algorithm. Propose a new adaptive memory genetic algorithm. The gene warehouse with appropriate scale is provided in thesis which is used to store gene encodings and individual fitness value of chromosomes repeated. The problem of repeatedly calculating individual fitness value of repeated chromosome is effectively solved through the previous method. The Logistic curve equation is applied to change crossover probability and mutation probability for improving algorithm's adaptability. Adopt typical TSP as application background to test. Test results show that the algorithm proposed can effectively reduce the time complexity of genetic algorithm and its speed up to about 49.70% than original algorithm. On convergence of the algorithm, its convergence speed is faster than simple genetic algorithm. Also, mean relative error of optimal solutions solved with the improved algorithm relative to optimal solution provided by TSPLIB is no more than 9.38%.

**Key words:** memory genetic algorithm; gene warehouse; adaptive; function optimization; TSP

## 0 引言

遗传算法(Genetic Algorithm, GA)是一种模仿生物自然进化过程的、自适应启发式的全局优化算法<sup>[1]</sup>。由于在利用遗传算法求解问题时仅需要很少的辅助信息,容易与其他领域的知识相结合,且具有较好的自适应性和并行性,使得遗传算法在组合优化、机器学习、自适应控制、人工生命等领域得到了广泛的应用<sup>[2]</sup>。

遗传算法的收敛性等理论成为了理论研究的热点<sup>[3]</sup>。虽然国内外都很重视遗传算法的理论和应用研究,并且取得了很多令人瞩目的研究成果,但是遗传算法的理论和方法还尚未成熟,如易发生早熟收敛、局部

搜索能力差、计算量大等,还需要进一步改进和完善<sup>[4-6]</sup>。

鉴于此,文中对遗传算法中的适应度值的求解过程进行改进,提出适当规模的有序基因库概念,降低遗传算法的计算复杂度,提高算法的性能。此外,在记忆遗传算法的基础上,引入自适应调整函数,以加快算法的收敛速度,提高算法性能。

最后应用改进后的遗传算法实现旅行商问题(Traveling Salesman Problem, TSP)的求解,对文中算法的性能进行进一步验证等。

收稿日期:2013-03-21

修回日期:2013-06-27

网络出版时间:2013-11-29

基金项目:国家自然科学基金资助项目(60970157)

作者简介:赵越(1979-),女,辽宁抚顺人,讲师,博士生,研究方向为人工智能、增强现实;徐鑫,教授,研究方向为人工智能。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20131129.1020.052.html>

## 1 基于基因库的遗传算法的改进

### 1.1 遗传算法

遗传算法是一种将生物进化引入到数学模型中的随机搜索算法,本质是一种求解问题的高度并行性全局搜索算法,具有隐含并行性和全局搜索性的特点<sup>[7]</sup>。

在遗传算法优化过程中,由于随着算法进化代数的不断增加,算法最终近似收敛于最优解,而这使得种群中基因相同的个体也不断地重复出现,导致出现个体适应度值的重复计算,特别是在用遗传算法解决复杂问题时,容易引起资源的浪费。

### 1.2 记忆遗传算法

针对遗传算法优化过程中,基因编码相同的个体重复出现的现象,文中引入了基因库 (Genetic Warehouse, GW) 的概念,对基本遗传算法中个体适应度的求解过程进行改进。对于基因库中已存在的重复个体直接从基因库中获取其相应的适应度值,从而省去了重复个体的解码、计算目标函数值及由目标函数值转化为个体适应度值等操作,以实现算法优化过程中重复个体适应度值的快速求解,降低算法的计算复杂度,以节省系统资源,提高算法的运行效率,其基本框架如图 1 所示。

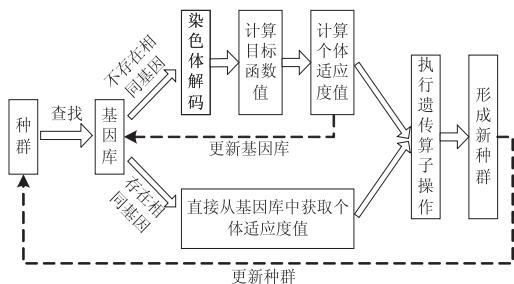


图 1 记忆遗传算法基本框架图

基因库:用于存储遗传算法优化过程中出现的较优个体的染色体编码及其对应的适应度值,并根据适应度值由大到小的规则排列集合。基因库的大小根据种群的规模进行合理的选取,一般情况下,基因库规模的取值为 0.1 ~ 0.2 倍的种群规模的大小。

遗传算法引入基因库后,对遗传算法中个体适应度的计算过程进行改进。首先,根据当前种群中的染色体的适应度值,构造初始基因库,将适应度值较优的染色体的编码和对应的适应度值添加到基因库中,并根据适应度值按照从大到小的规则进行排序。然后,在后代的算法优化中计算个体适应度时,先在基因库中查找是否存在基因编码相同的个体,如果存在则直接以基因库中的对应个体的适应度值作为当前个体的适应度值;如果不存在则根据适应度函数进行计算,并根据计算得到的适应度值与基因库中的适应度值进行比较,如果比基因库中最小的适应度值小则放弃入库

操作;否则将该个体的染色体编码和适应度值加入到基因库中,如果基因库中的规模达到了规定大小,则用新个体的信息替换掉基因库中适应度值最小的个体的信息,对基因库进行实时更新。此外由于遗传算法在进化前期,种群中个体的变化较大,重复率较低,所以可以选择在算法进行到一定优化代数后再进行基因库的构建工作,以尽量减少前期基因库利用效率过低的情况。

记忆遗传算法比基本遗传算法的求解过程要复杂一些,但是由于引入基因库后,使得算法具有记忆性,对于算法优化过程中出现的一些重复个体可以直接从基因库中获取其适应度值,而无需再利用适应度函数等操作重新计算。当目标函数越复杂时,越能体现出发挥记忆遗传算法的优越性能。

## 2 基于自适应遗传算法的改进

### 2.1 自适应遗传算法

从生物进化角度来讲,遗传算法虽然考虑了种群对环境的适应能力的模拟,但是却忽略了种群在进化过程中个体发育和遗传行为对环境变化的自适应特性,从而影响了基本遗传算法的性能和效率<sup>[8]</sup>。由于基本遗传算法采用固定的控制参数,无法满足遗传算法进化过程中参数动态变化的要求,尤其是交叉概率和变异概率的自适应调整。基本遗传算法的性能评价是一个复杂的过程,要选择恰当的固定参数是很困难的,对于不同的实际问题,其得出的结论和方法也是不同的。针对这些问题,在基本遗传算法的基础上产生了一种自适应遗传算法 (Adaptive Genetic Algorithm, AGA)。Srinivas 等人提出了一种根据适应度值动态调整交叉概率和变异概率的自适应遗传算法 (SAGA)<sup>[9]</sup>。

### 2.2 改进的自适应遗传算法的原理

#### 2.2.1 Logistic 曲线方程

Logistic 方程最早是由比利时数学家 Verhu 于 1838 年推导出来的,经过对其长期的应用表明,Logistic 方程能较好地描述某些有界增长现象。在预测学、信息科学、生物学、农学和经济学等领域都有较为广泛的应用。

Logistic 曲线方程的微分形式为:

$$\frac{dN}{dt} = rN(1 - \frac{N}{K}) \quad (1)$$

其积分形式为:

$$N = \frac{K}{1 + e^{-at}} \quad (2)$$

其中,  $N$  表示生物量 (Biomass)、生长量 (Growth) 或其他数量指标 (如发病数等);  $t$  表示时间 (或温度

等) 序列; $r$  是常数,表示内禀自然增长率或瞬时增长率(Instantaneous Growth Rate); $K$  也是常数,称为环境承载力或容纳量(Carrying Capacity); $e$  为自然对数底; $a$  为积分常数。式(2) 即为  $S$  形 Logistic 累计分布曲线方程<sup>[10]</sup>。

在 Logistic 曲线方程的基础上,令  $K = 1$ 、 $a = 0$ 、 $r = 1$ 、 $N = y$ 、 $t = \pm x$ ,则在公式(2) 的基础上得到以下两个简化的函数方程:

$$y = \frac{1}{1 + e^{-x}} \tag{3}$$

$$y = \frac{1}{1 + e^x} \tag{4}$$

可以看出方程(3) 随着  $x$  取值的增大函数值越来越大并最终趋向于一个固定值。因此通过调整系数,可以将其进一步改进为交叉概率  $P_c$  的调节公式。同样,根据方程(4) 可以进一步将其改进为变异概率  $P_m$  的调节公式。

2.2.2 相似系数

为了能够根据当前种群的分布情况进行分析,文中引入了相似系数的概念,对当前种群的整体优良情况进行有效地分析,以反映当前种群的相似程度。当相似系数较小时,表示种群个体之间的相似程度较低;当相似系数较大时,表示种群之间个体的相似程度较高。

由于遗传算法中随着算法进化代数的不断增加,种群中个体的平均适应度值越来越大,重复个体也越来越多,种群的搜索区域也逐渐减少,使得最优解趋近收敛,从而使得种群中个体适应度值的离散程度逐渐降低。因此,可以用种群中个体适应度值的期望  $EX$  来反映种群适应度值的平均水平,用方差  $DX$  来反映种群适应度值的离散程度<sup>[11]</sup>。

相似系数的计算公式如下式所示:

$$\varphi = \frac{EX + 1}{\sqrt{DX}} \tag{5}$$

$$EX = f_{avg} = \frac{f_1 + f_2 + \cdots + f_M}{M} \tag{6}$$

$$DX = \frac{f_1^2 + f_2^2 + \cdots + f_M^2}{M} - f_{avg}^2 \tag{7}$$

其中, $\varphi$  表示当前种群的相似系数; $M$  表示种群规模; $f_i(i = 1, 2, \cdots, M)$  表示种群中个体适应度值。

从公式(5) 中可以看出随着遗传算法进化代数的增加,种群适应度值的期望  $EX$  的值逐渐增大,方差  $DX$  的值却逐渐减少,从而相似系数  $\varphi$  的值也随着逐渐增大,种群中个体之间越来越相似。

2.2.3 改进的自适应遗传算法的自适应调节公式

根据自适应遗传算法中交叉概率和变异概率的设

计原则,结合相似系数的定义,引入了改进的自适应遗传算法的交叉概率和变异概率的动态调节公式:

$$P_c = 0.6 \times \frac{1}{1 + e^{-\frac{k_1}{\varphi}}} + 0.3 \tag{8}$$

$$P_m = \frac{k_2}{5(1 + e^{\frac{1}{\varphi}})} \tag{9}$$

其中, $k_1$ 、 $k_2$  为两个常数, $k_1 \in (0, +\infty)$ 、 $k_2 \in (0, 1)$ 。

从交叉概率和变异概率的自适应调节公式中可以看出,随着  $\varphi$  的增大,交叉概率  $P_c$  的值越来越小,而变异概率  $P_m$  的值越来越大,且  $P_c$  的变化范围为  $(0.6, 0.9)$ 、 $P_m$  的变化范围为  $(0, 0.1)$ 。因此,改进后的自适应遗传算法的交叉概率和变异概率的调节公式,符合交叉概率和变异概率的设计原则和变化规律,而且其变化范围也控制在合理的范围之内。

2.3 自适应记忆遗传算法(Adaptive Memory Genetic Algorithm, AMGA)的实现

在记忆遗传算法的基础上,根据上述调节公式(8)、(9),对算法进行改进。每一代种群在执行交叉算子和变异算子操作之前,先计算当前种群中个体的相似系数,利用调节公式对交叉概率和变异概率进行动态调整,按照调整后的交叉概率和变异概率的大小执行交叉算子和变异算子。

3 实验测试与结果

TSP(即旅行商问题)是一个典型的组合优化问题,已被证明为 NP 难问题<sup>[12]</sup>。实际应用中的电气布线、网络通讯、管道铺设、货物配送路线、加工调度、乘凉路径规划等许多复杂的实际问题,经过简化处理后都可建模为 TSP。因此,文中应用改进后的遗传算法实现 TSP 的求解,对算法的性能进行验证。

3.1 计算个体适应度值时间消耗方面性能测试

文中测试数据来源于 TSP 标准库 TSPLIB 中的城市坐标数据源文件。采用的测试的实例是 pr76,是 76 个城市的 TSP。

表1 求解不同 TSP 时基本遗传算法和自适应记忆遗传算法相关参数的设置

TSP	城市规模	种群规模	SGA 交叉概率	SGA 变异概率	最大终止代数
pr76. tsp	76	50、100、150	0.6	0.005	30 000

表 1 给出了基本遗传算法和文中改进后的算法求解 TSP 的相关参数的设置,其中自适应记忆遗传算法中基因库的规模记为 GWS,重复实验测试 30 次,对种群中计算个体适应度值消耗的时间求平均值。表 2 给出了以 pr76. tsp 为测试数据的基本遗传算法与自适应



记忆遗传算法计算个体适应度的时间消耗对比。

从表 2 中可以看出,相比于基本遗传算法,自适应记忆遗传算法在计算种群中个体适应度值方面消耗的时间更少。

表 2 解 pr76. tsp 时基本遗传算法和自适应记忆遗传算法计算个体适应度的时间消耗(s)

不同遗传算法	Popsize = 50	Popsize = 100	Popsize = 150
SGA	5.694 59	10.422 68	15.284 91
AMGA (GWS=0.1 * Popsize)	2.286 08	4.654 93	7.790 41
AMGA (GWS=0.2 * Popsize)	2.300 86	5.020 27	8.963 41
AMGA (GWS=0.3 * Popsize)	2.358 66	5.394 06	10.240 41
AMGA (GWS=0.4 * Popsize)	2.394 56	5.696 74	10.532 38
AMGA (GWS=0.5 * Popsize)	2.464 53	6.133 74	11.503 58
AMGA (GWS=1.0 * Popsize)	3.645 79	8.688 44	17.170 02

3.2 算法收敛性测试

以 p76 为 TSP 测试实例,对文中改进后的算法的收敛性能进行测试。测试实例进行 10 次重复测试实验,其收敛代数与最短路径长度的测试结果如表 3,不同算法求得的最优解与 TSPLIB 提供的最优解的平均相对误差如表 4 所示。

表 3 不同算法求解实例 pr76. tsp 的性能对比

不同算法	平均收敛代数	平均最短路径长度	最短路径长度	实际最短路径长度
SGA	2 067.9	189 815.138 8	179 497.362 5	108 159.438 3
Srinivas AGA	1 123.6	136 544.963 1	128 350.335 0	108 159.438 3
AMGA	1 521.7	113 108.868 0	109 679.316 4	108 159.438 3

表 4 不同算法求解的最优解与 TSPLIB 提供的最优解的平均相对误差

不同算法	bayg29. tsp/%	pr76. tsp/%	ch150. tsp/%
SGA	11.75	75.50	82.53
Srinivas AGA	4.66	26.24	21.91
AMGA	1.00	4.58	9.38

上述实验测试数据表明:文中的算法的收敛结果明显好于其他两种算法的收敛结果,最终收敛结果也接近于实际最优解,与 TSPLIB 提供的最优解的平均相对误差最大不超过 9.38%;其次,文中的算法具有较快的收敛速度。

(上接第 62 页)

Proc of WEB KDD 2001: Mining web log data across all customer touch points. [s.l.]: Springer-Verlag, 2002: 113-114.

[6] Dorigo M, Caro G D, Gamardella L M. Ant algorithm for discrete optimization[J]. Artificial life, 1999, 5(3): 137-172.

[7] 鲍 钰. 基于 Web 日志的个性化搜索引擎模型的发现[J]. 计算机应用研究, 2009, 26(5): 1806-1809.

[8] 郭 岩, 白 硕, 杨志峰, 等. 网络日志规模分析和用户兴趣挖掘[J]. 计算机学报, 2005, 28(9): 1483-1496.

[9] 马溪骏, 凌海峰, 刘业政, 等. 基于蚁群算法的群体用户兴

4 结束语

文中提出了一种自适应记忆遗传算法, 分别通过引入基因库和利用 Logistic 曲线方程对调节公式进行调整等策略对遗传算法进行改进, 经实验验证文中算法能够有效减少算法的时间复杂度且收敛速度较快。

参考文献:

[1] Holland J H. Adaptation in natural and artificial systems[M]. Ann Arbor: University of Michigan Press, 1975.

[2] 金 芬. 遗传算法在函数优化中的应用研究[D]. 苏州: 苏州大学, 2008.

[3] Sharapov R R, Lapshin A V. Convergence of genetic algorithms[J]. Pattern recognition and image analysis, 2006, 16(3): 392-397.

[4] 黄樟灿, 李 炜. 有界区域上多峰函数全局优化问题的改进演化算法[J]. 武汉大学学报(理学版), 2007, 53(1): 55-58.

[5] 陈一虎, 刘淳安. 基于实数编码的遗传算法收敛性研究[J]. 西南民族大学学报(自然科学版), 2006, 32(4): 666-669.

[6] 徐宗本, 聂赞坎, 张文修. 遗传算法的几乎必然强收敛性-鞅方法[J]. 计算机学报, 2002, 25(8): 785-793.

[7] Syswerda G, Schaffer J D. Uniform crossover in genetic algorithms[C]//Proceeding of 3rd international conference on genetic algorithms. USA: Morgan Kaufmann, 1989: 2-9.

[8] Kennedy J, Eberhart R C. Particle swarm optimization[C]//Proceeding of IEEE international conference on neural networks. USA: NJ, 1995: 1942-1948.

[9] Srinivas M, Patnaik L M. Adaptive probabilities of crossover and mutation in genetic algorithm[J]. IEEE transaction on systems man and cybernetics, 1994, 24(4): 656-667.

[10] 任子武, 伞 冶. 自适应遗传算法的改进及在系统辨识中的应用研究[J]. 系统仿真学报, 2006, 18(1): 41-43.

[11] 殷祚云. Logistic 曲线拟合方法研究[J]. 数理统计与管理, 2002, 21(1): 41-46.

[12] Applexgate D, Bixby R. Implementing the Danting-Fulkerman-Johnson algorithm for large traveling salesman problems[J]. Mathematical programming, 2003, 97(1-2): 91-98.

趣导航路径发现[J]. 中国管理科学, 2006, 14(3): 56-59.

[10] 凌海峰. 基于 ACO 的 Web 使用挖掘方法研究[D]. 合肥: 合肥工业大学, 2009.

[11] 霍晓钢. 由 Web 日志确定页面访问时间的方法研究[J]. 计算机工程与设计, 2010, 31(10): 2416-2419.

[12] 王有为, 张健斌. 一种新的层次结构网站用户兴趣模式变化识别算法[J]. 系统工程理论与实践, 2008, 28(10): 89-95.

[13] 刘清华. 基于 ACO 的 Web 日志挖掘研究[D]. 南昌: 江西师范大学, 2011.

# 自适应记忆遗传算法研究

作者:

[赵越](#), [徐鑫](#), [赵焱](#), [初雪宁](#), [ZHAO Yue](#), [XU Xin](#), [ZHAO Yan](#), [CHU Xue-ning](#)

作者单位:

[赵越, 徐鑫, 赵焱, ZHAO Yue, XU Xin, ZHAO Yan\(渤海大学 大学计算机教研部, 辽宁 锦州, 121013\), 初雪宁, CHU Xue-ning\(东北大学 信息学院, 辽宁 沈阳, 110819\)](#)

刊名:

[计算机技术与发展](#)

ISTIC

英文刊名:

[Computer Technology and Development](#)

年, 卷(期):

[2014\(2\)](#)

本文链接: [http://d.g.wanfangdata.com.cn/Periodical\\_wjz201402016.aspx](http://d.g.wanfangdata.com.cn/Periodical_wjz201402016.aspx)