第 29 卷第 8 期 计 算 机 应 用 研 究 Vol. 29 No. 8
2012 年 8 月 Application Research of Computers Aug. 2012

格雷码混合遗传算法求解 0-1 背包问题*

王则林[1]，吴志健[2]
(1. 南通大学 计算机科学与技术学院，江苏 南通 226000；2. 武汉大学 软件工程国家重点实验室，武汉 4300072)

摘　要：给出 0-1 背包问题的数学模型，修改传统二进制编码为格雷码混合遗传算法，使用贪心算法来解决约束问题，对每个个体使用价值密度来衡量，提高了算法搜索效率，同时使用精英保留机制来加速算法收敛的速度。最后通过数值实验证明了算法的有效性。

关键词：遗传算法；背包问题；格雷码；贪心算法；精英保留机制
中图分类号：TP18　　**文献标志码**：A　　**文章编号**：1001-3695(2012)08-2906-03
doi:10.3969/j. issn. 1001-3695. 2012. 08. 027

Gray coded hybrid genetic algorithm for 0-1 knapsack problem

WANG Ze-lin[1], WU Zhi-jian[2]
(1. *School of Computer Science & Technology, Nantong University, Nantong Jiangsu* 226000, *China*; 2. *State Key Laboratory of Software Engineering, Wuhan University, Wuhan* 430072, *China*)

Abstract: This paper gave an athematic mode of 0-1 knapsack problem, and modified the binary coding to establish a gray coded hybrid genetic algorithm used greedy algorithm to handle with the constraint conditions, And this paper proposed a value density operator to the individual, which could improve the search effciency, used the elitism mechanism to accelerate the convergence process. The numerical experiment proves the affectivity of the algorithm.

Key words: genetic algorithm(GA); knapsack problem; Gary code; greedy algorithm; elitism mechanism

近些年，背包问题吸引了很多理论和实际工作者对此问题进行深入的研究。主要是由于在工业上很多的实际应用，如资金运算、货仓装载、存储分配等，都是典型的应用例子[1~3]。

0-1 背包问题是一类组合优化问题，迄今已有 40 多年的研究历史，可广泛应用于碎片收集、作业调度、资金预算和货物装箱等领域；0-1 背包是一类 NP 问题，传统方法如持续松弛法、动态规划法、解析法、穷举法、分支界限法和一些近似算法等，一般仅能获得问题的近似最优解。近年来，不少学者将稳健的遗传算法应用于 0-l 背包问题的求解[4~8]，在问题求解质量方面收到了较好的效果。解决背包问题的关键在于两方面：a)约束条件的处理，目前比较流行的是采用惩罚函数的方式处理约束条件，当问题规模较大时，这种方法尽管可行，但搜索效率很低，因此本文在约束条件处理上将贪心法与遗传算法相结合，利用启发式搜索思想来提高效率；b)种群收敛速度和范围的处理，传统的二进制编码遗传算法中，种群收敛速度较慢，而且很容易陷入局部收敛，本文采用格雷码和遗传算法相结合[9,10]，使用精英保留机制[11]，提出了格雷码混合遗传算法，保持了种群在后期的多样性，避免种群过早地陷入局部收敛，提高了问题的求解质量和算法效率。

1　背包问题的数学模型

背包问题是一个 0-1 规划问题，假设有 n 个物品，每个物品的重量和价值分别记为 w_i、p_i($i=1,\cdots,n$)，背包能承受的最大重量是 c，当物品 i 被选择进入背包时，$x_i=1$，否则，$x_i=0$；背包的总重量记为 $\sum_{i=1}^{n} w_i x_i$，总价值记为 $\sum_{i=1}^{n} p_i x_i$。本文的目标旨在得到一个最优解 x_i($i=1,\cdots,n$)，能使得总价值最大，同时又不超过背包的重量限制。背包问题的数学模型可描述如下：

$$\text{maximize} \sum_{i=1}^{n} p_i x_i$$

$$\text{subject to } \sum_{i=1}^{n} w_i x_i \le c$$

$$x_i = 1\ 0 \quad i=1,\cdots,n$$

2　传统算法在求解背包问题中存在的问题

解决背包问题的传统方法有动态规划法、解析法和穷举法。动态规划算法通常用于求解具有某种最优解性质的问题，在这类问题中，可能会有许多可行解，每一个解都对应于一个值，本文希望找到具有最优解(最大值或最小值)的那个解。设计一个动态规划算法，通常可按以下几个步骤进行：

a)找出最优解的性质，并刻画其结构特征。

b)递归地定义最优值。

c)以自底向上的方式计算出最优值。

d)根据计算最优值时得到的信息，构造一个最优解。

步骤 a)～c)是动态规划算法的基本步骤。在只需求出最优值的情况下，步骤 d)可以省略；若需要求出问题的一个最优

收稿日期：2012-01-12；**修回日期**：2012-02-28　**基金项目**：国家自然科学基金资助项目(61070008)；国家教育部人文社科基金资助项目(11YJC870012)；江西省教育厅科技项目(GJJ11106)

作者简介：王则林(1973-)，男，江苏南通人，博士研究生，主要研究方向为进化计算、网络安全等(2011102110045@whu.edu.cn)；吴志健(1963-)，男，江西鄱阳人，教授，博导，博士，主要研究方向为进化计算、智能科学、并行算法等.

解，则必须执行步骤 d）。此时，在步骤 c）中计算最优值时，通常需记录更多的信息，以便在步骤 d）中根据记录的信息快速构造出一个最优解。

动态规划法存在的一个问题是算法需要时间太多，当背包容量 $c>2^n$ 时，动态规划法需要 $\Omega(n2^n)$ 的计算时间。

解析法是通过求解使目标函数梯度为零的一组非线性方程来进行搜索，一般而言，若目标函数连续可微，解的空间方程比较简单，解析法还是可以使用。但是，当方程的变量有几十或者几百个就无能为力了。

穷举法，即在一个连续的有限空间或离散无线搜索空间中，计算空间中每个点的目标函数值，且每次计算一个。显然，这种方法效率太低且鲁棒性不强，许多实际问题所对应的搜索空间很大，不允许一点一点地慢慢求解。

3　格雷码混合遗传算法求解背包问题

3.1　遗传编码

3.1.1　二进制编码

采用二进制 n 维解矢量 X 作为解空间参数的遗传编码，串 T 的长度等于 n。$x_i=1$ 表示该物体装入背包，$x_i=0$ 表示不装入背包。例如 $X=\{0,1,1,0,1,0,1\}$ 表示 2、3、5、7 这四个物体被选入包中。

3.1.2　格雷编码

使用传统的二进制编码时，GA 的局部搜索能力较差，相邻整数的二进制编码可能具有较大的 Hamming 距离。例如，31 和 32 的二进制分别表示为 011111 和 100000，因此算法的解要从 31 改进到 32 必须改变所有的位，这种缺陷将降低遗传算子的搜索效率。二进制编码的这一缺陷称为 Hamming 悬崖。为此，人们提出用格雷码（Gray code）来对个体编码。格雷码是这样一种编码，其连续的两个整数所对应的编码值之间仅仅只有 一个码位是不相同的，其余码位完全相同，如 31 和 32 的格雷码分别表示为 010000 和 110000。二进制编码和格雷码之间的对应关系是：

假设有二进制串 $\langle \alpha_1,\cdots,\alpha_n \rangle$，对应的 Gray 串为 $\langle \beta_1,\cdots,\beta_n \rangle$，则从二进制编码到 Gray 编码的变换为

$$\beta_K=\begin{cases}\alpha_1 & \text{如果 } K=1\\ \alpha_K \oplus \alpha_{K-1} & \text{如果 } K>1\end{cases}$$

从 Gray 编码到二进制编码的变换为

$$\alpha_k=\sum_{i=1}^{k}\beta_i \pmod 2$$

在遗传算法中，群体多样性对于保证搜索的全局性及提高搜索效率是至关重要的，而一直困扰着遗传算法的早熟收敛问题是因群体多样性的逐步丧失而导致的，因此，如果在交叉和变异的遗传操作中，每隔一定的代数变换使用一次格雷码编码，可以对多样性的逐步丧失有一定的减缓作用。

格雷码的主要优点是便于提高遗传算法的局部收索能力，变异和遗传操作也易于实现，符合最小字符集原则，便于用模式定理对算法进行理论分析。

3.2　约束条件的处理

如何处理约束条件是解决背包问题的关键，这里结合启发式搜索算法贪心法来处理背包问题的约束条件。贪心法是指将物品按价值密度 $p_i=v_i/w_i\,(i=1,2,\cdots,n)$ 的值降序排列，然后依据该顺序将相应的物品放入背包中，直到超出背包的容量为止。

由于在编码的过程中会产生很多不满足约束条件的无效染色体，对于这些不满足约束条件的基因编码串，本文利用贪心法的思想，优先装入 v_i/w_i 较大且 $x_i=1$ 的物品，直到背包的容量限制装不下为止。对于基因编码中指示应装入而实际上已装不下的物品，修改其基因编码串中对应的 x_i 为 0，由此可产生一些新的基因编码串，这些新的编码串相对来说是较好的。

3.3　遗传算子的设计

a）选择。采用轮盘赌选择方法，先计算种群中各个个体所对应的适应度总和 $\sum_{i=1}^{n} f_i$，再计算出每一个个体的相对适应度大小 $f_i/\sum_{i=1}^{n} f_i=1$，即为每一个个体遗传到下一代种群中的概率，每一个概率的值组成一个区域，全部的概率和为 1，最后再产生一个 0～1 之间的随机数，依据该随机数出现在上述哪个区域来确定各个个体被选中的次数。另外，还采用了精英保留机制，强行将上一代的最优个体直接保留进入下一代。这样，每进化一代，下一代的最优个体一定不劣于上一代的最优个体。

b）交叉。采用单点交叉，交叉点 k 的范围为 $[1,n-1]$，以随机产生的交叉点为分界相互交换分界点后面的代码串，由此产生两个新的个体。

c）变异。采用单点变异，变异位 k 的范围为 $[1,n-1]$，对系统随机产生的变异位进行翻转，即如果该位原来为 1，则变为 0，原来为 0，则变为 1，由此产生新个体。

3.4　混合遗传算法流程

本文所提出的混合遗传算法的流程如图 1 所示。

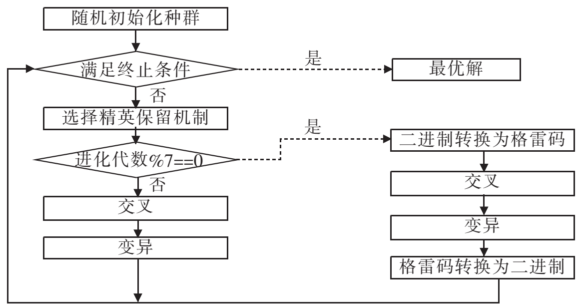

图 1　混合遗传算法流程

4　仿真实验

结合实例，将基本遗传算法、贪心遗传算法与格雷码混合遗传算法进行比较。实例中的问题规模为 80 个变量，背包最大的承受重量 $c=1505$ kg。

表 1　含有 80 个变量的背包问题数据表

p_i	350 340 310 305 300 295 290 287 285 283 280 275 272 270 268 265 260 255 251 245 240 235 230 227 220 215 212 207 203 202 200 198 196 193 190 184 182 181 177 175 160 157 155 154 143 140 137 133 132 128 125 117 110 105 101 92 83 77 75 73 72 70 69 66 60 58 45 40 38 36 33 31 27 23 20 19 10 9 4 1
W_i	135 133 130 11 128 123 20 75 9 66 105 43 18 5 37 90 22 85 9 80 70 31 25 37 41 97 85 57 53 17 20 17 60 35 57 35 61 40 8 50 32 40 72 35 100 2 7 19 28 10 22 27 30 11 5 7 41 22 9 14 19 21 23 88 91 47 68 108 10 12 42 11 20 37 17 4 3 21 10 67
C	1505

表 1 中列出了 80 个变量的背包问题测试数据,实验中遗传算法运行参数的选取为:种群大小 popsize = 200;交叉概率 $p_{cross}=0.6$;变异概率 $p_{mutation}=0.1$;终止代数 maxgen = 500。表 2 ~4 给出了实验结果。

表 2 GA 实验结果

最优解	总重量	总价值	次数
0001001111	1 505	9 120	7
0111101111			
1111100111			
1111111111			
1101011111			
1011110110			
1000000001			
0000000000			

表 3 GA + 贪心法实验结果

最优解	总重量	总价值	次数
0001001111	1 505	9 169	8
0111101111			
1111100111			
1111111110			
1101011111			
1111110111			
1100000000			
0000000000			

表 4 GA + 贪心法 + 格雷码实验结果

最优解	总重量	总价值	次数
0001001111	1 505	9 206	10
0111101011			
1111100111			
1111111111			
1101011111			
1111110111			
1110000010			
0000001000			

观察表 2 ~4 可知,格雷码混合遗传算法得到的最优解性能好于 GA 和贪心 GA,并且格雷码混合遗传算法可以得到性能稳定的最优解。

实验结果表明,GA 法收敛速度慢、效率低;GA 结合贪心算法的方法,其收敛性有较大提高;而本文提出的格雷码混合遗传算法则取得了非常好的效果,加快了收敛速度,很好地解决了早熟收敛问题。

5 结束语

格雷码混合遗传算法大大降低了迂回搜索的空间,提高了搜索效率,减缓了早熟收敛的发生。虽然使用格雷码会使迭代过程中的计算量有一定的增加,但加快了获得最优解的速度。数字仿真实验表明,格雷码混合遗传算法在求解 0-1 背包问题中确实具有良好的收敛性和较高的搜索效率。该算法在求解组合优化问题上是一种较好的方法。

参考文献:

[1] 周本达,陈明华,任哲. 一种求解 0-1 背包问题的新遗传算法[J]. 计算机工程与应用,2009,**45**(30):45-47.

[2] 胡欣,汪红星,康立三. 求解多维 0-1 背包问题的混合遗传算法[J]. 计算机工程与应用,1999,**35**(11):31-33.

[3] 庄中文,钱淑渠. 抗体修正免疫算法对高维 0/1 背包问题的应用[J]. 计算机应用研究,2009,**26**(8):35-37.

[4] 张晓琴,黄玉清. 基于禁忌搜索的启发式求解背包问题算法[J]. 电子科技大学学报,2005,**34**(3):359-362.

[5] 陈国良,王煦法,庄镇泉,等. 遗传算法及其应用[M]. 北京:人民邮电出版社,2001.

[6] 郑宗汉,郑晓明. 算法设计与分析[M]. 北京:清华大学出版社,2005.

[7] 张良杰,毛志杰,李衍达. 遗传算法中突变因子的分析及改进策略[J]. 电子学报,1996,**18**(6):590-595.

[8] 王小平,曹立明. 遗传算法——理论、应用与软件实现[M]. 西安:西安交通大学出版社,2002.

[9] YANG Xiao-hua, YANG Zhi-feng, YIN Xin-an. Chaos gray-coded genetic algorithm and its application for pollution source identifications in convection-diffusion equation[J]. Communications in Nonlinear Science and Numerical Simulation,2008,**13**(3):1676-1688.

[10] CHAKRABORTI N, MISHRA P, BANERJEE A. Optimization of aluminum oxynitride compaction process using a gray-coded genetic algorithm[J]. Materials Letters,2004,**58**(3):136-141.

[11] 曹晖,周延. 多种群精英共享遗传算法在异常光谱识别中的应用[J]. 光谱学与光谱分析,2011,**31**(7):1847-1851.