

DOI:10.3876/j.issn.1000-1980.2008.06.030

一种求解最大独立集的自学习进化算法

丁根宏¹ 李勤丰² 李尤丰³

(1.河海大学理学院,江苏 南京 210098;2.金陵科技学院基础部,江苏 南京 210001;
3.金陵科技学院信息技术学院,江苏 南京 210001)

摘要:为了求解最大独立集问题,通过对求解最大团问题 EA/G 算法的分析,从初始解选取、种群的构成、遗传策略等方面对 EA/G 算法进行了改进,提出了自学习进化算法,并在 DIMACS 基准图上进行了大量的实验.实验结果表明,该算法运算结果比 EA/G 算法所求结果有很好的改善.

关键词:遗传算法;EA/G 算法;最大独立集;最大团;自学习进化算法

中图分类号: O157.6 **文献标识码:** A **文章编号:** 1000-1980(2008)06-0863-04

最大独立集问题是图论中的经典组合优化问题,已被证明是 NP 完备的^[1].它在编码理论、故障诊断、计算机视觉、电台信息传送、生物学、聚类分析、信息检索、移动通信、VLSI 电路设计、排课表等方面有着广泛的应用^[2-8].基数最大的极大独立集称为图 G 的最大独立集^[9](the maximum independent set);基数最大的团称为图 G 的最大团^[9](the maximum clique). I 是图 G 的最大独立集当且仅当 I 是 G 的补图 $\bar{G}$ 的最大团.本文针对 EA/G 算法的不足,采用了一系列改进策略,提出了一种求解最大独立集的自学习进化算法.

1 EA/G 算法

遗传算法是基于种群进化思想的算法.迄今为止解决最大团问题性能最好的遗传算法是 EA/G (evolutionary algorithm with guided mutation)算法^[10].

1.1 EA/G 算法的思想和特征

EA/G 算法是基于种群优胜劣汰思想的算法,它与其他遗传算法比较有着自身的特征:

- 根据近似最优化理论^[11]中的一个假设(一个问题的质量好的解都具有相似结构),EA/G 算法引入了全局统计信息,并将之与当前解的局部信息结合,用于产生新的解.
- 变异操作(guided mutation)应用于当前种群中适应性最好的解上,而经过变异后的解有可能不是可行解,所以应用修复操作(heuristic repair)把变异后的解修复成可行解.
- 整个解空间分成不同的区域,算法在不同的阶段搜索不同的区域.

1.2 EA/G 算法的实现步骤

求解最大团问题的 EA/G 算法的步骤^[9]如下:

- 初始化参数 $N, \Delta, \alpha, \beta, \lambda$.
- 令 $t = 0$,随机选择 $x \in \Omega$,应用修复操作将 x 修改为团 U ,并令 $l = |U|$.
- 从 $\bigcup_{i=l+1}^{l+\Delta} \Omega_i$ 中随机选取 N 个二进制串,并对每个串进行修复操作,修复后的解构成种群 $Z(t)$,初始化概率向量 p .
- 从 $Z(t)$ 中选取最好的 N 个解构成父代 $F(t)$,更新概率向量 p .
- 对 $F(t)$ 中最适应的二进制串连续应用变异操作 $N/2$ 次形成新的解,并将这些解修复成团,然后把这 $N/2$ 个团加入 $F(t)$ 中构成种群 $Z(t+1)$.此时,若满足终止条件,则输出当前最大团.
- 令 $t = t + 1$,并设 S 是 $Z(t)$ 中的最大团,若 $|S| > l$ 则令 $l = |S|$,且转 c.
- 如果 $Z(t)$ 中的所有二进制串的适配值都相同,则转 c,否则转 d.

2 自学习进化算法

2.1 EA/G 算法的不足和改进策略

对 EA/G 算法的实时运行情况进行分析,发现算法中一些步骤运算效率不高,现从以下几个方面对 EA/G 算法进行改进.

2.1.1 初始解的选取

EA/G 算法的初始解是随机的,也即初始解可能只是一个具有很小势的独立集.在该独立集基础上进行后续操作就会造成很多的花费,因此这里改用随机搜索最小度算法来产生初始解.为了增加解的多样性,从度数最小的 3 个顶点中随机选择一个加入到独立集中.算法步骤如下:

- a. 令 $V(G)=V, I=\emptyset$.
- b. 从 $V(G)$ 中任取顶点 v , 使得 $I=I\cup v, G=G-\{v\}\cup N(v)$.
- c. 如果 $V(G)=\emptyset$ 则停止并输出 I , 否则从 $V(G)$ 中取顶点度数最小的 3 个顶点, 在这 3 个顶点中随机选择一个作为 v_k .
- d. $I=I\cup v_k, G=G-\{v_k\}\cup N(v_k)$ 转 c.

2.1.2 适应度相同时种群的选择

在 EA/G 算法步骤 g 中,如果 $Z(t)$ 中的所有二进制串的适配值都相同,则会重新产生初始解,从而造成前面运算所得信息的浪费.因此可采用保优策略,即保留第一个个体,重新产生其余个体.这种策略可提高收敛的速度.

2.1.3 种群的构成

EA/G 算法中的种群个体是从 $\bigcup_{i=1}^{l+\Delta} \Omega_i$ 中随机选取 N 个二进制串,然后对每个串进行修复,这样虽然可以增加算法随机性,但效率较低.因为 Ω_i 中的元素个数是 C_n^i ,在 $\bigcup_{i=1}^{l+\Delta} \Omega_i$ 中就有 ΔC_n^i 个元素,故选取范围过于庞大.因此改用连续使用随机搜索最小度算法共 N 次,相应得到 N 个极大独立集,用这 N 个极大独立集来构成种群 $Z(t)$.

2.1.4 动态调整策略

在修复(heuristic repair)操作中^[9],从 W 中随机选取顶点 k ,然后依据一定概率决定是否将 k 保留在集合中.在具体的操作中,本文改用了这样的策略:当顶点 k 从集合中被删除时,把与当前独立集不相邻的顶点集合中最后一个位置上的顶点放在 k 的位置上,所以当下一搜索到原来 k 所在的位置时,里面的顶点已经不是原来的顶点了,这样就把顶点序列打乱,从而增加解的多样性.以下举例说明实现过程.

设与当前独立集 I 不相邻的顶点集合为 $T=\{3, 7, 8, 11, 23, 27, 35, 86\}$,且设 T 的元素放置情况为

$$\begin{matrix} 3 & 27 & 7 & 11 & 35 & 86 & 23 & 8 \end{matrix}$$

当从 W 中删除顶点 35 时, T 改变为 $T=\{3, 7, 8, 11, 23, 27, 86\}$,元素放置时将原本放在最后的顶点 8 放置在顶点 35 的位置上,具体为

$$\begin{matrix} 3 & 27 & 7 & 11 & 8 & 86 & 23 \end{matrix}$$

所以下一次搜索到位置 5 时,里面的元素已经变为顶点 8 了.重复上述过程多次后,顶点序列与最初的序列大相径庭,从而增加解的多样性.

2.1.5 更换操作

对 $F(t)$ 中的任意 2 个个体,因为直接交叉得到的顶点集合不一定是独立集,所以这里按以下方式来取代遗传算子中交叉操作.对个体 x 来说,当顶点 v_i 在独立集中时, x 的第 i 个位置上的元素为 1. 所以统计 x 中 1 的位置就能知道当前独立集.当用随机搜索最小度算法构成极大独立集时,在与当前独立集都不相邻的点集合 T 中任意选取顶点加到独立集中,并及时更新个体 x . 重复此过程直到 x 为极大独立集.重复整个操作 m 次就可以得到 m 个新的极大独立集,把这 m 个独立集和临时种群 $F(t)$ 一起构成子代种群 $Z(t+1)$.

2.2 自学习进化算法求解最大独立集问题的步骤

根据上述改进策略得到自学习进化算法,具体步骤如下:

- a. 初始化参数种群大小 N 、学习率 λ 、迭代次数 $i_{\max}$.

- b. 令 $t = 0$,使用随机搜索最小度算法产生初始解 U ,并令 $f_0 = |U|$.
- c. 连续使用随机搜索最小度算法共 $N = 2m$ 次 ,产生初始化种群 $Z(t)$,并初始化概率向量 p .
- d. 从 $Z(t)$ 中选取最好的 m 个解用来构成父代 $F(t)$,然后更新概率向量 p .
- e. 利用更换操作 ,重新使用随机搜索最小度算法产生 m 个解 ,加入到 $F(t)$ 中构成种群 $Z(t + 1)$.此时 ,若满足终止条件 ,则输出当前最大独立集 .
- f. 令 $t = t + 1$,并设 S 是 $Z(t)$ 中的最大独立集 ,若 $|S| > f_0$,则令 $f_0 = |S|$ 并且转 c .
- g. 如果 $Z(t)$ 中的所有二进制串的适应度函数值都相同 ,则采用择优操作 ,然后转 c ,否则转 d .

2.3 实验结果

为了测试自学习进化算法的有效性 ,用此算法对 DIMACS 提供的基准图进行了大量实验 ,得到了问题的近似最优解 ,部分结果见表 1 .其余基准图如 c-fat200-1 ,c-fat200-2 ,c-fat200-5 ,c-fat500-1 ,c-fat500-2 ,c-fat500-5 ,c-fat500-10 ,keller4 ,keller5 ,p_hat300-1 ,p_hat300-2 ,p_hat300-3 ,p_hat500-1 ,p_hat500-2 ,p_hat500-3 ,p_hat700-1 ,p_hat700-2 ,p_hat700-3 ,p_hat1000-1 ,p_hat1000-2 ,p_hat1000-3 ,hamming6-2 ,hamming6-4 ,hamming8-2 ,hamming8-4 ,hamming10-2 ,san200_0.7_1 ,san200_0.7_2 ,san200_0.9_1 ,san200_0.9_2 ,san200_0.9_3 ,san400_0.5_1 ,san400_0.7_1 ,san400_0.7_2 ,san400_0.7_3 ,san400_0.9_1 ,sanr200_0.7 ,sanr200_0.9 ,sanr400_0.5 ,sanr400_0.7 ,johnson8-2-4 ,johnson8-4-4 ,johnson16-2-4 ,johnson32-2-4 ,MANN_a9 ,MANN_a27 均获得了目前最好的结果 ,为节省篇幅未在表中列出 .

分析表 1 数据可以看到 ,自学习进化算法相当稳定且有效 ,在参数选取要求不太高的情况下就能得到绝大多数算例的最好结果 ,非常令人满意 .

另外要说明的是 ,用自学习进化算法运算时 ,对图 brock200_4 ,brock400_3 选择的参数分别为 $\lambda = 0.5$ 和 $\lambda = 0.9$,图 keller5 ,brock400_2 ,brock400_4 ,brock800_2 ,brock800_3 ,brock800_4 ,san400_0.7_3 选择的参数 $\lambda = 0.99$,且这些图的迭代次数均大于 1 000 .

3 结 语

通过分析 EA/G 算法存在的不足 ,提出了自学习进化算法 ,给出了自学习进化算法的计算步骤 ,并在 DIMACS 上进行了实验验证 .从大量实验结果来看 ,自学习进化算法运算结果比 EA/G 算法所求结果有很好的改善 ,且运行时间少 ,效率高 ,有一定的应用价值 .

参考文献 :

[1] GAREY M R ,JOHNSON D S .Computer and intractability :a guide to the theory of NP-Completeness[M] .San Francisco :W .H .Freeman ,1979 .

[2] SLONE N J A .Unsolved problem in graph theory arising from the study of codes[J] .Journal Graph Theory Notes of New York ,1989 ,18 :11-20 .

[3] BERMAN P ,PELC A .Distributed fault diagnosis for multiprocessor system[C]//Proceeding of the 20th Annual International Symposium on Fault-Tolerant Computing .Newcastle ,UK [s .n .] ,1990 :340-346 .

[4] BALLARD D H ,BROWN M .Computer visior[M] .Englewood Cliffs ,New Jersey :Prentice-Hall ,1982 .

[5] YOSHIO S .Broadcasting network problem on CATV[J] .Transaction Institute Electronics Committee Engrs ,Japan ,1975 ,58(4) :376-386 .

[6] PARDALOS P M ,XUE Jue .The maximum clique problem[J] .Journal of Global Optimization ,1994 ,1(4) :301-328 .

[7] BASAGNI S .Finding a maximal weighted independent set in wireless networks[J] .Telecommunication Systems ,2001 ,18(1) :155-168 .

表 1 自学习进化算法在 benchmark 基准图上的实验结果

Table 1 Experimental results of the self-learning evolution algorithm on DIMACS benchmark

基准图	边密度	顶点数	EA/G 算法的结果	自学习进化算法的结果	最好结果
brock200_1	0.742	200		21	21
brock200_2	0.496	200	12	12	12
brock200_3	0.603	200		15	15
brock200_4	0.657	200	17	17	17
brock400_1	0.747	400		25	27
brock400_2	0.749	400	25	29	29
brock400_3	0.746	400		31	31
brock400_4	0.748	400	33	33	33
brock800_1	0.649	800		19	23
brock800_2	0.651	800	21	24	24
brock800_3	0.649	800		25	25
brock800_4	0.649	800	21	26	26

- [8] 李勤丰. 最大独立集在高校排课表系统中的应用[J]. 广西科学院学报, 2006, 22(4): 339-341.
- [9] BONDY J A, MURTY U S R. 图论及其应用[M]. 吴望名, 译. 北京: 科学出版社, 1984.
- [10] ZHANG Qing-fu, SUN Jian-yong, TSANG E. An evolutionary algorithm with guided mutation for the maximum clique problem[J]. IEEE Transactions on Evolutionary Computation, 2005, 9(2): 192-200.
- [11] GLOVER F, LAGUNA M. Tabu search[M]. Boston: Kluwer Academic Publishers, 1997.

Self-learning evolution algorithm for the maximum independent set

DING Gen-hong¹, LI Qin-feng², LI You-feng³

(1. College of Sciences, Hohai University, Nanjing 210098, China ;

2. Basic Department, Jinling Institute of Technology, Nanjing 210001, China ;

3. School of Information Science and Technology, Jinling Institute of Technology, Nanjing 210001, China)

Abstract : In order to solve the problems concerning the maximum independent set, the EA/G algorithm was improved by analyzing the algorithm for the solutions to the maximum clique problems from the aspects of selecting the initial solution, constructing populations and genetic strategies. A self-learning evolution algorithm was put forward. After abundant experiments were performed on DIMACS benchmark. The experimental results show that the results obtained by the self-learning evolution algorithm are much better than those by the EA/G algorithm.

Key words : genetic algorithm ; EA/G algorithm ; maximum independent set ; maximum clique ; self-learning evolution algorithm

《水利水电科技进展》征订启事

(邮发代号 28-244, CN32-1439/TV, ISSN1006-7647, 双月刊, A4 开本)

《水利水电科技进展》由河海大学主办, 是全国中文核心期刊, 中国科技核心期刊, 全国水利系统优秀期刊, 华东地区优秀期刊, 江苏省优秀期刊. 主要刊登水科学、水工程、水资源、水环境、水管理方面的科技论文, 主要栏目有水问题论坛、研究探讨、工程技术、水管理、专题综述、国外动态等, 适合与水利、水电、水科学、水工程、水资源、水环境有关的科研、工程、管理人员以及大专院校师生阅读.

本刊由邮局发行, 邮发代号 28-244, 2009 年每期定价 10 元, 全年 6 期共计 60 元. 可在全国各地邮局订阅, 也可直接向编辑部订阅.

编辑部地址: 210098 南京市西康路 1 号 《水利水电科技进展》编辑部.

电话/传真: 025-83786335

E-mail: jz@hhu.edu.cn

网址: http://kbb.hhu.edu.cn/index_jz.htm