改进遗传优化的贝叶斯网络结构学习①

张　亮，章　兢

(湖南大学 电气与信息工程学院，长沙 410082)

摘　要：针对贝叶斯网络结构学习提出了一种改进的遗传算法，和传统遗传算法相比，该改进算法针对贝叶斯网络结构学习问题增加了优化变异和修正非法图两个新的算子。新算子不但保持了贝叶斯网络学习的多样性和正确性，而且还能保证算法快速搜索到全局最优的网络结构。将该改进遗传算法用于贝叶斯网络结构学习的仿真结果表明，和传统 K2 算法、GS/GES 算法、遗传算法和粒子群算法等算法相比，该算法具有更好的全局搜索能力和收敛速度。

关键词：贝叶斯网络；结构学习；全局最优；遗传算法；粒子群算法

Structure Learning of BN Based on Improved Genetic Algorithm

ZHANG Liang, ZHANG Jing

(College of Electrical and Information Engineering, Hunan University, Changsha 410082, China)

Abstract: An improved genetic algorithm (IGA) is proposed in this paper for structure learning of Bayesian Network (BN). Compared with the traditional GA, two new operators named optimized mutation and illegal figure modification are proposed in the improved GA, which aim to solve the BN structure learning problem. The two new operators can simultaneously maintain the diversity and correctness of BN structure learning as well as the algorithm convergence speed of searching the global optimal network structure. In simulation, compared with the traditional algorithms such as K2 algorithm, GS/GES algorithms, normal GA, PSO, etc., the proposed GA shows better performance in global searching and convergence speed.

Key words: Bayesian network; structure learning; global optimization; GA; PSO

在人工智能领域，不确定性知识的推理和决策一直是一个重要的研究问题。而贝叶斯网络正是对不确定性问题模拟和推理的一种有效工具[1,2]。贝叶斯网络学习主要包括参数学习和结构学习两个部分的内容[3]，在实际应用中其网络结构的规模会随着变量的数目和每个变量的状态数量呈指数级增长，从而导致网络结构学习计算量也会随之而呈指数级增长。为了克服构建网络结构中的计算和搜索的困难，许多学者进行了大量的探索工作。1992 年 Cooper 等首先提出的最经典的基于爬山搜索算法的 K2 算法[4]。2002 年 Chickering 提出的贪婪算法(greedy search，GS)[5]该算法易陷入局部优值。2007 年 Sahin 等提出分布式离散粒子群算法[6]，将其用于贝叶斯网络学习。文献[6]的仿真结果证明其具有收敛速度快的特点，但其和贪婪算法一样容易陷于局部最优。针对已有文献[4-6]的利弊，本文提出一个改进的遗传算法来专门处理贝叶斯网络结构的学习问题。和传统遗传算法相比，该算法在始终保留最优个体的情况下，通过多点交叉、优化变异、修正非法图等操作来代替经典遗传算法中的交叉，选择，变异等算子操作。和传统方案[4]和[5]相比，本算法可以缩小网络结构的搜索空间，提高结构的学习效率，从而避免收敛到次优网络模型。本文最后通过典型的 Asia 网络，验证了该算法得有效性，并与经典算法[4,5]、PSO 和传统的 GA 等智

① 基金项目:国家自然科学基金(60634020);长沙市科技计划(K1005018-11)
收稿时间:2011-01-06;收到修改稿时间:2011-03-02

能优化算法相比取得了较好的实验效果。

1 贝叶斯网络

1.1 贝叶斯网络概念

贝叶斯网络是基于概率分析、图论的一种不确定性知识表达和推理模型[2]。贝叶斯网络又称信度网络，是 Bayes 方法的扩展，也是目前不确定知识表达和推理领域最有效的理论模型之一。

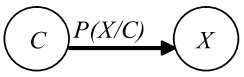

图 1 贝叶斯网络

如图 1 所示，一个贝叶斯网络可以看作一个有向无环图，由变量节点及连接这些节点的有向边构成。节点(C,X)代表随机变量，节点间的有向边代表了节点间的互相关系(由父节点指向其后代节点)，用条件概率(P(X/C))进行表达关系强度，没有父节点的则用先验概率进行信息表达。节点变量可以是任何问题的抽象，如：测试值，观测现象，意见征询等。贝叶斯网络不仅适用于表达和分析不确定性和概率性的事件，还可以应用于有条件地依赖多种控制因素的决策。此外，其还可以从不完全，不精确或不确定的知识或信息中做出推理。

1.2 贝叶斯网络的学习

如前所叙，贝叶斯网络的学习包含参数学习和结构学习。参数学习分为处理完整数据和处理缺失数据两类。结构学习分为对已知结构和对未知结构的学习，因此对贝叶斯网络学习有四类情况[7]：1)已知结构，数据完整；2)已知结构，数据缺失；3)未知结构，数据完整；4)未知结构，数据缺失。通常在网络结构已知情况下，常见的参数学习方法有最大似然估计算法、贝叶斯估计算法、不完备数据下参数学习等。即用 MLE 公式和 BE 公式、EM 来求参数，故第 1，2 种情况参数学习已经相当成熟。结构学习相对而言就要困难得多，也是贝叶斯网络研究的热点和难点，第 4 种情况常用 SEM 算法和 MCMC 算法来实现结构学习，其复杂程度高，计算量都很大。本文主要侧重于在数据已知的情况下的结构学习。

在数据完备情况下，结构学习算法可以分为基于搜索和评分的方法(Search and Score based Method)和基于独立性测试的方法(Conditional Independence Testing based Method)。本文主要采用的是基于搜索和评分的方法。它是选定一个适当的记分函数（本文选用 BIC 测度作为计分函数），通过不断地改变网络结构计算出相应的记分值，结合一些优化算法在结构空间进行启发搜索，直至搜索到的结构具有最高记分为止。

1.3 贝叶斯信息标准测度(BIC)

1978 年 Schwarz[8]提出了贝叶斯信息标准测度(BIC，Bayesian information criterion)来评价贝叶斯网络结构。其公式为：

$$Q_{BIC} = LL(B \mid D) - 1/2 \log N * Dim(B) \quad (1)$$

其中，B 表示所学得的贝叶斯网络结构，D 是训练数据集，

$$LL(B \mid D) = \sum_{i=1}^{n} \log p(B \mid D) \quad (2)$$

是基于概率分布描述 D 所需要的比特数的度量，1/2 表示每个参数使用的比特数，贝叶斯网络的维度：

$$Dim(B) = \sum_{i=1}^{n}(r_i - 1)q_i = \sum_{i=0}^{n}(r_i - 1)\prod_{x \subseteq pa_i} r_i \quad (3)$$

是指明随机变量 X 的联合概率分布所需要的自由参数的数目。

2 本文算法

遗传算法模拟生物进化的优胜劣汰规则与染色体的交换机制，通过选择、交叉、变异三种基本操作寻求最优个体，是处理复杂优化问题的一类方法，具有极高的鲁棒性和广泛适用性[9]。但是,遗传算法亦表现出迭代次数多、收敛速度慢、易陷入局部极值的现象。本文将改进的遗传算法用于贝叶斯网络结构的学习，对贝叶斯结构进行矩形编码，多点交叉，优化变异，修正非法图实现全局搜索。

2.1 矩形编码

本文需要对贝叶斯网络结构进行结构矩阵编码，采用 Poza[10]等介绍的编码方式(如图 2 所示)，具体方式如下：

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (4)$$

$$a(i,j) = \begin{cases} 1 & i\text{为}j\text{的父节点} \\ 0 & \text{否则} \end{cases}$$

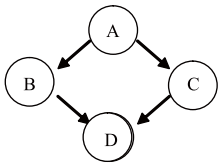

图 2　贝叶斯结构图和编码

2.2 多点交叉

交叉又叫重组，是指把两个父代个体部分结构加以交替，重组而生成新个体。由于矩形结构编码本质是采取二进制编码，所以本文采用文献[9]所提多点交叉[11]，在个体基因中随机设置多个交叉点，能有效避免早熟，进入局部收敛。本文交叉实现方法是把矩形结构编码的结构矩阵变成行向量，再进行多点交叉。

2.3 优化变异

变异是指生物体子代与亲代之间的差异，子代个体之间的差异的现象。本文针对贝叶斯网络的结构提出了一种新的优化变异算子，其包括三种变异方式：1)增加边：一条边增加到网络中，以增加新的依赖。2)删除边：删除一条已有的依赖关系。3)反转边：为了改变现有网络中某两个节点的相互依赖特性而反转其连接边方向。

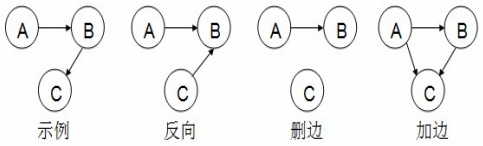

图 3　变异后的拓扑图

如图 3 所示，显示了三种基本变异的拓扑图，而优化变异的思想如下：在决定增加边或删除边时，应考虑这些边所邻接点间的贝叶斯信息标准测度，1)当决定增加有向边时，就选择所有可添加边中 BIC 测度最高的边，2)当决定删除有向边时，就选择所有可删除边中 BIC 测度最低的边，3)边反向操作可以随机操作。这样才能保证无论是哪一种变异，都能保证变异后评分函数值最大，达到全局最优。

2.4 修正非法图

如图 4 所示，在经过多点交叉和优化变异后，必然会出现一系列的非法拓扑图和非法结构如图，从而会严重影响到下一代的遗传进化，所以修正的操作必不可少。

修正步骤如下：

1) 根据网络结构图所对应的矩阵求其传递闭包。

2) 查看闭包对角线上的元素是否全为 0。若不是，保留主对角线上不为 0 的元素所对应的节点(这些节点全位于环内)的 i(i 表示网络结构矩阵的第 i 行)。此时的矩阵为传递闭包前的矩阵，求其父节点，删除或反向这些父节点指向该节点的任一边。

3) 执行去环操作，使拓扑图中不存在有向环。

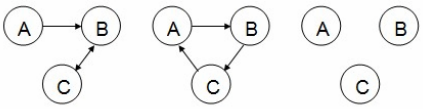

图 4　三种常见非法图

3　实现步骤

3.1 贝叶斯网络工具箱 BNT

基于 MATLAB 的贝叶斯网络工具箱 BNT 是 Kevin P. Murphy 基于 MATLAB 语言开发的关于贝叶斯网络学习的软件包[12]，BNT 是个完全免费的软件包,其代码完全公开,系统的可扩展性良好。本文算法实现都是用 MATLAB 实现编程。

3.2 实现步骤

步骤如下：

1) 初始化种群数、交叉变异概率等相关参数。

2) 根据贝叶斯网络结构进行矩阵编码，并生成该图的所有邻近图，选取一定数量的结构图为初始种群。

3) 计算所有初始网络结构的计分值作为个体极值，并找出全局最大计分值。

开始循环

4) 将更新个体和局部极值个体的编码进行变换，多点交叉更新个体编码。

5) 将更新个体编码还原成矩阵进行优化变异操作

6) 修正非法图。

7) 计算全部更新后的个体对应的记分值，并将新产生的个体记分值与原来对应的极值进行比较。若大于或小于原来的极值且在允许的误差范围内，则更新个体及其相应的记分值；否则，维持不变。

8) 找出新的全局最大记分值和新的全局最优个体。

9) 若新的全局最大记分值大于旧的全局最大记分值。则保留该记分值及相应的个体，无效迭代次数

清零；否则，全局最大记分值维持不变，无效迭代次数加 1。

10) 判断重复迭代次数。若大于最大重复迭代次数，则跳出循环。

11) 输出全局最优个体，得到最终的拓扑结构图。

为提高算法运行的速度，本文在所用程序中设置如果寻找到的最佳个体的计分值连续 q 次都未改变，或者迭代次数超过允许的最大迭代次数，则算法结束。由于遗传算法受初始种群选择的影响大,在产生初始种群时，采用 Chow 等提出的 MWST 算法来产生贝叶斯网络最初的边集，形成与贝叶斯网络拟合最优(即后验概率最大)的树结构（由函数 learn_ struct_mwst 实现）。由该结构作为初始结构图模型，通过对该图任意加、删或反向一边产生该图的所有近邻矩阵（由函数 mk_nbrs_of_ dag_ topo 实现），在其中选取一定数量的矩阵作为初始种群。

4 算法性能测试

对基于改进遗传算法的贝叶斯网络结构学习，本文以经典的 Asia 网络为例，进行 MATLAB 编程，并通过 Asia 数据库学习得到相应的贝叶斯网络结构。为了验证本文算法用于贝叶斯网络结构学习的有效性，将本文改进算法与经典 K2 算法和 GS2(Greedy Search)算法,GES(Greedy Equivalence Search)算法[4,5]进行对比。初始值设置：粒子种群数为 20，交叉变异概率为 0.8，连续无效迭代次数为 10 次，允许的最大迭代次数为 100 次。

如图 5 所示，从左往右依次为理想的 Asia 结构图，MWST 初始结构图，和本文算法优化后的结构图。由图 5 和图 6 所示，很明显可以看出经 K2，GS2 和 GES 算法优化后的贝叶斯网络都出现了多边的情况，并且 GS 算法的有向边出现了反向边的情况，明显不是最优结构。所以本文算法优化后的结构图与理想的结构图不但相似度最高，而且打分值（BIC）也最接近理想 Asia 网络。这说明了本文算法用于贝叶斯网络结构学习的有效性。

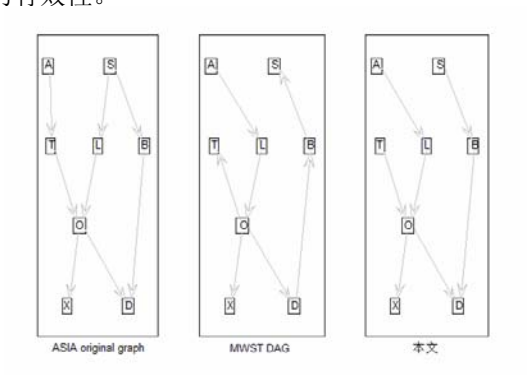

图 5 优化后的结构

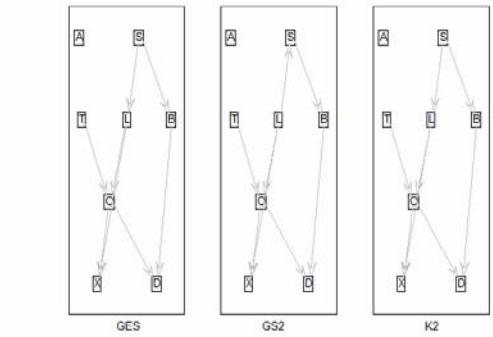

图 6 与三种算法的比较

为了验证本文算法的学习性能，在相同初始条件下，我们将本文改进算法与现有智能优化算法：PSO，GA，HPGA[7](粒子群算法与遗传算法的结合)，进行比较，如图 7 和表 1 所示为多种智能优化算法用于贝叶斯网络结构优化的比较。

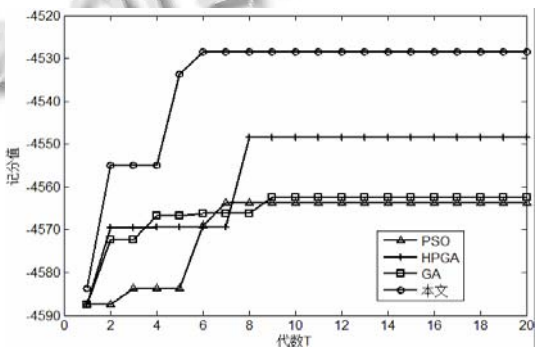

图 7 多种智能算法比较

表 1 多种算法比较

种群规模	GA		PSO		HPGA		本文	
	记分值	时间/代数	记分值	时间/代数	记分值	时间/代数	记分值	时间/代数
20	-4562.4	9.19/9 代	-4570.8	6.71/9 代	-4548.2	9.45/8 代	-4530 .4	9.57/6 代
15	-4568.6	7.38	-4575.7	6.43	-4555.8	7.62	-4532.7	8.65
10	-4573.9	6.93	-4581.1	5.67	-4562.2	7.01	-4534.4	7.92

如图 7 和表 1 所示，随着种群规模的增大，本文算法能够保持较高的计分值，且迭代次数也比较少。而 GA，PSO 和 HPGA 算法受种群规模影响比较大，而且易于过早收敛。由此证明本文算法搜索能力的优越性。而由计分值和迭代次数充分说明本文算法既能保持良好的搜索精度，又能快速收敛到最优结构。

5 结语

本文提出一种适用于贝叶斯网络学习的改进的遗传算法的。相对于传统遗传算法，该改进算法提出了两个适用于贝叶斯网络结构学习的操作算子：优化变异和修正非法图。仿真证明使用该改进的遗传算法对数据库进行学习，可得到全局最优的贝叶斯网络结构。相对于已有文献传统算法[4-6]和智能优化算法(PSO,GA,HPGA)，该算法具有更好的学习能力和收敛速度，从而具有更好的搜索能力。但是由于现有评分函数的局限性会导致在不同网络结构上出现相同的记分值，从而导致优化后的结构离最优结构有一定距离，如何设计一个更有效的评分函数将是未来研究的努力方向。

参考文献

1 Bultan T, Fu X, Hull R, Su J. Conversation specification: A new approach to design and analysis of e-service composition. Proc. of 12th Int'l World Wide Web Conf., May 2003.

2 王双成,冷翠平,李小琳.小数据集的贝叶斯网络学习.自动化学报,2009,35(8):1064-1070.

3 孙岩,唐一源.新的贝叶斯网络结构学习方法.计算机工程与设计,2008,29(5):1238–1240.

4 Cooper GF, Herskovis E. A Bayesian method for the induction of probabilistic networks from data. Machine Learning, 1992, 9(4):309–347.

5 Chickering DM. Optimal structure identification with greedy search. Journal of Machine Learning Research, 2002,11(3): 507–554.

6 Sahin F, Yavuz M, Amavu Z. et al. Fault diagnosis for airplane engines using Bayesian network and distributed particle swam optimization. Parallel Computing, 2007, 11(33):124-–143.

7 许丽佳,黄建国,王厚军,龙兵.混合优化的贝叶斯网络结构学习.计算机辅助设计与图形学报,2009,21(5):633–639.

8 Friedman N, Goldszmidt M. Building classifiers using Bayesian network. Proc. Nation Conference on Artificial Intelligence. Menlo park, CA: AAAI Pres, 1996:1227–1284.

9 尹作海,邱洪泽,周万里.基于改进变异算子的遗传算法求解柔性作业车间调度.计算机系统应用,2009,18(10):156–159.

10 Larranaga P, Poza M, Yurramendl Y, et al. Structure learning of Bayesian networks by genetic algorithms: a performance analysis of control parameters. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1996,18(9):912–926.

11 邓宏贵,罗安,曹建,丁家峰,王会海.基因多点交叉遗传算法在变压器故障诊断中的应用.电网技术,2004,28(24):467–470.

12 Kevin P, Murphy. The Bayes Net Toolbox for Matlab.[2001]. http://www. cs. ubc. ca/~murphyk/

(上接第 85 页)

用资源减少，可为在 FPGA 上实现其他功能预留资源。并且在 FPGA 上实现中值滤波算法的硬件结构简单，集成度高，可靠性强，时序固定，延时小等优点。由于 FPGA 可编程的特点，增加了系统的灵活性，稍作修改就可以适合于不同的系统，有很强的通用性，所以，本文提出的滤波器设计方法具有很强的实用性。

参考文献

1 Gavin L,Saeid N.FPGA implementation of a median filter. TENCON' 97 IEEE Region 10 Annual Conference. Australia, 1997:437–440.

2 陈加成,徐熙平,吴琼.基于 FPGA 的中值滤波算法研究与硬件设计.长春:长春理工大学,2008.

3 潘松,黄继业.EDA 技术实用教程.北京:科学出版社,2005.

4 刘皖,何道君,谭明.FPGA 设计与应用.北京:清华大学出版社,2006.

5 付强.基于 FPGA 的图像处理算法的研究与硬件设计.南昌:南昌大学,2006.