

文章编号: 1001-0920(2013)04-0618-05

基于无约束优化和遗传算法的贝叶斯网络结构学习方法

汪春峰^{1,2}, 张永红¹

(1. 河南师范大学 数学与信息科学学院, 河南 新乡 453007; 2. 西安电子科技大学 理学院, 西安 710071)

摘要: 基于无约束优化和遗传算法, 提出一种学习贝叶斯网络结构的限制型遗传算法. 首先构造一无约束优化问题, 其最优解对应一个无向图. 在无向图的基础上, 产生遗传算法的初始种群, 并使用遗传算法中的选择、交叉和变异算子学习得到最优贝叶斯网络结构. 由于产生初始种群的空间是由一些最优贝叶斯网络结构的候选边构成, 初始种群具有很好的性质. 与直接使用遗传算法学习贝叶斯网络结构的效率相比, 该方法的学习效率相对较高.

关键词: 贝叶斯网络; 结构学习; 无约束优化; 遗传算法

中图分类号: TP301

文献标志码: A

Bayesian network structure learning based on unconstrained optimization and genetic algorithm

WANG Chun-feng^{1,2}, ZHANG Yong-hong¹

(1. College of Mathematics and Information Science, He'nan Normal University, Xinxiang 453007, China; 2. Department of Mathematical Sciences, Xidian University, Xi'an 710071, China. Correspondent: WANG Chun-feng, E-mail: wangchunfeng09@126.com)

Abstract: Based on unconstrained optimization and genetic algorithm, this paper presents a constrained genetic algorithm(CGA) for learning Bayesian network structure. Firstly, an undirected graph is obtained by solving an unconstrained optimization problem. Then based on the undirected graph, the initial population is generated, and selection, crossover and mutation operators are used to learn Bayesian network structure. Since the space of generating the initial population is constituted by some candidate edges of the optimal Bayesian network, the initial population has good property. Compared with the methods which use genetic algorithm(GA) to learn Bayesian network structure directly, the proposed method is more efficiency.

Key words: Bayesian network; structure learning; unconstrained optimization; genetic algorithm

0 引言

贝叶斯网络是一种将贝叶斯概率方法和有向无环图的网络拓扑结构有机结合的表示模型, 描述了数据库中数据项及其相互之间的依赖关系. 构建贝叶斯网络结构时, 若单纯依靠人类专家给出贝叶斯网, 当节点个数较多时, 构造过程将费时费力, 且在很多情况下是不可能的. 近十几年来, 人们对于如何通过现有数据学习构造贝叶斯网或者自动求精由人类专家提供的贝叶斯网进行了深入研究^[1-2].

从数据集中学习贝叶斯网络包括网络结构学习和网络参数学习, 其中结构学习是核心内容. 本文

主要关注于网络结构学习. 贝叶斯网络结构学习的关键是如何确定节点间的弧及弧的方向. 目前网络结构学习所采用的策略可归为3类: 基于依赖独立测试的方法^[3-5], 基于打分搜索的方法^[6-8]和混合算法^[9-10]. 基于依赖独立测试方法的关键是选用合适的测度方法度量变量之间的依赖关系, 然后将依赖关系最强的一些边连接起来, 完成网络结构的学习. 如Chow等^[3]提出的用树结构表示概率网的近似算法; Sprites等^[4]提出的从完全图删除边的PC算法. 这类算法的复杂度相对较低, 但对训练数据集的依赖较大, 如果训练数据集数据量不是充分大或者存在误差, 则所得结果通常是不可靠的. 基于打分搜索方法通常作为最优化解

收稿日期: 2011-11-02; 修回日期: 2012-02-10.

基金项目: 国家自然科学基金项目(60974082, 11171094); 河南省基础与前沿技术研究计划项目(102300410264); 河南省教育厅基础研究项目(2010A110010); 河南师范大学博士科研启动课题(qd12103).

作者简介: 汪春峰(1978—), 男, 讲师, 博士, 从事贝叶斯网络结构学习、最优化理论方法及应用的研究; 张永红(1977—), 女, 硕士生, 从事贝叶斯网络结构学习的研究.

题来处理,即被看作是一个在贝叶斯网络空间内由打分函数引导搜索最优网络结构的过程,所采用的打分函数一般被设计来评价网络结构与数据集合的匹配程度.比较常用的打分函数有BIC、MDL和BDe等,如Suzuki^[7]提出的运用MDL打分函数学习贝叶斯网络结构的算法.这类算法的效率通常较低,而且易于陷入局部最优解,当变量较多时难以实现.为了降低搜索空间,基于打分搜索的方法通常要求节点有序并使用一些启发式搜索方法.混合算法是以上两类算法思想的结合,因其能克服前两种算法的缺陷,目前已成为研究热点之一.

遗传算法是一类处理复杂优化问题的方法.它通过模拟生物进化的优胜劣汰规则与染色体的交换机制,使用选择、交叉和变异3种基本操作寻求最优个体,具有极高的鲁棒性和广泛的适应性.近年来,有学者将此类方法应用于贝叶斯网络结构的学习^[11-12].但是,现有的大多数学习贝叶斯网络结构的遗传算法需要事先给定节点顺序,以减少算法的搜索空间,从而大大降低了算法的通用性.

基于无约束优化和遗传算法,本文提出一种学习贝叶斯网络结构的限制型遗传算法.该方法由两个阶段构成:在第1阶段,构造无约束优化问题并求解,确定一无向图,为第2阶段产生初始种群提供一个较小的空间;在第2阶段,首先产生初始种群,然后使用遗传算法对初始种群进行选择、交叉和变异操作,完成贝叶斯网络结构的学习.由于在第2阶段中遗传算法初始种群的产生是以无向图中的边作为候选边,减小了搜索空间的规模,该方法的效率要比直接使用遗传算法学习的效率高.数值实验表明无论是学习效果,还是计算效率,本文方法都取得了较好的结果.

1 贝叶斯网络

贝叶斯网络是根据马尔可夫假设寻找的满足条件独立性限制的模型结构,又称信度网(BN).贝叶斯网可以表示为一个二元组 $BN = (G, \theta)$.其中: $G = (V, E)$ 是一个非循环的有向无环图(DAG),包括节点变量集合 $V = \{v_1, v_2, \dots, v_n\}$ 和连接节点的有向边集 E ,它捕捉结构特性,降低了推理、决策和学习的复杂度; θ 是一个条件概率分布集, $\theta_i \in \theta$ 表示贝叶斯网络在给定节点 v_i 的父节点时的条件概率.两部分组合在一起定义了唯一的联合概率分布.贝叶斯网是自然、紧凑的联合概率分布的表示形式.

对于具有 n 个属性节点 $\{v_1, v_2, \dots, v_n\}$ 的贝叶斯网络,由联合概率公式可得

$$P(v_1, v_2, \dots, v_n) = \prod_{i=1}^n P(v_i | v_1, v_2, \dots, v_{i-1}). \quad (1)$$

假定对于每个节点 v_i ,在给定其父节点集 $Pa(v_i) \subseteq \{v_1, v_2, \dots, v_{i-1}\}$ 时,满足 v_i 与 $\{v_1, v_2, \dots, v_{i-1}\} \setminus Pa(v_i)$ 条件独立,即有

$$P(v_i | v_1, v_2, \dots, v_{i-1}) = P(v_i | Pa(v_i)). \quad (2)$$

则由式(1)和(2)可知,联合概率 $P(v_1, v_2, \dots, v_n)$ 满足

$$P(v_1, v_2, \dots, v_n) = \prod_{i=1}^n P(v_i | Pa(v_i)). \quad (3)$$

式(3)表明贝叶斯网络实质上是一个联合概率分布 $P(v_1, v_2, \dots, v_n)$ 所有条件独立的图形化表示.

给定离散变量集 $\{v_1, v_2, \dots, v_n\}$ 上的数据集 D ,贝叶斯网络结构学习是指尽可能地结合先验知识,找到与样本数据拟合最好的网络拓扑结构.通常的做法是定义一个评分函数,评判某个具体结构反映的独立关系和样本的匹配程度,选择适宜的搜索算法搜索分值最高的网络模型,即

$$G^* = \operatorname{argmax}_{G \in G_n} f(G : D). \quad (4)$$

其中: $f(G : D)$ 是度量网络 G 与数据集 D 拟合程度的打分函数, G_n 是定义在节点集 V 上的所有DAG.

目前存在的打分函数都是在统计数据集合的基础上设计的,大体上可以分为依据信息论原理设计的打分函数和依据贝叶斯方法设计的打分函数.这些打分函数通常具有一个很重要的性质:可分解性,即

$$f(G : D) = \sum_{i=1}^n f(V_i, Pa(v_i) : N_{v_i, Pa(v_i)}), \quad (5)$$

其中 $N_{v_i, Pa(v_i)}$ 是 v_i 和 $Pa(v_i)$ 在给定数据集中的统计量.

利用打分函数的可分解性,当网络的某节点处发生了边的添加、删除和反向操作,需要计算新网络的得分时,只需计算该节点处的局部分值即可,因为其他节点处未发生改变.这样可以大大减少计算量,在设计算法方面具有重要作用.

n 个节点构成的所有有向无环图都可能作为贝叶斯网络的结构,且 n 个变量可能构成的贝叶斯网络数目的计算公式如下:

$$g(n) = \sum_{i=1}^n (-1)^{i+1} C_i^n 2^{i(n-i)} g(n-i). \quad (6)$$

由式(6)可知,当 $n = 10$ 时,需要搜索的模型个数将达到约 4.17×10^{18} ,可见搜索空间巨大.若直接使用遗传算法(GA)进行贝叶斯网络结构学习,则效率必然受到严重影响.鉴于此,基于一个无约束优化问题,本文先对遗传算法的初始种群进行优化,然后再使用遗传算法学习贝叶斯网络结构.由于初始种群性能较好,本文提出的限制性遗传算法(CGGA)可以较快地学习到最优贝叶斯网络结构.

2 贝叶斯网络结构学习的限制型遗传算法 (CGA)

在给出贝叶斯网络结构学习的限制型遗传算法 (CGA) 之前, 先介绍用于产生较好初始种群的无约束优化问题的构造.

2.1 无约束优化问题

令 $\chi_{ij|k}^2$ 表示自由度为 $df = (n_i - 1)(n_j - 1)n_k$ 的 χ^2 分布. 其中: n_i 是 v_i 的可取值个数; $\chi_{df,\alpha}^2$ 表示自由度为 $df = (n_i - 1)(n_j - 1)n_k$, 显著性水平为 α 的 χ^2 分布值.

定义 1 (1-步依赖系数) 考虑变量 v_i 和 v_j , 称 $c_{ij\alpha}^{(1)} = \min_{k \neq i,j} \{\chi_{ij|k}^2 - \chi_{df,\alpha}^2\}$ 为 v_i 与 v_j 之间的 1-步依赖系数.

注 1 显然, 如果 $c_{ij\alpha}^{(1)} > 0$, 则无论是否有其他变量的加入, 变量 v_i 和 v_j 之间都满足统计学意义上以 $(1-\alpha)\%$ 依赖的, 即它们之间极大可能存在一条边. 如果 $c_{ij\alpha}^{(1)} \leq 0$, 则它们之间可能是条件独立的. 在实际编程实现时, 为了计算 $c_{ij\alpha}^{(1)}$, 根据文献 [13] 中定理 1, $c_{ij\alpha}^{(1)}$ 可用如下的近似量替代:

$$c_{ij\alpha}^{(1)} = \min_{k \neq i,j} \{2NMI_D(v_i, v_j | v_k) - \chi_{df,\alpha}^2\}.$$

其中: N 为样本数据条数, $MI_D(v_i, v_j | v_k)$ 为节点变量 v_i , v_j 与条件节点变量 v_k 之间的条件互信息.

例 1 考虑如图 1 所示的 4 个节点的贝叶斯网络.

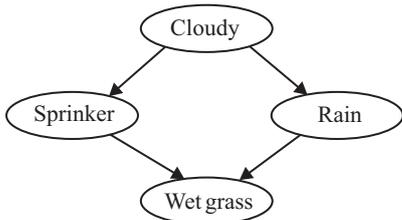


图 1 四节点的贝叶斯网络

将图 1 中的 Cloudy、Sprinker、Rain 和 Wet Grass 分别简记为 C 、 S 、 R 和 W , 并假定每个节点变量可以取两个值. 为了判断节点 S 与 W 之间是否相互依赖, 即节点 S 与 W 之间是否存在边, 可以计算两变量间的 1-步依赖系数 $c_{SW\alpha}^{(1)}$. 假定 χ^2 分布的显著性水平 $\alpha = 0.01$, 此时 χ^2 分布的自由度为 $df = (2 - 1)(2 - 1)2 = 2$, 对应的 χ^2 分布值为 9.2103, 从而根据定义 1 可计算 $c_{SW\alpha}^{(1)} = \min\{\chi_{SW|C}^2, \chi_{SW|R}^2\} - 9.2103$. 如果 $c_{SW\alpha}^{(1)} > 0$, 则节点 S 与 W 之间以 99% 的概率存在一条边.

定义 2 1-步依赖系数矩阵 $C = (c_{ij})$ 定义如下:

$$c_{ij} = \begin{cases} c_{ij\alpha}^{(1)}, & i \neq j; \\ 0, & i = j. \end{cases}$$

注 2 注意到 $\chi_{ij|k}^2 = \chi_{ji|k}^2$, 因此有 $c_{ij\alpha}^{(1)} = c_{ji\alpha}^{(1)}$, 即 C 是对称矩阵.

定义 3 如果存在节点 v_k , 使得 $\chi_{ij|k}^2 < \chi_{df,\alpha}^2$ 成立, 则称节点 v_i 与 v_j 以显著性水平 α 局部条件独立 (LCI).

根据 $c_{ij\alpha}^{(1)}$ 以及局部条件独立定义, 有以下结论:

引理 1 节点 v_i 与 v_j 以显著性水平 α 局部条件独立 LCI 当且仅当 $c_{ij\alpha}^{(1)} < 0$.

定义 4 若对于 $\forall k \neq i, j$, 总有 $\chi_{ij|k}^2 < \chi_{df,\alpha}^2$ 成立, 则称节点 v_i 与 v_j 以显著性水平 α 全局条件独立 (GCI).

一个贝叶斯网络结构可以表示成一个邻接矩阵 $X = (x_{ij})$ 的形式, 其中

$$x_{ij} = \begin{cases} 1, & (v_i, v_j) \in E; \\ 0, & (v_i, v_j) \notin E. \end{cases}$$

通过使用邻接矩阵 $X = (x_{ij})$ 和 1-步依赖系数, 可以构造一个度量网络全局依赖的量.

定义 5 对于一个邻接矩阵为 $X = (x_{ij})$ 的贝叶斯网络, 其 1-步全局依赖度量为

$$H(X, \alpha) = \sum_{i=1}^n \sum_{j=1, j \neq i}^n c_{ij} x_{ij}.$$

在定义 5 的基础上, 可构造如下无约束优化问题:

$$\max H(x, \alpha) = \sum_{i=1}^n \sum_{j=1, j \neq i}^n c_{ij} x_{ij}. \quad (7)$$

因为系数矩阵是一对称阵, 所以由上述无约束优化问题确定的邻接矩阵对应的图必是一无向图. 再根据 $c_{ij\alpha}^{(1)}$ 的定义可知, 无向图中所保留的边是一些可能为最优贝叶斯网络结构的候选边.

2.2 限制型遗传算法 (CGA)

在得到无约束优化问题所对应无向图的基础上, 给出本文所提出的限制型遗传算法.

遗传算法是一类模拟自然过程, 特别是模拟生物界自然进化和遗传过程的随机搜索算法. 它是一个以适应度为依据, 通过对群体中的个体进行选择、交叉、变异等遗传操作来实现群体内个体结构重组的迭代处理过程. 由于贝叶斯网络结构可以用邻接矩阵的形式表示, 如图 1 表示成邻接矩阵的形式

$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

在本文算法中, 种群个体用邻接矩阵表示. 当个体之间进行交叉和变异操作时, 首先将矩形结构编码的结构矩阵变成行向量后再进行. 如一个贝叶斯网络结构的邻接矩阵为 $X = (x_{ij})$, 它所对应的矩形结构编码的结构矩阵变成行向量后可表示为

$$x_{11}x_{12} \cdots x_{1n}x_{21}x_{22} \cdots x_{2n} \cdots x_{n1}x_{n2} \cdots x_{nn}.$$

从而, 对应图1的邻接矩阵变成行向量后为

0 1 1 0 0 0 0 1 0 0 0 1 0 0 0 0.

下面先对限制型遗传算法(CGA)所涉及的初始种群的设定, 种群个体间的选择、交叉和变异操作做一介绍; 然后给出该算法的具体描述.

首先通过求解如上所述的无约束优化问题, 获得一无向图. 以此无向图为基础网络, 随机生成初始种群, 也即初始种群个体所对应网络结构的边只能在无向图中的边中随机产生, 而无向图中的这些边恰好是最优贝叶斯网络结构的候选边, 从而极大地减弱了单纯使用遗传算法学习对初始群体的依赖性, 而且算法的收敛速度也有很大的提高.

算法采用最佳个体保留法对个体进行选择. 最佳个体保留法能够保证遗传算法终止时得到的结果一定是历代出现过的具有最高适应度的个体.

选择操作选出的个体随机配对, 并按交叉概率 p_c 决定是否进行交叉操作. 交叉操作是产生新个体的主要方法, 决定了遗传算法的全局搜索能力. 本文采用较为简单的单点交叉方法, 例如, 假设两个个体所对应矩形结构编码的结构矩阵变成行向量后分别为 10000100 和 11111111, 随机确定的位置为第4位, 则交叉后新产生个体所对应矩形结构编码的结构矩阵的行向量分别为 10001111 和 11110100.

变异操作可以随机改变个体的性状, 增加群体的多样性以避免局部极值. 具体做法是按变异概率 p_m 随机改变某些个体串的某些基因座的基因值, 即 $1 \rightarrow 0$ 或 $0 \rightarrow 1$. CGA 算法中个体的基因 $\text{Gene}(v_i)$ 表示变量 v_i 在贝叶斯网络中的父亲节点集, 由此可定义算法的基本变异算子为: 逆转有向弧 $X_i \rightarrow X_j$.

此外, 在算法中, 群体规模、遗传操作的交叉概率 p_c 、变异概率 p_m 以及算法的结束条件等应根据算法处理的具体情况设定. 算法终止条件可以设置为: 已经进化了 NC 代或连续 L 代最佳的网络结构没有变化.

CGA 算法描述如下:

1) 初始化. 设定无约束优化中的显著性水平 α , 遗传算法中的交叉概率 p_c 和变异概率 p_m , 迭代次数 NC, 迭代计数器 $t = 0$.

2) 确定无向图. 构造无约束优化问题并求解, 获得一无向图.

3) 产生初始种群. 以无向图为基础, 产生初始种群 $\text{Pop}(0)$.

4) 主循环. 当 $t \leq \text{NC}$ 时:

对种群 $\text{Pop}(t)$ 执行选择、交叉和变异算子, 产生子代群体 $\text{Pop}(\text{son})$.

从 $\text{Pop}(t)$ 和 $\text{Pop}(\text{son})$ 中按最佳个体保留原则选择个体组成新一代群体 $\text{Pop}(t + 1)$.

置 $t = t + 1$.

终止循环.

5) 输出最优贝叶斯网络结构.

3 数值实验

下面以图1为例说明本文方法学习贝叶斯网络结构的过程. 首先由图1中的网络生成 800 条数据, 并计算 1-步依赖系数矩阵, 如表 1 所示.

表 1 1-步依赖系数

	Cloudy	Sprinker	Rain	Wet Grass
Cloudy	0	130.921 4	365.535 8	41.259 8
Sprinker	130.921 4	0	-5.739 7	392.845 5
Rain	365.535 8	-5.739 7	0	392.845 5
Wet Grass	41.259 8	392.845 5	329.265 3	0

然后, 调用 Matlab 中的 `bintprog` 函数求解无约束优化问题 (7), 获得无向图如图 2 所示.

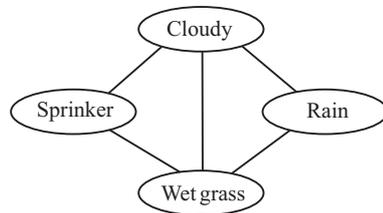


图 2 CGA 方法学习得到的无向图

最后, 以此无向图为基础产生初始种群, 并对初始种群进行选择、交叉和变异等遗传操作学习贝叶斯网络结构. 通过使用 Matlab 7.0, 在数值实验平台为 Pentium 4, 3.06 GHz CPU, 512 M 的微机, 显著性水平 $\alpha = 0.5\%$, 交叉概率 $p_c = 0.99$, 变异概率 $p_m = 0.001$, 种群规模为 200, 迭代次数 $\text{NC} = 15$ 的设置下, 采用 BIC 打分函数, 本文提出的 CGA 方法运行 71.9690 s 后学习得到了图 1 中的真实网络.

为了进一步测试本文方法的性能, 对限制型遗传算法 CGA 和遗传算法 GA 学习贝叶斯网络结构的效果和运行效率做了比较. 实验中除种群规模改为 300 外, 其他参数设置同上. 选取 Asia 网络 (图 3) 为学习对象, 学习结果比较见图 4.

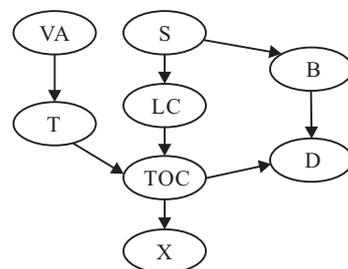


图 3 真实的 Asia 网络结构

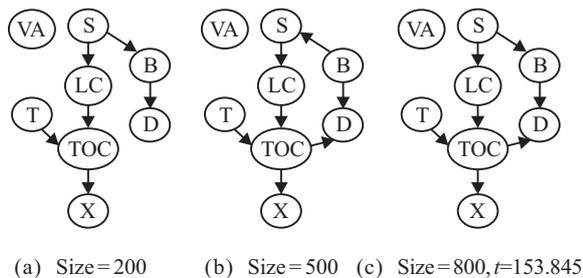


图4 CGA方法在不同样本尺寸下得到的最优贝叶斯网络图

通过比较可以发现,限制型遗传算法CGA在样本尺寸为800,运行时间为153.8450s时即可获得较好的网络结构;而非限制下的遗传算法GA要获得同样好的网络结构则需要样本数据为3000,运行时间为210.9032s.实验结果表明,本文所提出的限制型遗传算法CGA可以在小样本数据下,以较短的时间获得较优的贝叶斯网络结构.此外,两种方法每次迭代的最优贝叶斯网络相应得分随迭代次数的变化结果(图5)显示,CGA方法在开始时即可获得较好的贝叶斯网络结构,这进一步说明,限制型遗传算法CGA在使用遗传算法进行边及边的方向确定前,已经对许多较差的贝叶斯网络结构进行了排除.因此,本文方法无论从学习效果,还是运行效率,都比直接使用遗传算法GA的性能好.

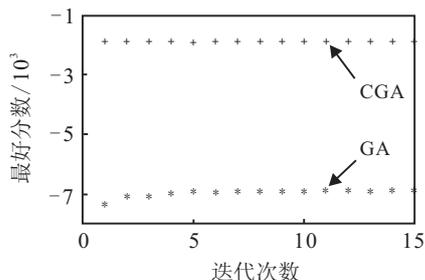


图5 CGA和GA方法在样本尺寸分别为800和3000时每次迭代所得最好分数与迭代次数的关系

4 结论

为有效学习贝叶斯网络结构,本文基于无约束优化和遗传算法提出了一种限制性遗传算法(CGAs).首先,通过引入依赖变量和全局依赖度量等概念,构造了一个无约束优化问题,其最优解对应一无向图;然后以此无向图为基础产生遗传算法的初始种群,进而通过使用遗传算法中的选择、交叉和变异算子学习获得最优贝叶斯网络结构.数值实验表明,本文方法可以在小样本数据下,以较短时间获得较好的结果.关于参数的设置,以及不同交叉算子的使用对算法的影响,将是进一步研究的内容.

参考文献(References)

[1] Pinto P C, Nagele A, Dejori M, et al. Using a local discovery ant algorithm for Bayesian network structure

learning[J]. IEEE Trans on Evolutionary Computation, 2009, 13(4): 767-779.

[2] Sangeeta B E, Shou-Ching S, Marco F R, et al. Use of Bayesian networks to probabilistically model and improve the likelihood of validation of microarray findings by RT-PCR[J]. J of Biomedical Informatics, 2009, 42(2): 287-295.

[3] Chow C K, Lin C N. Approximating discrete probability distribution with dependence tree[J]. IEEE Trans on Information Theory, 1968, 14(3): 462-467.

[4] Sprites P, Glymour C. An algorithm for fast recovery of sparse causal graphs[J]. Social Science Computing Review, 1991, 9(1): 62-72.

[5] Chen X W, Gopalakrishna A, Lin X T. Improving Bayesian network structure learning with mutual information-based node ordering in the k_2 algorithm[J]. IEEE Trans on Knowledge and Data Engineering, 2008, 20(5): 1-13.

[6] Lobna B, Afif M, Faiez G, et al. Improving algorithms for structure learning in Bayesian networks using a new implicit score[J]. Expert Systems with Applications, 2010, 37(7): 5470-5475.

[7] Suzuki J. A construction of Bayesian networks from database based on a MDL scheme[C]. Proc of the 9th Conf on Uncertainty in Artificial Intelligence. San Mateo: Morgan, 1993: 266-273.

[8] Yan L J, Cercone N. Bayesian network modeling for evolutionary genetic structures[J]. Computers and Mathematics with Applications, 2010, 59(8): 2541-2551.

[9] Tsamardinos I, Brown L E, Aliferis C F. The max-min hillclimbing bayesian network structure learning algorithm[J]. Machine Learning, 2006, 65(1): 31-78.

[10] Schulte O, Frigo G, Greiner R, et al. A new hybrid method for Bayesian network learning with dependency constraints[C]. IEEE Symposium on Computational Intelligence and Data Mining. Nashville, 2009: 53-60.

[11] Larrañaga P, Poza M, Yurramendi, et al. Structure learning of Bayesian network by genetic algorithm: A performance analysis of control parameters[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1996, 18(9): 912-926.

[12] Campos L M, Gámez J A, Moral S. Partial abductive belief networks-an evolutionary computation approach by using problem-specific genetic operators[J]. IEEE Trans on Evolutionary Computation, 2002, 6(2): 105-131.

[13] de Campos L. A scoring function for learning Bayesian networks based on mutual information and conditional independence tests[J]. J of Machine Learning Research, 2006, 7(12): 2149-2187.