

一种通过结构边界进行贝叶斯网络学习的算法

刘广怡* 李 鸥 张大龙

(解放军信息工程大学 郑州 450000)

摘要: 贝叶斯网络是智能算法领域重要的理论工具, 其结构学习问题被认为是 NP-hard 问题。该文通过混合学习算法的方式, 从分析低阶条件独立性测试提供的信息入手, 给出了构造目标网络结构空间边界的方法, 并给出了完整的证明。在此基础上执行打分搜索算法获得最终的神经网络结构。仿真结果表明该算法与同类算法相比具有更高的精度和更好的执行效率。

关键词: 贝叶斯网络; 结构学习; 有向无圈图; 条件独立

中图分类号: TP18

文献标识码: A

文章编号: 1009-5896(2015)04-0894-06

DOI: 10.11999/JEIT140786

Learning Bayesian Network from Structure Boundaries

Liu Guang-yi Li Ou Zhang Da-long

(The PLA Information Engineering University, Zhengzhou 450000, China)

Abstract: Bayesian network is an important theoretical tool in the artificial algorithm field, and learning structure from data is considered as NP-hard. In this article, a hybrid learning method is proposed by starting from analysis of information provided by low-order conditional independence testing. The methods of constructing boundaries of the structure space of the target network are given, as well as the complete theoretical proof. A search & scoring algorithm is operated to find the final structure of the network. Simulation results show that the hybrid learning method proposed in this article has higher learning precision and is more efficient than similar algorithms.

Key words: Bayesian Network (BN); Structure learning; Directed acyclic graph; Conditional Independence (CI)

1 引言

作为人工智能领域表示不确定性知识和推理的一种重要方法, 近些年来, 贝叶斯网络(Bayesian Network, BN)一直是众多国内外学者研究的热点, 目前在诸如故障检测、医疗诊断、交通管理等方面已经取得了成功的运用^[1-3]。从贝叶斯网络发展的过程看, 一开始人们关注的是寻求一种数据结构以对联合概率密度进行压缩表示, 并针对该数据结构开发快速的概率推理算法, 因此诞生了贝叶斯网络。在取得成功之后, 人们开始转而关注针对数据的网络结构自学习算法研究。本质上看, 贝叶斯网络的结构学习问题属于组合优化问题, 并已经被证明为 NP-hard^[4], 尽管如此, 一些启发式算法还是得到了成功的应用^[5-7]。

目前, BN 结构学习算法大致可分为基于独立性测试的方法和基于打分-搜索机制的方法。近来, 一些研究者结合上述两种方法提出一类混合算法, 该类算法首先利用条件独立测试(Conditional

Independence testing, CI-testing)降低搜索空间的复杂度, 然后执行评分搜索找到最佳网络。这类算法由于有选择性地压缩了搜索空间, 因此可以提高结构学习算法整体收敛速度, 具有较强的实用性。

对于目前大多数针对贝叶斯网络结构的混合学习算法而言, 都是先通过 CI 测试获得目标网络的局部结构。文献[8]中通过 CI 测试分析目标网络中可能存在的 V 结构, 但是这种方法需要事先获得网络的框架, 这一条件在实际的贝叶斯网络学习中往往不容易获得。文献[9]中提出了一种 MMHC(Max-Min Hill Climbing)算法, 该算法分为 2 部分, 第 1 部分称为 MMPC(Max-Min Parents and Children), 它通过 CI 测试获得每个节点的父节点和子节点集, 从而给出待学习目标网络的一个局部框架, 进而在算法的第 2 部分执行打分搜索确定网络中存在的边及方向。由于算法的第 1 部分使用了高阶 CI 测试, 而这种测试的准确性无法保证^[10], 因此在搜索阶段并没有严格依据 MMPC 给出的先验结构, 而是进行了相当程度的开放搜索, 这样就导致了计算资源的浪费。文献[11]提出了 BNEA 算法, 它是一种基于 MMHC 和最大主子图分解(Maximal Prime Decomposition, MPD)的方法, 通过 MMHC 给出的

2014-06-17 收到, 2014-12-20 改回

国家科技重大专项(2014ZX03006003)资助课题

*通信作者: 刘广怡 liu.guangyi@outlook.com

节点 Markov 边界^[12]进行 MPD, 从而在较低的维度上尝试获得网络的 Markov 等价结构。在一定程度上, BNEA 算法获得的 Markov 等价结构比直接使用 MMPC 得到的局部结构要精确, 但由于调用 MMPC 算法无法避免高阶 CI 测试, 因此 BNEA 算法在计算量和精度的平衡上还有待进一步提高。

本文提出基于混合方式的一种上下界候选学习 (Upper-lower Bounds Candidate sets Searching, UBCS) 算法, 该方法首先通过低阶 CI 测试确定待学习网络结构空间的上下界, 从而提供一种更有指导意义的候选网络集, 在此基础上使用贪婪搜索算法确定最终的网络结构, 仿真结果表明该方法能够在保证精度的同时有效降低学习算法的时间复杂度。

2 相关知识

2.1 相关定义

定义 1 贝叶斯网络定义 给定随机变量 x_1, x_2, \dots, x_n 的联合分布 P , 和有向无圈图 $G = (V, E)$, 若 $V = \{v_1, v_2, \dots, v_n\}$ 与随机变量 x_1, x_2, \dots, x_n 一一对应, 且 (G, P) 满足 Markov 条件^[13], 则称二元组 $BN = (G, P)$ 为一个贝叶斯网络。

由于贝叶斯网络中节点和随机变量一一对应, 因此本文中对两者不加区分, 除特殊说明外一律统称节点。此外, 对于图结构中的有向边 $v_i \rightarrow v_j$ 用 e_{ij} 表示, 无向边 $v_i - v_j$ 用 e_{ij} 表示, 下文不再赘述。

定义 2 V 结构 贝叶斯网络 $BN = (G, P)$ 中, $\forall v_i, v_j, v_k \in V$, 若 $e_{ik} \in E$, $e_{jk} \in E$, $e_{ij} \notin E$ 且 $e_{ji} \notin E$, 则称 v_i, v_j, v_k 构成 V 结构。

定义 3 条件互信息 $MI(X, Y | Z)$ 随机变量集 (X, Y, Z) 的条件互信息 $MI(X, Y | Z)$ 定义为

$$\begin{aligned} MI(X, Y | Z) &= H(X | Z) - H(X | Z, Y) \\ &= \sum_{x_i \in X, y_i \in Y} P(x_i, y_j | Z) \\ &\quad \cdot \lg \frac{P(x_i, y_j | Z_k)}{P(x_i | Z)P(y_j | Z)} \end{aligned} \quad (1)$$

其中

$$H(X | Z) = \sum_{x_i \in X, z_k \in Z} P(x_i, z_k) \lg \frac{1}{P(x_i | z_k)} \quad (2)$$

$$\begin{aligned} H(X | Z, Y) &= \sum_{x_i \in X, y_j \in Y, z_k \in Z} P(x_i, y_j, z_k) \\ &\quad \cdot \lg \frac{1}{P(x_i | y_j, z_k)} \end{aligned} \quad (3)$$

由于 $MI(X, Y | Z) = 0$ 意味着, 随机变量集 X, Y 在给定 Z 时条件独立, 即 $\text{Ind}(X, Y | Z)$, 因此一般用 $MI(X, Y | Z)$ 进行随机变量的条件独立测试 (CI 测

试), 并将 $\|Z\|$ 称为 CI 测试的阶数, 若 $Z = \Phi$, 称为 0 阶 CI 测试。

定义 4 Markov 等价 两个具有相同节点集的有向无圈图 G_1 和 G_2 图等价, 当且仅当: (1) G_1 和 G_2 具有相同骨架; (2) G_1 和 G_2 具有相同的 V 结构。

称两个贝叶斯网 $BN_1 = (G_1, P_1)$ 和 $BN_2 = (G_2, P_2)$ Markov 等价, 若图 G_1 和 G_2 图等价。

按上述关系可以将同一个结点集上的所有有向无环图分成多个等价类, 称作 Markov 等价类, 每一个等价类唯一决定一个统计模型, 且每一个等价类可以用一个部分有向无环图来表示, 称之为完备 PDAG。文献[14]给出了 Markov 等价的特征^[14], 文献[15]将此特征扩展到有向无环图, 并证明两个有向无环图是 Markov 等价当且仅当它们有同样的构架, 并且有同样的“V”结构。

3 算法描述

贝叶斯网络结构学习算法是在给定数据集 D 的情况下寻找贝叶斯网 $BN = (G, P)$ 的最佳网络结。文献[16]证明了包含 n 个节点可能的 BN 结构数为

$$f(n) = \sum_{i=1}^n (-1)^{i+1} C_n^i 2^{i(n-i)} f(n-1) \quad (4)$$

从式(4)可以看出, 随着节点的增加, 可能的网络结构空间呈指数级增长, 因此, 寻找好的候选网络结构集是有效降维的办法。基于此, 本文给出一种 UBCS 算法, 该算法通过构造目标网络 PDAG 的上下界给出候选网络集合, 并通过打分搜索得到最终网络模型。下文中首先给出 UBCS 的第 1 部分 SGLA 算法, 并证明其输出结果 G^+ 是目标网络道德图的上界, 之后引入 0 阶互信息量不增原理, 从而给出 UBCS 的第 2 部分 LGLA 算法, 即目标网络 PDAG 下界的学习算法, 在这之后讨论搜索算法。

3.1 候选网络学习算法

首先给出 SGLA 描述(表 1)。

由 SGLA 算法过程可以知道 G^+ 是弦化图, 对于弦化图有定理 1 成立:

定理 1 对于任意无向图 G , 若它是完备 PDAG 当且仅当 G 是弦化的。

定理证明参见文献[17]。由定理知 G^+ 是完备 PDAG, 即在最好的情况下, 由 SGLA 算法得到的 G^+ 就是目标网络的完备 PDAG, 但此条件太强, 下面给出一个更为一般性的定理。

定理 2 给定数据集 D , 设待求贝叶斯网 $BN = (G, P)$ 在数据集 D 下可能的结构为 G_D , 其对应的道德图为 G_{D_m} , 则由 SGLA 算法生成的无向图 G^+ 是由 G_{D_m} 构成的偏序集 $GM = (G_{D_m}, \subseteq)$ 的上界。

表1 SGLA 算法

算法 1 SGLA	
步骤 1	输入: 数据集 D ; 变量集 $V = \{v_1, v_2, \dots, v_n\}$;
步骤 2	初始化: 无向图 $G^+ = (V, E)$, 其中 $E = \Phi$;
步骤 3	0 阶 CI 测试:
for	每一对变量 $v_i, v_j \in V, i = 1, 2, \dots, n$
if	$\text{Ind}(v_i, v_j)$ 成立, 则 $E = E \cup \{e_{ij}\}$, 其中 $e_{ij} = (v_i, v_j)$ 为无向边
end for;	
步骤 4	1 阶 CI 测试
For	每一个有序三元组 $v_i, v_j, v_k \in V$
If	$e_{ij} \in E, e_{ik} \in E, e_{jk} \notin E$, and $\neg \text{Ind}(v_j, v_k v_i)$ 成立, 则 $E = E \cup \{e_{jk}\}$
End for;	
步骤 5	三角化无向图 G^+ , 得到 G_T^+ ;
步骤 6	调用 Junction tree 算法对 G_T^+ 进行最大主子图分解得到 $G_{\text{sub}}^1, G_{\text{sub}}^2, \dots, G_{\text{sub}}^k$;
步骤 7	输出 G^+ and $G_{\text{sub}}^1, G_{\text{sub}}^2, \dots, G_{\text{sub}}^k$.

证明 只需证明 $\forall G_{D_{-m}}^i \in \text{GM}$, 有 $G_{D_{-m}}^i \subseteq G^+$, 其中 $G_{D_{-m}}^i = (V, E_{D_{-m}}^i), G^+ = (V, E^+)$ 。由定理 1 知, 当 G_D 的完备 PDAG 为弦化图时, $G_{D_{-m}}^i = G^+$ 。接下来要证明 $G_{D_{-m}}^i \subset G^+$, 由于所涉及的所有图中都具有相同的节点, 因此只需要证明对任意无向边 $e_{jk} \in E_{D_{-m}}^i$, 有 $e_{jk} \in E^+$ 。由于 $E_{D_{-m}}^i$ 中所有的无向边一共可以分为两类, 一类是由 G_D 中的有向边无向化得到的, 设为 $\bar{E}_{D_{-m}}^i$, 另一类是 G_D 中具有相同子节点的节点间添加的道德边, 设为 $\tilde{E}_{D_{-m}}^i$ 。对于 $\forall e_{jk} \in \bar{E}_{D_{-m}}^i$, 0 阶 CI 测试保证了一定有 $e_{jk} \in E^+$; 对于 $\forall e_{jk} \in \tilde{E}_{D_{-m}}^i$, 虽然 0 阶 CI 测试不能保证一定有 $e_{jk} \in E^+$, 但 SGLA 算法第 4 步保证了此时一定有 $e_{jk} \in E^+$ 。证毕

定理 3 存在于 $\text{BN} = (G, P)$ 中的任意 V 结构, 一定存在于 G^+ 的 MPD 分解^[18]的某一子图中。

定理 3 的证明参见文献[11], 该定理保证了 SGLA 输出结果中 $G_{\text{sub}}^1, G_{\text{sub}}^2, \dots, G_{\text{sub}}^k$ 覆盖原图所有的 V 结构, 为下文将要给出的 LGLA 算法提供初始值。

上文讨论了 $\text{GM} = (G_D, \subseteq)$ 的上界, 从而为搜索提供了可能的备选, 下面讨论待学习网络结构空间的下界, 从而为搜索算法给出一个较为精确的初始值。

首先引入一个引理:

引理 1 对于任意两个离散随机变量 $v_i, v_j \in V$ 和 V 的子集 $S \subseteq V \setminus \{v_i, v_j\}$, 有 $H(v_i | v_j, S) \leq H(v_i | v_j)$ 成立, 等号成立的条件为当且仅当: $\text{Ind}(v_i, v_j | S)$ 。

证明从略。

定理 4 给定贝叶斯网 $B = (G, P)$, 其中 $G =$

(V, E) , 对于任意 $v_i, v_j, v_k \in V$, 有 $\text{MI}(v_i, v_k) = \min \left\{ \bigcup_{v_l, v_n \in V', v_l \neq v_n} \text{MI}(v_l, v_n) \right\}$, 其中 $V' = \{v_i, v_j, v_k\} \subseteq V$

成立, 如果存在 $e_{ij} \in E, e_{kj} \in E, e_{ik} \notin E, e_{ki} \notin E$ 。

证明 由于存在 $e_{ij} \in E, e_{kj} \in E$, 不失一般性, 不妨设 $\text{MI}(v_i, v_j) = \min\{\text{MI}(v_i, v_j), \text{MI}(v_k, v_j)\}$, 只需要证明 $\text{MI}(v_i, v_k) > \text{MI}(v_i, v_j)$, 由互信息定义: $\text{MI}(v_i, v_k) = H(v_i) - H(v_i | v_k)$, $\text{MI}(v_i, v_j) = H(v_i) - H(v_i | v_j)$, 则有

$$\begin{aligned} \text{MI}(v_i, v_k) - \text{MI}(v_i, v_j) &= H(v_i) - H(v_i | v_k) \\ &\quad - H(v_i) + H(v_i | v_j) \\ &= H(v_i | v_k) - H(v_i | v_j) \end{aligned} \quad (5)$$

由 v_i, v_j, v_k 关系可知节点 $v_j \in S$, 其中 S 满足 $\text{Ind}(v_i, v_k | S)$, 因此式(5)可以写为:

$$\begin{aligned} &\text{MI}(v_i, v_k) - \text{MI}(v_i, v_j) \\ &= H(v_i | v_j, \{v_k\} \cup S / v_j) - H(v_i | v_j) \end{aligned}$$

由引理 1 可得 $\text{MI}(v_i, v_k) \leq \text{MI}(v_i, v_j)$, 其中等号成立的条件是 $S / v_j = \Phi$ 。证毕

定理 4 称为 0 阶互信息量不增性原理。从定理的条件可知, 对于存在 V 结构的情况该定理是不适用的。即对于如图 1 所示的结构而言, 定理 4 无法判断 $\text{MI}(a, d)$ 和 $\text{MI}(a, b)$ 的大小关系。如果首先在网络中排除所有的 V 结构, 再来考虑节点 a 和节点 cbe 的可能连接关系, 若 $\text{MI}(a, c)$ 最大, 则一定有 $\text{MI}(a, c) = \text{MI}(a, b) > \text{MI}(a, e)$, 此时由 1 阶 CI 测试可以直接排除掉 ac 连接的可能性, 因此就能直接确定 ab 是相连的。通过以上分析可以看出, 0 阶互信息量不增性原理实际上提供了一种除 V 结构外的确定网络中无向边存在性的思路。

基于以上分析给出 G^- 学习算法(LGLA)示于表 2。

LGLA 算法中涉及的 V 结构测试算法 VSTA (V-Structure Test Algorithm)描述如表 3 所示。

VSTA 是一种“尽力而为”的检测算法, 由于仅使用了 0 阶和 1 阶 CI 测试, 其低阶检测的准确性保证了检测出的 V 结构的存在性, 而对于构成 V 结构的两个父节点间通路多于一条时, 将不予检测。这种检测方式避免了高计算量和干扰边的引入。

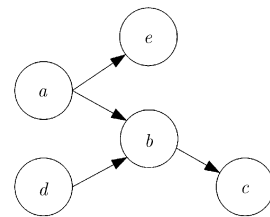


图1 BN 中存在 V 结构的情况

表 2 LGLA 算法

算法 2 LGLA

步骤 1 输入：数据集 D ；变量集 $V = \{v_1, v_2, \dots, v_n\}$ ；
 $G_{\text{sub}}^1, G_{\text{sub}}^2, \dots, G_{\text{sub}}^k$ ；

步骤 2 初始化： $G^- = (V, E^-)$ ，其中 $E^- = \Phi$ ；

步骤 3 For 每一个 G_{sub}^i 若其中存在圈，调用 VSTA；

步骤 4 For 每一个 $v_i \in V$ ，if $v_j \in V$ 使得 $MI(v_i, v_j)$ 最大，则 $E^- = E^- \cup \{e_{ij}\}$ ；

步骤 5 If 存在另一个 $v_k \in V$ 使得 $MI(v_i, v_j) = MI(v_i, v_k)$ 最大
 If $\text{Ind}(v_i, v_k | v_j)$ is hold, then
 $E^- = E^- \cup \{e_{ik}\} / \{e_{ij}\}$
 End if;

步骤 6 输出 $G^- = (V, E^-)$ 做为最终结果。

表 3 VSTA 算法

算法 3 VSTA

步骤 1 输入：数据集 D ；变量集 $V = \{v_1, v_2, \dots, v_n\}$ ； $G^- = (V, E^-)$ ； G_{sub}^i ，其中 $G_{\text{sub}}^i[V] = \{v_i, v_j, v_k\}$ ；

步骤 2 If $\text{Ind}(v_i, v_j)$ and $\neg \text{Ind}(v_i, v_j | v_k)$
 在 G^- 中增加 $v_i \rightarrow v_j$ 并删除 $v_i - v_j$
 Else if $\neg \text{Ind}(v_i, v_j | v_k)$ and $\text{Ind}(v_i, v_j | v_l)$,
 $v_l \in V / \{v_i, v_j, v_k\}$ ，则
 在 G^- 中增加 $v_i \rightarrow v_k$ 并删除
 $v_i - v_j$ in G^- ，
 End if;

步骤 3 输出 G^- 。

对于 LGLA 的输出结果 G^- ，有定理 5 成立。

定理 5 设待求贝叶斯网 $BN = (G, P)$ 中 G 的所有可能的完备 PDAG 结构和包含关系构成偏序集 $PG = (G_{\text{PDAG}}, \subseteq)$ ，则由 LGLA 算法生成的 G^- 是 PDAG。

证明 G^- 包含有向边和无向边，对于无向边，上文给出的 0 阶互信息量不增原理保证了对任意无向边 $e_{ij} \in G^-[E^-]$ 一定有 $e_{ij} \in G[E]$ 成立。 G^- 中的有向边由 VSTA 给出。由定理 3 知 $BN = (G, P)$ 中的任意 V 结构，一定存在于某个 G_{sub}^i 中，由 VSTA 算法的性质可知任意存在于 G^- 中的 V 结构一定存在于 G 中，反之则不一定成立。对于无圈图而言，删除任意有向边得到的仍是无圈图。因此 G^- 是 PDAG。

证毕

定理 6 当 VSTA 返回结果完全准确时， G^- 是 PG 的下界。

证明略。

定理 6 的条件较强。事实上，在很多情况下，可以近似认为 G^- 是 PG 的下界。以 Asia 网络为例。

通过图 2 可以看到， G^- 中的任意边都存在于原始网络的 PDAG 中。

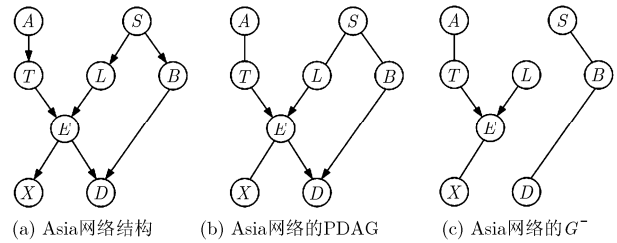


图 2 Asia 网络结构及其对应的 PDAG 和 G^-

3.2 打分搜索学习算法

爬山法是关于贝叶斯网络结构学习中基于打分搜索的一种常用贪婪搜索算法，其搜索算子共有 3 种：加边、减边和反转边。在 UDCS 算法中，同样使用爬山搜索算法，只是搜索过程受到 SGLA 和 LGLA 算法给出的上下界的限制，即搜索算子产生的新的网络结构如果超出 SGLA 和 LGLA 给定的边界，则丢弃该结构。为了避免搜索过快的陷入局部最优结构，引入次优竞争机制，每一轮搜索保留前 N 个得分最高的结构进入下一轮迭代。原则上 N 的大小由网络结构的规模决定，网络规模越大， N 越大，但过大的次优候选集将导致算法时间复杂度的增加。对于规模如 Alarm 网络的待学习网络而言，推荐经验值 $N=10$ 。为了便于与文献[11]中的 BENA 算法进行比较，采用 BDeu 评分函数作为搜索的目标函数。

4 实验结果与分析

使用标准的 Alarm 网络对本算法进行测试，并与 BNEA 和 MMHC 算法进行比较。首先比较的是在不同样本数 3 种算法得到的网络结构的评分，如表 4 所示，其中，为了便于对比观察，将评分结果进行了归一化处理，仿真结果示于表 4。

表 4 的结果是重复实验 10 次的平均结果。从表 4 中可以看出，UBCS 算法得到的网络结构评分在小样本数量下优于 BENA 和 MMHC，随着样本数的增加，3 种算法学习出的网络结构评分趋于一致，在样本数达到 5000 时已经和真实网络的评分非

表 4 UBCE, BENA, MMHC 在 Alarm 网络数据集上的评分对比

算法	样本数 SS			
	500	1000	2000	5000
UBCS	1.0000	1.0000	1.0000	1.0000
BENA	1.0801	1.0309	1.0286	0.9991
MMHC	1.2000	1.0981	1.0848	1.0029

常接近。UBCS 算法中包含 SGLA 和 LGLA 两个主要算法,其中 LGLA 中由于 VSTA 算法并不能充分保证检测结果真实性,但是从仿真结果来看,对最终学习性能几乎没有影响,这一方面由于 SGLA 提供的上界具有很高的稳定性,另一方面在受约束的打分搜索学习过程中这种影响被进一步减弱所致。应当指出的是算法在样本数较大时(样本数大于 5000 以上)容易出现过学习现象,一般认为过学习现象出现在小样本阶段,但对于如贝叶斯网络结构学习而言的高维组合优化问题,由于组合空间的指数级增长,一般情况下很难使用充分的样本进行学习,且目前的算法在巨大样本数面前所付出的时间性能的代价也是不容忽视的问题,因此,对于维数较高的贝叶斯网络结构学习问题,应当探索在适当样本数下精度和泛化能力平衡的学习算法。UBCS 算法中由于 SGLA 和 LGLA 的使用对泛化能力起到了一定的保证,因此在学习过程中没有发现过学习现象的发生。

3 种算法所消耗的时间如图 3 算法所示。仿真使用的计算机是主频为 2.50 GHz 的 Intel 4 核处理器,内存 4G。从图 3 算法中可以看出,UBCS 相比于其他两种算法在时间复杂度上具有明显的优势,与 MMHC 相比,BNEA 和 UBCS 都使用了 MDP 分解降低了搜索空间维度,从而压缩了学习时间,而 BNEA 由于调用了 MMHC 中的 MMPC 算法,在最坏情况下 BNEA 可能与 MMHC 具有相同的复杂度,BNEA 的空间分解从头至尾仅依赖分解空间中的 0 阶和 1 阶 CI 测试,因而具有更好的时间复杂度性能。

5 结束语

本文提出了一种混合方式的贝叶斯网络结构学习算法(UBCS)。该算法利用低阶的 CI 测试获得待学习网络结构空间的上下界,从而提供一种较好的约束结构,进而通过打分搜索的方式确定网络的最终结构。理论证明和仿真结果都表明了该算法的有

效性。与同类混合结构学习算法相比,UBCS 算法具有时间复杂度上的明显优势,并可以达到令人较为满意的学习精度。下一步的研究目标为进一步完善 UBCS 算法中的上下界学习算法(SGLA 和 LGLA),并尝试应用到诸如增量学习和不完备数据下的学习等贝叶斯网络结构学习的其他问题中。

参考文献

- [1] Pourali M and Mosleh A. A functional sensor placement optimization method for power systems health monitoring[J]. *IEEE Transactions on Industry Applications*, 2013, 49(4): 1711-1719.
- [2] Andrzej W and Edyta W. Integrating Bayesian networks into fuzzy hypothesis testing problem-case based presentation[C]. 2013 IEEE 15th International Conference on e-Health Networking, Application & Services (Healthcom), Lisbon, 2013: 100-104.
- [3] Neumann T, Bohnke, P L, and Tcheumadjeu T. Dynamic representation of the fundamental diagram via Bayesian networks for estimating traffic flows from probe vehicle data[C]. 2013 16th International IEEE Conference on Intelligent Transportation Systems (ITSC), The Hague, 2013: 1870-1875.
- [4] Chickering D M, Geiger D, and Heckerman D. Learning Bayesian networks is NP-hard[R]. Technical Report MSR-TR-94-17, Microsoft Research, 1994: 1-22.
- [5] Carvalho A. A cooperative coevolutionary genetic algorithm for learning bayesian network structures[OL]. <http://arxiv.org/abs/1305.6537>. 2013: 1131-1138.
- [6] Aouay S, Jamoussi S, and Ben Ayed Y. Particle swarm optimization based method for Bayesian Network structure learning[C]. 2013 5th International Conference on Modeling, Simulation and Applied Optimization (ICMSAO), Hammamet, 2013: 1-6.
- [7] Basak A, Brinster I, Ma X, et al. Accelerating Bayesian network parameter learning using Hadoop and MapReduce [C]. Proceedings of the 1st International Workshop on Big Data, Streams and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications, ACM, Beijing, 2012: 101-108.
- [8] 贾海洋, 陈娟, 朱允刚, 等. 基于混合方式的贝叶斯网弧定向算法[J]. *电子学报*, 2009, 37(8): 1842-1847.
Jia Hai-yang, Chen Juan, and Zhu Yun-gang. A Hybrid method for orienting edges of Bayesian network[J]. *Acta Electronica Sinica*, 2009, 37(8): 1842-1847.
- [9] Tsamardinos I, Brown L F, and Aliferis C F. The max-min hill climbing BN structure learning algorithm[J]. *Machine Learning*, 2006, 65(1): 31-78.

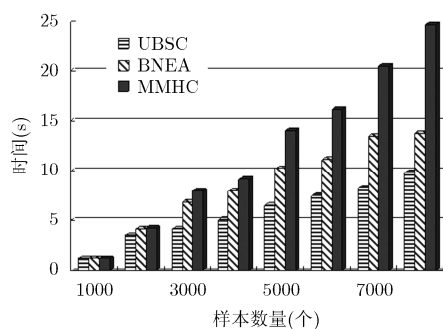


图3 算法时间复杂度比较

- [10] Wong M L and Leung K S. An efficient data mining method for learning Bayesian networks using an evolutionary algorithm-based hybrid approach[J]. *IEEE Transactions on Evolutionary Computation*, 2004, 8(4): 378-404.
- [11] 朱明敏, 刘三阳, 杨有龙. 基于混合方式的贝叶斯网络等价类学习算法[J]. 电子学报, 2013, 41(1): 98-104.
Zhu Ming-min, Liu San-yang, and Yang You-long. Structural learning bayesian network equivalence classes based on a hybrid method[J]. *Acta Electronica Sinica*, 2013, 41(1): 98-104.
- [12] De Morais S R and Aussem A. A novel Markov boundary based feature subset selection algorithm[J]. *Neurocomputing*, 2010, 73(4): 578-584.
- [13] Neapolian R E. Learning Bayesian Networks [M]. Englewood Cliffs, NJ, US, Prentice-Hall, Inc., 2004: 31-39.
- [14] Frydengberg M. The chain graph markov property[J]. *Scandinavian Journal of Statistics*, 1990, 17(4): 333-353.
- [15] Verma T and pearl J. An algorithm for deciding if a set of observed independencies has a causal explanation[C]. Proceedings of the Eighth Annual Conference on Uncertainty in Artificial Intelligence, Stanford, California, 1992: 323-330.
- [16] Robinson R W. Counting Unlabeled Acyclic Digraphs[M]. Berlin Heidelberg, Springer, 1977: 28-43.
- [17] Andersson S A, Madigan D, and Perlman M D. A characterization of Markov equivalence classes for acyclic digraphs[J]. *The Annals of Statistics*, 1997, 25(2): 505-541.
- [18] Wu D. Maximal prime subgraph decomposition of Bayesian networks: a relational database perspective[J]. *International Journal of Approximate Reasoning*, 2007, 46(2): 334-345.
- 刘广怡: 男, 1982年生, 博士生, 研究方向为网络数据智能处理、智能算法、贝叶斯网络.
- 李 鸥: 男, 1961年生, 教授, 博士生导师, 研究方向为人工智能、计算机网络协议、认知网络.
- 张大龙: 男, 1976年生, 博士, 讲师, 研究方向为网络数据智能处理、网络协议分析.