

基于V-结构&对数似然函数定向与禁忌爬山的贝叶斯网络结构算法

刘浩然^{*①②} 王念太^{①②} 王毅^{①②} 张力悦^{①②} 苏昭玉^{①②} 刘文^③ 赵旭丹^③

^①(燕山大学信息科学与工程学院 秦皇岛 066004)

^②(河北省特种光纤与光纤传感重点实验室 秦皇岛 066004)

^③(北京市机电研究院 北京 100027)

摘要: 针对爬山算法搜索空间过大和易陷入局部最优的问题, 该文提出基于V-结构&对数似然函数定向与禁忌爬山的贝叶斯网络结构算法(VTH)。该算法利用定向最大支撑树约束搜索空间, 在最大支撑树定向过程中, 提出V-结构与对数似然函数(VLL)结合的定向策略; 在评分搜索过程中, 提出禁忌爬山(VTH)评分搜索策略, 该策略将禁忌表清空机制与爬山搜索的局部择优准则结合, 在提高全局寻优能力的同时也能保证搜索效率。该算法与其他算法在Asia, Car, Child和Alarm 4种标准网络中进行仿真实验, 对比汉明距离、F1值、平衡评分函数(BSF)值、运行时间4个指标, 验证了该算法的有效性。

关键词: 贝叶斯网络结构; 爬山算法; 禁忌搜索; 定向最大支撑树

中图分类号: TP18

文献标识码: A

文章编号: 1009-5896(2021)11-3272-10

DOI: 10.11999/JEIT210032

Bayesian Network Structure Algorithm Based on V-structure & Log-Likelihood Orientation and Tabu Hill Climbing

LIU Haoran^{①②} WANG Niantai^{①②} WANG Yi^{①②} ZHANG Liyue^{①②}

SU Zhaoyu^{①②} LIU Wen^③ ZHAO Xudan^③

^①(School of Information Science and Engineering, Yanshan University, Qinhuangdao 066004, China)

^②(The Key Laboratory for Special Fiber and Fiber Sensor of Hebei Province, Yanshan University, Qinhuangdao 066004, China)

^③(Beijing Institute of Mechanical and Electrical Engineering, Beijing 100027, China)

Abstract: Hill climbing algorithm has too large search space and is easy to fall into local optimum. In this paper, a new Bayesian network structure algorithm based on V-structure & log-likelihood orientation and Tabu Hill (VTH) climbing is proposed. The algorithm limits the search space by using the oriented maximum weight spanning tree. In the process of maximum weight spanning tree orientation, the orientation strategy based on V-structure and Log-Likelihood (VLL) function is proposed. Tabu Hill Climbing (THC) scoring search strategy is established during the process of search, it combines the tabu list clearing mechanism with the local optimization criteria of hill climbing, the strategy not only ensures the search efficiency, but also improves the global optimization ability. By comparing Hamming distance, F1-value, Balanced Scoring Function(BSF) value and Time with other algorithms in Asia, Car, Child and Alarm standard networks, the effectiveness of the proposed algorithm is verified.

Key words: Bayesian network structure; Hill climbing algorithm; Tabu search; Oriented maximum weight spanning tree

收稿日期: 2021-01-11; 改回日期: 2021-04-21; 网络出版: 2021-05-07

*通信作者: 刘浩然 liu.haoran@ysu.edu.cn

基金项目: 国家重点研发计划(2019YFB1707301), 河北省人才工程培养资助项目(A201903005)

Foundation Items: The National Key R&D Program of China (2019YFB1707301), The Hebei Talent Engineering Training Support Project (A201903005)

1 引言

贝叶斯网络(Bayesian Networks, BN)源于人们对人工智能领域不确定性问题的研究,是进行不确定性问题推理和数据分析的重要工具^[1]。它将有向无环图与概率论知识相结合,既直观揭示问题关联结构又降低概率计算复杂度,在工业生产^[2]、医疗卫生^[3]、生态学^[4]、金融分析^[5]等众多领域得到广泛应用。

贝叶斯网络学习分为结构学习和参数学习两部分,贝叶斯网络结构对参数学习和推理的精度有较大影响,因此从数据中学习精确的贝叶斯网络结构成为研究热点,但从数据中学习网络结构已被证明为非确定性多项式时间难题(Non-deterministic Polynomial-time hard, NP-hard),为此,国内外专家提出了许多贝叶斯网络结构学习算法,其中最常见的是基于评分搜索的算法^[6]。

爬山算法^[7]是经典的基于评分搜索的贝叶斯网络结构学习算法,其通过局部择优的启发式搜索策略提高了搜索效率,但爬山算法搜索空间随变量数量增加呈指数增长和易陷入局部最优的问题也较为突出^[8]。针对爬山算法存在的问题,许多学者提出了不同的改进算法,Tsamardinos等人^[9]提出最大最小爬山(Max-Min Hill-Climbing, MMHC)算法,该算法利用条件独立性构建网络的结构框架,利用爬山搜索获得最终网络结构,有效约束了搜索空间,但其输出结果的质量仍有较大提升空间;冀俊忠等人^[10]提出基于禁忌搜索的贝叶斯网络结构学习算法,该算法利用禁忌搜索策略提高了算法全局寻优能力,但并没有解决搜索空间过大的问题;Gómez等人^[11-13]先后提出约束爬山(Constrained Hill Climbing, CHC)算法、快速约束爬山(Fast Constrained Hill Climbing, FastCHC)及其改进算法,以上算法虽然通过动态限制候选结构缩小了搜索空间,但学习到的结构准确性仍有待提高;刘浩然等人^[14]提出最大支撑树(Maximum Weight Spanning Tree, MWST)与改进爬山策略相结合的简化爬山(Simplify Hill-Climbing, SHC)算法,该算法虽然约束了搜索空间,但需要标准节点顺序作为最大支撑树定向条件,此外,SHC中缺少减边操作导致结构中多边不能被消除,冗余边过多问题有待解决;刘彬等人^[15]利用最大支撑树约束搜索空间来提高搜索效率,利用改进的交叉算子和变异算子替代爬山搜索中的减边算子来扩展全局寻优能力,但其默认按照标准节点顺序对最大支撑树定向,且全局寻优能力有待提高;Constantinou^[16]利用扩展最大生成图(Extended Maximum Spanning Graph, EMSG)约束搜索空间,在评分搜索中将爬山搜索

得到的结构作为禁忌搜索策略的输入,经过禁忌搜索后得到最终结构,由于爬山搜索与禁忌搜索在搜索过程中独立进行,仍不能解决爬山算法易陷入局部最优的问题。

为约束爬山算法搜索空间并提高其全局寻优能力,本文提出基于V-结构&对数似然函数定向与禁忌爬山的贝叶斯网络结构算法(Bayesian network structure algorithm based on V-structure & log-likelihood orientation and Tabu Hill climbing, VTH)。VTH利用定向最大支撑树约束搜索空间,为解决最大支撑树定向问题,提出V-结构与对数似然函数结合(V-structure & Log-Likelihood, VLL)的定向策略;为提高算法全局寻优能力,提出禁忌爬山(Tabu Hill Climbing, THC)评分搜索策略,THC将禁忌表清空机制与爬山策略的局部择优准则结合,提高全局寻优能力的同时也能保证搜索效率。仿真实验结果表明,VTH算法相比于其他算法在没有大幅增加时间开销的情况下提高了全局寻优能力。

2 VTH算法研究

在VTH算法中,首先计算网络节点间的互信息,进而生成最大支撑树,然后利用条件互信息与互信息的差值构造判别参数,通过判别参数检验出最大支撑树中所有的V-结构并定向,对剩余未定向的边利用对数似然函数进行定向,从而得到完全定向的最大支撑树;以定向最大支撑树作为初始结构进入爬山搜索阶段,在爬山搜索达到停止条件后,以当前最优结构为搜索起点进入禁忌搜索阶段,若此阶段中搜索到更优的结构,则清空禁忌表,并以当前最优结构为搜索起点再回到爬山搜索阶段;否则,在禁忌表达到一定长度时,停止搜索,输出当前最优结构。

2.1 VTH算法构建

将贝叶斯网络结构表示为有向无环图结构 $G = (\mathbf{V}, \mathbf{E})$, $\mathbf{V} = (X_1, X_2, \dots, X_n)$ 表示网络节点 $X_i (i = 1, 2, \dots, n)$ 的集合, \mathbf{E} 表示网络节点间有向边构成的集合^[17]。设变量 e_{ij} 表示任意两个节点 X_i 和 X_j 之间的连接关系, e_{ij} 取值范围为 $\{-1, 0, 1, \infty\}$, 当存在有向边 $X_j \rightarrow X_i$ 时 e_{ij} 取 -1 , 存在无向边 $X_i - X_j$ 时 e_{ij} 取 0 , 存在有向边 $X_i \rightarrow X_j$ 时 e_{ij} 取 1 , X_i 和 X_j 之间不存在边时 e_{ij} 取 ∞ , 如式(1)所示

$$e_{ij} = \begin{cases} -1, & \text{存在边 } X_j \rightarrow X_i \\ 0, & \text{存在边 } X_i - X_j \\ 1, & \text{存在边 } X_i \rightarrow X_j \\ \infty, & X_i \text{ 和 } X_j \text{ 间无边} \end{cases} \quad (1)$$

计算节点集合V中任意两节点 X_i 和 X_j 的互信息 $I(X_i; X_j)$ ^[18], 如式(2)所示, 其中, $P(X_i, X_j)$ 表示节点 X_i 和 X_j 的联合概率, $P(X_i), P(X_j)$ 分别表示节点 X_i, X_j 的概率。取节点集合V中任意节点 X_r 加入最大支撑树^[19]节点集合 $U = \{\}$, 从集合 $V - U$ 中搜索节点 X_t , 使其满足 $I(X_t; X_u)(X_u \in U)$ 最大, 如式(3)所示, 将搜索到的节点 X_t 加入 U , 将 $e_{tu} = 0$ 加入最大支撑树边集合 E_T , 以此类推, 当 U 中包含网络中全部节点时, 得到最大支撑树 $T = (U, E_T)$

$$I(X_i; X_j) = \sum_{X_i, X_j} P(X_i, X_j) \lg \frac{P(X_i, X_j)}{P(X_i)P(X_j)} \quad (2)$$

$$X_t = \arg \max_{X_t \in V-U} (I(X_t; X_u)), \forall X_u \in U \quad (3)$$

搜索 $T = (U, E_T)$ 中满足 X_i, X_j 与 X_k 相连且 X_i 和 X_j 不相连的3节点组合, 3节点连接如图1所示。根据节点 X_i, X_j 在条件 X_k 下的条件互信息^[18]与互信息 $I(X_i; X_j)$ 的差值构造判别参数 Δ 如式(4)所示

$$\Delta = \sum_{x_k \in X_k} \sum_{x_j \in X_j} \sum_{x_i \in X_i} Q(x_i, x_j, x_k) - I(X_i; X_j) \quad (4)$$

$$Q(x_i, x_j, x_k) = P_{X_i, X_j, X_k}(x_i, x_j, x_k) \cdot \lg \frac{P_{X_k}(x_k)P_{X_i, X_j, X_k}(x_i, x_j, x_k)}{P_{X_i, X_k}(x_i, x_k)P_{X_j, X_k}(x_j, x_k)}$$

其中, $P_{X_i, X_j, X_k}(x_i, x_j, x_k)$ 表示节点 X_i, X_j 与 X_k 的联合概率, $P_{X_i, X_k}(x_i, x_k)$ 表示节点 X_i 和 X_k 的联合概率, $P_{X_k}(x_k)$ 表示节点 X_k 的概率。

若 $\Delta > 0$, 表明已知节点 X_k 状态时节点 X_i 和 X_j 之间的依赖关系增强, 即节点 X_i 与 X_j 在 X_k 处形成V-结构^[20], 于是3节点间的无向边定向为 $X_i \rightarrow X_k, X_j \rightarrow X_k$; 否则, 3个节点之间不形成V-结构, 不进行定向, 如式(5)所示。将最大支撑树中所有符合图1所示连接关系的3节点组合都进行V-结构判定并定向, 得到部分定向最大支撑树(Partially Oriented maximum weight spanning Tree, POT)。

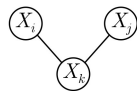


图1 最大支撑树中3节点连接关系

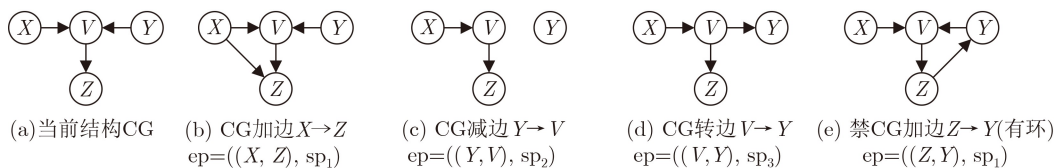


图2 ep修改CG得到候选结构

$$\begin{cases} \{e_{ik} = 1, e_{jk} = 1\}, \Delta > 0 \\ \{e_{ik} = 0, e_{jk} = 0\}, \text{其他} \end{cases} \quad (5)$$

对于POT中仍未定向的任意边 $X_i - X_j$, 将该边定向为 $X_i \rightarrow X_j$ 时得到结构 G_1 , 定向为 $X_i \leftarrow X_j$ 时得到结构 G_2 , 删除 G_1 和 G_2 中无向边并分别代入对数似然函数^[8] $l(G, \theta|D)$ (如式(6)所示)中, 得到对应函数值 l_1 与 l_2 , 当 l_1 小于 l_2 时, $X_i - X_j$ 定向为 $X_i \leftarrow X_j$, 即令 $e_{ij} = -1$; 否则, $X_i - X_j$ 定向为 $X_i \rightarrow X_j$, 即令 $e_{ij} = 1$, 如式(7)所示。将POT中所有为未定向边经过数似然函数定向后, 输出完全定向的最大支撑树结构(Oriented maximum weight spanning Tree, OT), 并将OT作为后续评分搜索的初始结构

$$l(G, \theta|D) = \sum_{i=1}^n \sum_{j=1}^{q_i} \sum_{k=1}^{r_i} m_{ijk} \lg \theta_{ijk} \quad (6)$$

$$\begin{cases} e_{ij} = -1, l_1 < l_2 \\ e_{ij} = 1, \text{其他} \end{cases} \quad (7)$$

其中, n 表示网络中节点总个数, q_i 表示父节点状态组合总数量, r_i 为变量 X_i 的所有取值的总数量, m_{ijk} 为数据集 D 中满足 $X_i = k$ 且父节点状态组合为 j 的样本数量。

将当前结构(Current Structure, CG)和目前最优结构best_G初始化为OT, 进入爬山搜索阶段, 通过边操作对CG进行局部修改(加边、减边、转边)产生候选结构, 边操作ep表达式如式(8)所示, 其中 sp_i 为搜索算子, nodes为 sp_i 修改边对应的两个节点

$$ep = (\text{nodes}, sp_i) \quad (8)$$

$$sp_i = \begin{cases} sp_1, \text{加边} \\ sp_2, \text{减边} \\ sp_3, \text{转边} \end{cases} \quad (9)$$

通过ep对CG修改的具体过程如图2所示, 设CG为图2(a)所示结构, CG加边 $X \rightarrow Z$ 得到的候选结构如图2(b), CG减边 $Y \rightarrow V$ 得到的候选结构如图2(c), CG转边 $V \rightarrow Y$ 得到的候选结构如图2(d), 禁止CG加边 $Z \rightarrow Y$, 因为产生的候选结构中出现非法有向环, 如图2(e)所示。

根据CG节点间的连接状态结合不同的搜索算子, 建立边操作集合 $EP = \{ep_1, ep_2, \dots, ep_w\}$, 利用

不同的边操作修改CG，得到候选结构集如式(10)，其中CG + ep_i表示ep_i对CG进行局部修改

$$G = \{G_i = CG + ep_i | ep_i \in EP\} \quad (10)$$

利用贝叶斯信息准则(Bayesian Information Criterion, BIC)^[21]对CG和G = {G₁, G₂, ..., G_w}中元素进行评分，对应得到评分值S_c和评分值集合S，如式(11)和式(12)所示，以CG的评分作为目前最优评分best_S

$$S_c = BIC(CG|D) \quad (11)$$

$$S = \{S_i | S_i = BIC(G_i|D)\}, G_i \in G \quad (12)$$

其中，BIC(G|D)为贝叶斯信息准则评分函数，如式(13)所示

$$BIC(G|D) = \sum_{i=1}^n \sum_{j=1}^{q_i} \sum_{k=1}^{r_i} m_{ijk} \lg \frac{m_{ijk}}{m_{ij}} - \frac{q_i(r_i - 1)}{2} \lg m \quad (13)$$

其中，n表示网络中节点总个数，q_i表示父节点状态组合总数量，r_i为变量X_i的所有取值的总数量，m为总样本数量，m_{ijk}为数据集D中满足X_i = k且父节点状态组合为j的样本数量，m_{ij} = ∑_{k=1}^{r_i} m_{ijk}。

求S中最大的评分值S_{max}，如式(14)所示，将S_{max}对应的候选结构和边操作分别记为G_{max}和ep_{max}。当S_{max} > best_S时，初始化(清空)禁忌表TL(默认为空)，即令TL = {}，将CG和best_G更新为G_{max}和best_S更新为S_{max}，以更新后的CG为搜索起点再次进行爬山搜索；否则，爬山搜索达到停

止条件，将当前最优结构及其对应的边操作记为G_{temp}和ep_{temp}，进入禁忌搜索阶段，如式(15)所示

$$S_{max} = \max(S) \quad (14)$$

$$\left. \begin{aligned} CG = G_{max}, best_G = G_{max}, best_S = S_{max}, \\ \mathbf{TL} = \{\}, S_{max} > best_S \\ \{G_{temp} = G_{max}, ep_{temp} = ep_{max}\}, \text{其他} \end{aligned} \right\} \quad (15)$$

在禁忌搜索阶段，以ep为禁忌对象。首先检验ep_{max}是否被禁忌，若ep_{max} ∉ TL，表明其不被禁忌，则将当前结构更新为G_{max}，将ep_{max}加入禁忌表TL，如式(16)所示；若ep_{max} ∈ TL，表明其被禁忌，则检测S是否为空集，若S非空，将S_{max}从集合S中删除，将更新后的S代入式(14)求得最大评分值S_{max}，利用S_{max}对应的结构及边操作更新G_{max}和ep_{max}，再检测更新的ep_{max}是否被禁忌；若S为空，将当前结构更新为G_{temp}，将ep_{temp}加入禁忌表TL中，如式(17)所示。

最后判断禁忌表长度(TL中元素的个数)是否达到规定的最大长度ML，如果|TL| ≤ ML(|TL|表示集合TL中元素个数)，回到爬山搜索阶段；否则，输出最优结构best_G

$$\{CG = G_{max}, \mathbf{TL} = \mathbf{TL} \cup \{ep_{max}\}\}, ep_{max} \notin \mathbf{TL} \quad (16)$$

$$\left. \begin{aligned} S = S \setminus \{S_{max}\}, S \neq \{\} \\ CG = G_{temp}, \mathbf{TL} = \mathbf{TL} \cup \{ep_{temp}\}, \text{其他} \end{aligned} \right\} \quad (17)$$

2.2 VTH算法流程图

VTH算法流程如图3所示。

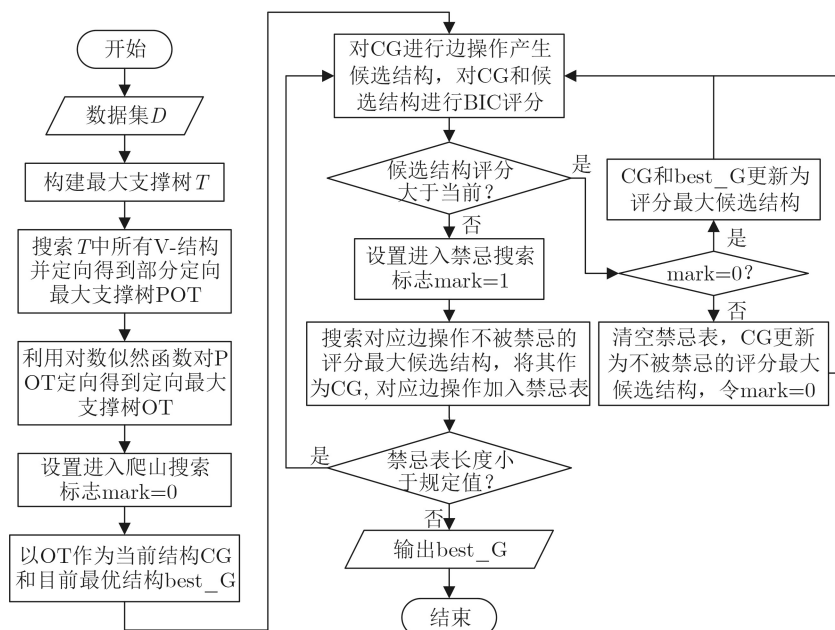


图 3 VTH算法流程图

3 实验结果与分析

实验运行环境为: Intel(R) Core™ i5-8500 CPU, 主频3.00 GHz, 内存8.00 GB, Windows10 64 bit操作系统。在MATLAB 2020a平台中基于bnt-master工具箱进行实验, 本算法(VTH算法)及对比算法(HC算法^[7]、MMHC算法^[9]、Tabu算法^[10]、CHC算法^[11]、FastCHC算法^[12]、SHC算法^[14]和SaiyanH算法^[16]) 在经典的Asia, Car, Child和Alarm标准网络^[22]数据集下进行仿真分析, 对比以下常用评价指标^[23]:

(1)F1值(F1-score), 如式(18)所示, 其值越大表明学习到的网络越接近标准网络, 其中Re(Recall)表示召回率, 计算公式如式(19); Pr(Precision)表示准确率, 计算公式如式(20)

$$F1 = 2 \frac{\text{Pr} \cdot \text{Re}}{\text{Pr} + \text{Re}} \quad (18)$$

$$\text{Re} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (19)$$

$$\text{Pr} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (20)$$

(2)汉明距离(Hamming Distance, HD), 如式(21)所示, HD越小则学习到的网络与标准网络差异越小

$$\text{HD} = \text{FP} + \text{FN} \quad (21)$$

(3)BSF值, 如式(22)所示, 其取值范围为[-1, 1], 值越大表明学习到的网络与标准网络越接近。其中 a 表示标准网络中存在的边数, i 定义为式(23), \mathbf{X} 为标准网络中所有节点集合, $|\mathbf{X}|$ 表示集合 \mathbf{X} 中元素数量

$$\text{BSF} = \frac{1}{2} \left(\frac{\text{TP}}{a} + \frac{\text{TN}}{i} - \frac{\text{FP}}{i} - \frac{\text{FN}}{a} \right) \quad (22)$$

$$i = \frac{1}{2} |\mathbf{X}| (|\mathbf{X}| - 1) \quad (23)$$

(4)运行时间(Time), 算法学习贝叶斯网络结构所需时间。

在式(19)–式(23)中, TP表示学习到的网络结构与标准网络结构中存在的边数, TN表示学习到的网络结构与标准网络结构中都不存在的边数, FP表示学习到的网络结构中存在的边数而标准网络结构中不存在的边数, FN表示学习到的网络结构中不存在而标准网络结构中存在的边数。

3.1 VTH算法实验分析

为检验VLL策略对最大支撑树的定向效果, 选取CRAE策略^[24]进行对比实验, 利用Asia, Car, Child, Alarm 4种标准网络分别生成样本量为1000,

5000, 10000的数据集, 每组数据中独立进行100次试验, 取均值作为最终结果。两种策略定向结果如表1所示, 其中 T 表示定向正确的边数, F 表示定向错误的边数, acc 表示定向准确率, acc 通过式(24)计算

$$\text{acc} = \frac{T}{T + F} \times 100\% \quad (24)$$

由表1可见, 在Asia, Car和Alarm网络实验中, VLL策略相比于CRAE策略正确定向的边数均有提升, 错误定向边数相应减少, 定向准确率分别提升了35.73%, 23.58%, 31.95%。VLL策略将V-结构与对数似然函数相结合, 利用V-结构特殊的连接形式进行前期定向, 利用对数似然函数对边方向的评判性进行后期定向, 最终获得了较优的定向结果。在Child网络中, 由于网络中V-结构较少, VLL策略定向结果稍差, 但相比于CRAE策略定向准确率仅下降1.77%。

为检验THC策略寻优能力, 在Asia, Car, Child和Alarm 4种标准网络中, 选择样本量为5000的数据集, 每种网络都从同一初始结构出发, 禁忌表最大长度设置为 $\text{ML}=15$, 将THC策略与HC策略寻优过程进行对比, 结果如图4所示。

由图4可见, 在前期迭代中THC进行爬山搜索, 所以其曲线与HC的曲线重合, 由于HC策略的停止条件使得其过早停止迭代, 而陷入局部最优, THC在爬山搜索不能使评分进一步增大时, 进行禁忌搜索来试图摆脱局部最优状态, 由于THC策略在禁忌搜索阶段允许接受劣解, 所以在曲线中会有评分下降的过程, 但全局最优结构评分有所增大, 通过对比可以看出, THC策略在前期有与

表1 4种网络中VLL策略和CRAE策略定向结果

网络	数据量	VLL策略			CRAE策略		
		T	F	acc(%)	T	F	acc(%)
Asia	1000	5.56	0.58	90.58	3.58	2.56	58.33
	5000	6.02	0.47	92.81	3.62	2.87	55.82
	10000	6.27	0.47	93.07	3.71	3.02	55.12
Car	1000	7.93	0.96	89.25	5.87	3.02	66.00
	5000	8.18	0.82	90.86	6.00	2.98	66.83
	10000	8.11	0.89	90.12	6.00	3.00	66.67
Child	1000	13.78	4.98	73.46	14.31	4.44	76.30
	5000	13.93	5.07	73.33	14.20	4.80	74.74
	10000	13.98	5.02	73.57	14.18	4.82	74.62
Alarm	1000	33.49	0.44	98.69	23.40	10.53	68.96
	5000	34.02	0	100.00	22.84	11.18	67.15
	10000	34.00	0	100.00	22.69	11.31	66.73

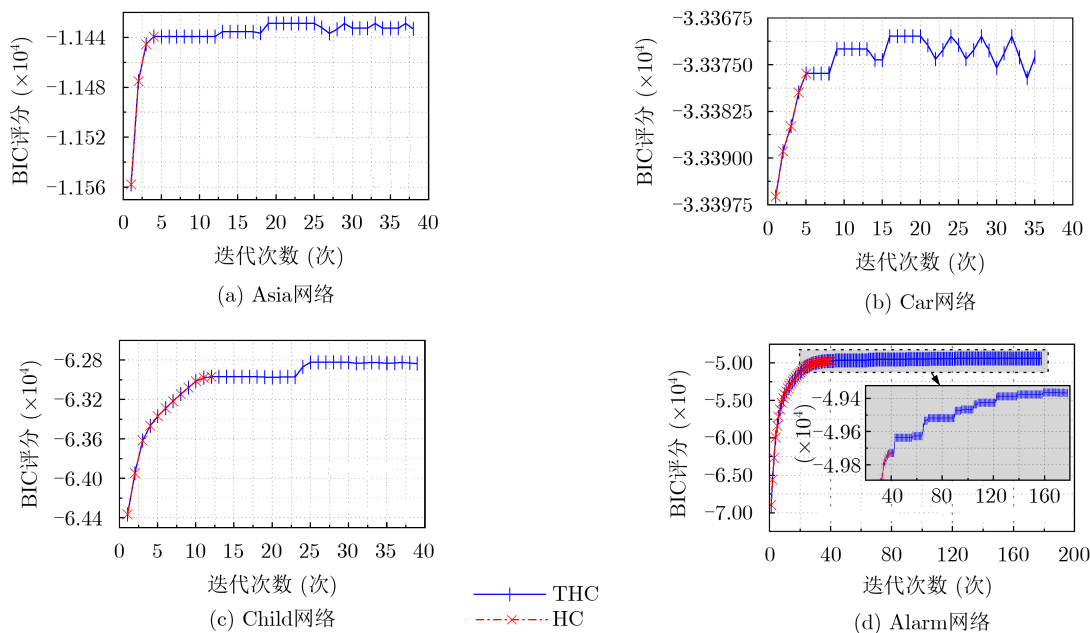


图 4 HC与THC在4种网络中寻优过程对比

HC策略相同的搜索效率，在后期拥有更强的全局寻优能力。此外，在Asia, Car, Child网络的寻优后期都出现最大评分值在较多次迭代中不变的情况，表明目前所设禁忌表长度足够大；由于Alarm网络较大，寻优后期仍有上升趋势，但上升幅度不大，出于时间考虑不再增大禁忌表长度，所以后续试验中ML均设为15。

3.2 VTH与其他算法对比

为分析VTH算法性能，将其分别与HC算法^[6]、MMHC算法^[9]、Tabu算法^[10]、CHC算法^[11]、Fast-CHC算法^[12]、SHC算法^[14]和SaiyanH算法^[16]进行对比，算法介绍如表2所示。

利用Asia, Car, Child和Alarm标准网络分别生成1000, 5000, 10000样本量的数据集，本算法与其他算法在不同网络和数据量的条件下独立实验100次，取均值作为最终结果，4种标准网络中不同数据量下的运行时间实验结果如图5所示，汉明距

离实验结果如图6所示，F1值和BSF值实验结果如表3—表6所示。

由图5看出，在Asia, Car, Child和Alarm网络中，本算法和同样将爬山算法与禁忌搜索结合的SaiyanH算法相比运行时间大大降低，与HC, Tabu算法相比，在Child网络中运行时间有所增加，其余网络中运行时间均有所减少，与CHC, FastCHC算法相比用时基本持平，本算法用时减少主要得益于利用定向最大支撑树约束搜索空间，其次是爬山算法局部择优特点与简化禁忌搜索结合的策略提高了搜索效率。相比于MMHC算法，VTH算法运行时间有所增加。与同样利用最大支撑树限制搜索空间的SHC算法相比，VTH算法时间开销略有增加，是因为SHC按照标准节点顺序指定最大支撑树边方向(即定向与标准网络相同)省去定向时间，且VTH对最大支撑树不完全正确定向会增加后续评分搜索的运行时间。

由图6可以看出，在Asia, Car和Alarm网络中本算法汉明距离小于其他算法，VTH在评分搜索过程中利用简化的禁忌搜索辅助爬山算法寻找全局最优结果，有效改善了爬山算法易陷入局部最优的问题。由于Child网络中VTH对最大支撑树定向结果不理想，导致VTH汉明距离高于按照标准节点顺序(实际中标准节点顺序是未知的)指定最大支撑树边方向的SHC算法。

由表3—表6数据可知，在Asia, Car, Child和Alarm网络实验中，VTH学习到的结果相比于HC, F1值平均提高43.72%, 34.55%, 21.69%, 68.20%，BSF值平均提高50.79%, 27.23%, 19.74%，

表 2 对比算法介绍

算法	描述
HC算法	经典爬山算法
Tabu算法	利用禁忌表提高爬山全局寻优能力的禁忌搜索算法
MMHC算法	依赖分析与爬山算法结合的经典混合算法
CHC算法	通过动态限制提高爬山效率的高效算法
FastCHC算法	对CHC算法的改进算法
SHC算法	标准节点序定向最大支撑树与简化爬山结合的有效算法
SaiyanH算法	新近提出的爬山算法和禁忌搜索结合的混合算法

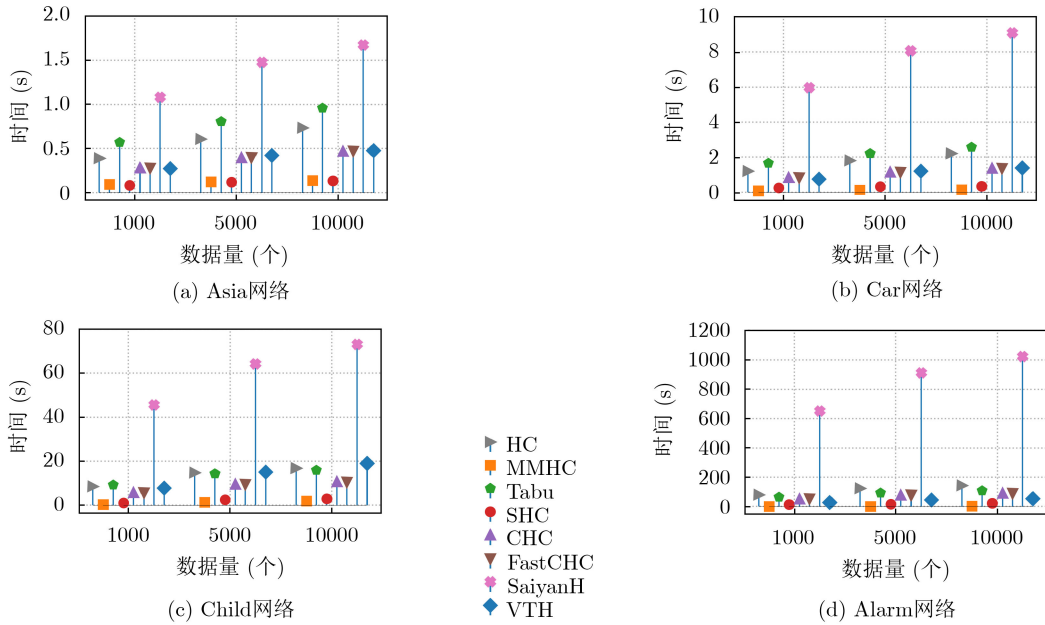


图 5 4个网络中不同算法运行时间结果

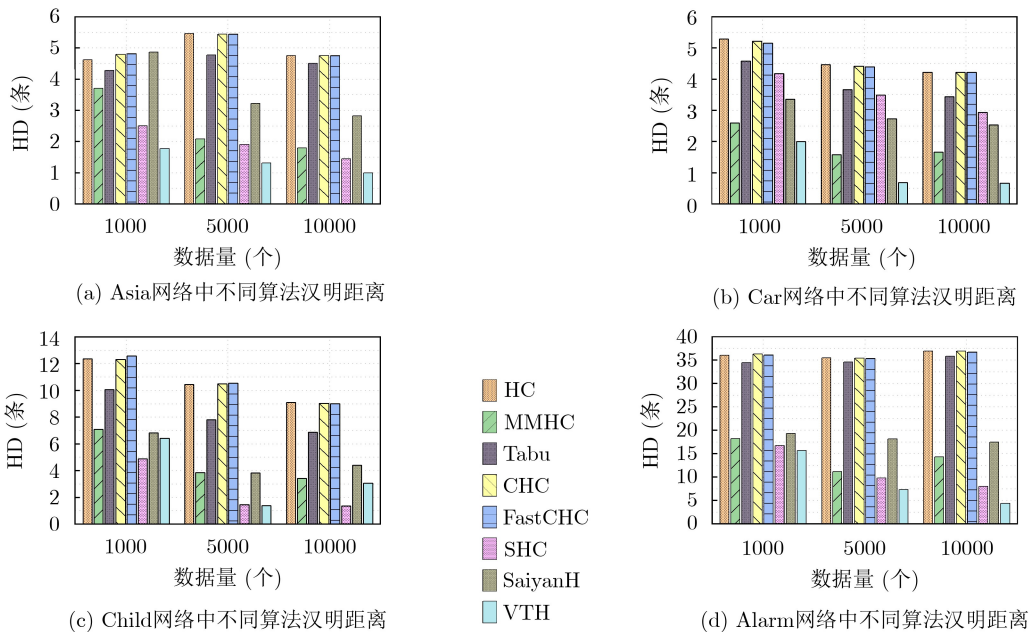


图 6 4个网络中不同算法汉明距离结果

37.55%，相较于MMHC, Tabu, CHC, FastCHC和 SaiyanH, F1值与BSF值也均有不同程度提高；VTH与按照标准节点顺序定向最大支撑树的SHC相比，在Asia网络中F1值与BSF值分别提高4.18%，2.58%，在Car网络中F1值与BSF值分别提高15.45%，7.64%，在Child网络中F1值与BSF值分别降低2.55%，2.27%，在Alarm网络中F1值与BSF值分别提高3.25%，2.64%。VTH在不需要任何先验条件的情况下，在4种网络中学习到的结果要优于多数算法，这是由于VTH在爬山搜索达到停

止条件后，利用禁忌搜索以目前最优解为基础在一定步数内进行可接受劣解的寻优搜索，最终增强了算法的全局寻优能力。

4 结束语

针对爬山算法搜索空间大和评分搜索过程中易陷入局部最优问题，本文提出基于VLL定向与禁忌爬山的贝叶斯网络结构算法，通过VLL策略定向的最大支撑树结构作为后续评分搜索的初始结构，缩小了搜索空间，提高了算法效率；

表 3 Asia网络中不同算法F1值和BSF值结果

数据量	对比项	HC	MMHC	Tabu	SHC	CHC	FastCHC	SaiyanH	VTH
1000	F1	0.6579	0.7450	0.6862	0.8369	0.6424	0.6401	0.6319	0.8788
	BSF	0.4964	0.5696	0.5192	0.6730	0.4841	0.4823	0.4782	0.6950
5000	F1	0.5900	0.8549	0.6481	0.8739	0.5923	0.5923	0.7687	0.9118
	BSF	0.4512	0.6812	0.4966	0.7137	0.4530	0.4530	0.5990	0.7325
10000	F1	0.6526	0.8748	0.6721	0.9046	0.6526	0.6526	0.7974	0.9335
	BSF	0.5069	0.7052	0.5226	0.7468	0.5069	0.5069	0.6300	0.7605

表 4 Car网络中不同算法F1值和BSF值结果

数据量	对比项	HC	MMHC	Tabu	SHC	CHC	FastCHC	SaiyanH	VTH
1000	F1	0.6383	0.8317	0.6962	0.7690	0.6456	0.6495	0.7852	0.8810
	BSF	0.6357	0.7589	0.6742	0.7686	0.6410	0.6430	0.7380	0.8203
5000	F1	0.7133	0.9042	0.7673	0.8133	0.7161	0.7179	0.8299	0.9609
	BSF	0.6989	0.8333	0.7345	0.8118	0.7006	0.7020	0.7785	0.8958
10000	F1	0.7346	0.8982	0.7844	0.8462	0.7346	0.7346	0.8443	0.9617
	BSF	0.7164	0.8277	0.7485	0.8426	0.7164	0.7164	0.7911	0.8919

表 5 Child网络中不同算法F1值和BSF值结果

数据量	对比项	HC	MMHC	Tabu	SHC	CHC	FastCHC	SaiyanH	VTH
1000	F1	0.7054	0.8367	0.7644	0.8930	0.7068	0.6999	0.8467	0.8554
	BSF	0.6193	0.7250	0.6649	0.7736	0.6209	0.6152	0.7308	0.7386
5000	F1	0.7666	0.9167	0.8291	0.9698	0.7654	0.7636	0.9187	0.9715
	BSF	0.7273	0.8422	0.7776	0.9014	0.7273	0.7251	0.8526	0.9061
10000	F1	0.8000	0.9267	0.8514	0.9717	0.8016	0.8021	0.9061	0.9366
	BSF	0.7607	0.8569	0.8006	0.9030	0.7621	0.7626	0.8466	0.8775

表 6 Alarm网络中不同算法F1值和BSF值结果

数据量	对比项	HC	MMHC	Tabu	SHC	CHC	FastCHC	SaiyanH	VTH
1000	F1	0.5128	0.7767	0.5369	0.7951	0.5068	0.5095	0.7616	0.8090
	BSF	0.5680	0.7644	0.5807	0.7457	0.5642	0.5658	0.7413	0.7545
5000	F1	0.5463	0.8690	0.5589	0.8866	0.5485	0.5485	0.7891	0.9165
	BSF	0.6413	0.8474	0.6485	0.8488	0.6425	0.6421	0.7955	0.8755
10000	F1	0.5313	0.8268	0.5484	0.9093	0.5308	0.5329	0.7972	0.9514
	BSF	0.6402	0.8183	0.6505	0.8856	0.6399	0.6381	0.8072	0.9174

评分搜索策略中将禁忌表清空机制与爬山算法的局部择优准则结合,使得算法在评分搜索过程的前期具有较高的搜索效率,搜索过程的后期具有较强的跨越坪区和摆脱局部最优的能力。仿真实验表明,VTH算法在多数标准网络的实验中,能够在保证效率的同时获得到较优结构,但在最大支撑树定向不佳的网络中,仍然会有运行时间上升和陷入局部最优的问题,后续可用不同方式对初始结构进一步优化,以提高算法效率和准确性。

参考文献

- [1] 李硕豪,张军. 贝叶斯网络结构学习综述[J]. 计算机应用研究, 2015, 32(3): 641-646. doi: [10.3969/j.issn.1001-3695.2015.03.001](https://doi.org/10.3969/j.issn.1001-3695.2015.03.001).
LI Shuohao and ZHANG Jun. Review of Bayesian networks structure learning[J]. *Application Research of Computers*, 2015, 32(3): 641-646. doi: [10.3969/j.issn.1001-3695.2015.03.001](https://doi.org/10.3969/j.issn.1001-3695.2015.03.001).
- [2] ZHAO Xuan, PENG Benhong, ELAHI E, et al. Optimization of Chinese coal-fired power plants for cleaner

- production using Bayesian network[J]. *Journal of Cleaner Production*, 2020, 273: 122837. doi: [10.1016/j.jclepro.2020.122837](https://doi.org/10.1016/j.jclepro.2020.122837).
- [3] MCLACHLAN S, DUBE K, HITMAN G A, *et al.* Bayesian networks in healthcare: Distribution by medical condition[J]. *Artificial Intelligence in Medicine*, 2020, 107: 101912. doi: [10.1016/j.artmed.2020.101912](https://doi.org/10.1016/j.artmed.2020.101912).
- [4] CARRIGER J F and BARRON M G. A Bayesian network approach to refining ecological risk assessments: Mercury and the Florida panther (*Puma concolor coryi*)[J]. *Ecological Modelling*, 2020, 418: 108911. doi: [10.1016/j.ecolmodel.2019.108911](https://doi.org/10.1016/j.ecolmodel.2019.108911).
- [5] ANDERSON B. Using Bayesian networks to perform reject inference[J]. *Expert Systems with Applications*, 2019, 137: 349–356. doi: [10.1016/j.eswa.2019.07.011](https://doi.org/10.1016/j.eswa.2019.07.011).
- [6] SCANAGATTA M, SALMERÓN A, and STELLA F. A survey on Bayesian network structure learning from data[J]. *Progress in Artificial Intelligence*, 2019, 8(4): 425–439. doi: [10.1007/s13748-019-00194-y](https://doi.org/10.1007/s13748-019-00194-y).
- [7] HECKERMAN D, GEIGER D, and CHICKERING D M. Learning Bayesian networks: The combination of knowledge and statistical data[J]. *Machine Learning*, 1995, 20(3): 197–243. doi: [10.1007/BF00994016](https://doi.org/10.1007/BF00994016).
- [8] 曹杰. 贝叶斯网络结构学习与应用研究[D].[博士论文], 中国科学技术大学, 2017.
CAO Jie. Bayesian network structure learning and application[D].[Ph.D. dissertation], University of Science and Technology of China, 2017.
- [9] TSAMARDINOS I, BROWN L E, and ALIFERIS C F. The max-min hill-climbing Bayesian network structure learning algorithm[J]. *Machine Learning*, 2006, 65(1): 31–78. doi: [10.1007/s10994-006-6889-7](https://doi.org/10.1007/s10994-006-6889-7).
- [10] 冀俊忠, 张鸿勋, 胡仁兵, 等. 基于禁忌搜索的贝叶斯网结构学习算法[J]. 北京工业大学学报, 2011, 37(8): 1274–1280.
JI Junzhong, ZHANG Hongxun, HU Renbing, *et al.* A Tabu-search based Bayesian network structure learning algorithm[J]. *Journal of Beijing University of Technology*, 2011, 37(8): 1274–1280.
- [11] GÁMEZ J A, MATEO J L, and PUERTA J M. Learning Bayesian networks by hill climbing: Efficient methods based on progressive restriction of the neighborhood[J]. *Data Mining and Knowledge Discovery*, 2011, 22(1-2): 106–148. doi: [10.1007/s10618-010-0178-6](https://doi.org/10.1007/s10618-010-0178-6).
- [12] GÁMEZ J A, MATEO J L, and PUERTA J M. One iteration CHC algorithm for learning Bayesian networks: An effective and efficient algorithm for high dimensional problems[J]. *Progress in Artificial Intelligence*, 2012, 1(4): 329–346. doi: [10.1007/s13748-012-0033-7](https://doi.org/10.1007/s13748-012-0033-7).
- [13] ARIAS J, GÁMEZ J A, and PUERTA J M. Structural learning of Bayesian networks via constrained hill climbing algorithms: Adjusting trade-off between efficiency and accuracy[J]. *International Journal of Intelligent Systems*, 2015, 30(3): 292–325. doi: [10.1002/int.21701](https://doi.org/10.1002/int.21701).
- [14] 刘浩然, 吕晓贺, 李轩, 等. 基于Bayesian改进算法的回转窑故障诊断模型研究[J]. 仪器仪表学报, 2015, 36(7): 1554–1561. doi: [10.3969/j.issn.0254-3087.2015.07.014](https://doi.org/10.3969/j.issn.0254-3087.2015.07.014).
LIU Haoran, LÜ Xiaohu, LI Xuan, *et al.* A study on the fault diagnosis model of rotary kiln based on an improved algorithm of Bayesian[J]. *Chinese Journal of Scientific Instrument*, 2015, 36(7): 1554–1561. doi: [10.3969/j.issn.0254-3087.2015.07.014](https://doi.org/10.3969/j.issn.0254-3087.2015.07.014).
- [15] 刘彬, 刘永记, 刘浩然, 等. 基于改进遗传爬山算法的篦冷机熟料换热二次风温故障诊断[J]. 计量学报, 2018, 39(5): 651–658. doi: [10.3969/j.issn.1000-1158.2018.05.10](https://doi.org/10.3969/j.issn.1000-1158.2018.05.10).
LIU Bin, LIU Yongji, LIU Haoran, *et al.* Fault diagnosis of secondary air temperature of grate cooler cement clinker heat transfer based on improved genetic hill climbing algorithm[J]. *Acta Metrologica Sinica*, 2018, 39(5): 651–658. doi: [10.3969/j.issn.1000-1158.2018.05.10](https://doi.org/10.3969/j.issn.1000-1158.2018.05.10).
- [16] CONSTANTINOU A C. Learning Bayesian networks that enable full propagation of evidence[J]. *IEEE Access*, 2020, 8: 124845–124856. doi: [10.1109/ACCESS.2020.3006472](https://doi.org/10.1109/ACCESS.2020.3006472).
- [17] CHEN Guangjie and GE Zhiqiang. Robust Bayesian networks for low-quality data modeling and process monitoring applications[J]. *Control Engineering Practice*, 2020, 97: 104344. doi: [10.1016/j.conengprac.2020.104344](https://doi.org/10.1016/j.conengprac.2020.104344).
- [18] YANG Jing, GUO Xiaoxue, AN Ning, *et al.* Streaming feature-based causal structure learning algorithm with symmetrical uncertainty[J]. *Information Sciences*, 2018, 467: 708–724. doi: [10.1016/j.ins.2018.04.076](https://doi.org/10.1016/j.ins.2018.04.076).
- [19] PRIM R C. Shortest connection networks and some generalizations[J]. *The Bell System Technical Journal*, 1957, 36(6): 1389–1401. doi: [10.1002/j.1538-7305.1957.tb01515.x](https://doi.org/10.1002/j.1538-7305.1957.tb01515.x).
- [20] DAI Jingguo, REN Jia, and DU Wencai. Decomposition-based Bayesian network structure learning algorithm using local topology information[J]. *Knowledge-Based Systems*, 2020, 195: 105602. doi: [10.1016/j.knosys.2020.105602](https://doi.org/10.1016/j.knosys.2020.105602).
- [21] BEHJATI S and BEIGY H. Improved K2 algorithm for Bayesian network structure learning[J]. *Engineering Applications of Artificial Intelligence*, 2020, 91: 103617. doi: [10.1016/j.engappai.2020.103617](https://doi.org/10.1016/j.engappai.2020.103617).
- [22] ZHAO Jianjun and HO S S. Improving Bayesian network local structure learning via data-driven symmetry correction methods[J]. *International Journal of Approximate Reasoning*, 2019, 107: 101–121. doi: [10.1016/j.ijar.2019.02.004](https://doi.org/10.1016/j.ijar.2019.02.004).
- [23] CONSTANTINOU A C. The importance of temporal information in Bayesian network structure learning[J]. *Expert Systems with Applications*, 2021, 164: 113814. doi: [10.1016/j.eswa.2021.113814](https://doi.org/10.1016/j.eswa.2021.113814).

[10.1016/j.eswa.2020.113814](https://doi.org/10.1016/j.eswa.2020.113814).

- [24] JIANG Jinke, WANG Juyun, YU Hua, *et al.* Poison identification based on Bayesian network: A novel improvement on K2 algorithm via Markov Blanket[C]. The 4th International Conference on Advances in Swarm Intelligence, Harbin, China, 2013: 173–182.

刘浩然：男，1980年生，教授，博士生导师，研究方向为无线传感器网络、工业故障检测及预测。

王念太：男，1996年生，硕士生，研究方向为贝叶斯网络、工业故障检测及预测。

王毅：男，1996年生，硕士生，研究方向为贝叶斯网络、工业故障检测及预测。

张力悦：男，1994年生，博士生，研究方向为贝叶斯网络、工业故障检测及预测。

苏昭玉：女，1994年生，硕士生，研究方向为贝叶斯网络、工业故障检测及预测。

刘文：女，1966年生，学士，研究方向为机床工艺参数感知与预测。

赵旭丹：女，1981年生，硕士，研究方向为财务会计理论与方法。

责任编辑：余蓉