

基于蚁群优化的贝叶斯网络学习

高晓光, 赵欢欢, 任 佳

(西北工业大学电子信息学院, 陕西 西安 710072)

摘要: 针对贝叶斯网络学习中的混合算法容易缩小搜索空间, 同时易陷入局部最优等缺点, 提出了基于蚁群优化的贝叶斯网络学习算法。首先应用最大最小父子节点集合算法(max-min parents and children, MMPC)来构建无向网络的框架, 然后利用蚁群优化算法进行评分-搜索, 通过平衡“开发”和“探索”力度来修补搜索空间并确定网络结构中边的方向。最后应用本算法学习逻辑报警还原机理网(a logical alarm reduction mechanism, ALARM), 结果显示本算法减少了丢失边的数量, 得到了更接近真实结构的贝叶斯网络。

关键词: 贝叶斯网络; 结构学习; 蚁群优化算法

中图分类号: TP 181

文献标志码: A

DOI: 10.3969/j.issn.1001-506X.2010.07.037

Bayesian network learning on algorithm based on ant colony optimization

GAO Xiao-guang, ZHAO Huan-huan, REN Jia

(School of Electronics and Information, Northwestern Polytechnical Univ., Xi'an 710072, China)

Abstract: According to the hybrid Bayesian networks learning algorithms which are easy to narrow the search space and fall into local optimum, a Bayesian network learning algorithm based on ant colony optimization is proposed. Firstly, this paper applies max-min parents and children (MMPC) to construct the framework of the undirected network, and then uses ant colony optimization to score-search, by balancing the “exploitation” and “exploration” to repair the search space and determine the direction of edges in the network. Finally applying the algorithm to learn a logical alarm reduction mechanism (ALARM) network shows that it reduces the number of missing edges, and gets closer to the real structure of Bayesian network.

Keywords: Bayesian networks; structure learning; ant colony optimization

0 引言

贝叶斯网络是一种帮助人们将概率统计应用于复杂领域进行不确定性推理和数据分析的工具。构造贝叶斯网络的主要目的是进行概率推理。但是如何构建一个合理的贝叶斯网络仍存在很多问题, 因此, 作为通过专家确定网络结构的方法的补充, 发展一种从数据中自动构建网络结构的方法仍然具有重要的研究意义。

目前, 贝叶斯网络结构学习算法大多基于两种观点: (1) 基于条件独立性检验的方法^[1-2], 以 d-separation 准则为依据, 通过条件独立性检验确定两个结点之间边的存在性, 并利用碰撞识别等方法确定边的方向; (2) 基于评分-搜索的方法^[3-5], 这类方法学习过程简单规范, 但打分函数的运算复杂程度和结构搜索空间大小随变量增加指数增长(完全搜索是 N-P 难问题)。此外, 一些研究者结合了上述两种方法, 提出了一些混合算法^[6-8], 这类算法首先利用独立性测试降低搜索空间的复杂度, 然后执行评分搜索找到最佳

网络, 如最大最小爬山(max-min hill-climbing, MMHC)算法。但是, 该类算法在降低搜索空间的复杂度的同时缩小了搜索空间, 而评分搜索算法由于搜索空间的约束而不会修正这个错误, 又容易陷入局部最优, 从而将导致无法学习到最佳的网络结构。

针对上述文献存在的问题, 本文提出了基于蚁群优化的贝叶斯网络学习算法, 首先利用最大最小父子节点集合算法(max-min parent and child, MMPC)构建无向网络的框架, 然后利用蚁群优化算法进行评分-搜索, 通过平衡“开发”(确定候选父子节点集中的边的方向)和“探索”以一定的概率将父子节点集中的节点的父子节点加入搜索空间), 来修补搜索空间, 在构建解的同时进行局部更新, 提高了结构学习的质量。

1 贝叶斯网络结构学习

贝叶斯网络^[9]表示为一个有向无环图 $B=(V, E)$, 其中节点集合 $V=\{x_1, x_2, \dots, x_n\}$ 表示变量, 有向边集合 E

收稿日期: 2009-07-07; 修回日期: 2009-11-24。

基金项目: 国家自然科学基金(60774064)资助课题

作者简介: 高晓光(1957-), 女, 教授, 博士, 主要研究方向为航空火力控制、智能决策与优化。E-mail: xggao@nwpu.edu.cn

表示变量间的直接依赖关系,条件概率分布存储在变量节点中。对于每一个节点 $x_i \in V$ 都有一个条件分布族 $P(x_i | Pa(x_i))$,其中 $Pa(x_i)$ 表示变量 x_i 的父节点集合。由此,我们将联合概率分布分解为

$$P(x_1, x_2, \dots, x_n) = \prod_{i=1}^n P(x_i | Pa(x_i)) \quad (1)$$

依赖与独立关系使得上式的分解成为可能,并且通过变量间的直接相连与否在网络结构中体现。

贝叶斯网络结构学习的问题表述如下:设 x_1, x_2, \dots, x_n 是一组随机变量, $D = (D_1, D_2, \dots, D_n)$ 是关于这些随机变量的一组独立同分布的数据,寻找在某种意义上与数据匹配最相近的结构模型。首先定义一个关于网络结构的评分函数,然后再寻找全局最优解。本文采取通用的贝叶斯评分函数

$$L(B | D) = \prod_{i=1}^n \prod_{j=1}^{q_i} \frac{\Gamma(\alpha_{ij*})}{\Gamma(\alpha_{ij*} + m_{ij*})} \prod_{k=1}^{r_i} \frac{\Gamma(\alpha_{ijk} + m_{ijk})}{\Gamma(\alpha_{ijk})} \quad (2)$$

式中, m_{ijk} 是 D 中满足 $x_i = k, Pa(x_i) = j$ 的样本个数,

$m_{ij*} = \sum_{k=1}^{r_i} m_{ijk}, \alpha_{ij*} = \sum_{k=1}^{r_i} \alpha_{ijk}$, 取一个等价样本,其数量为 N ,

$$\alpha_{ijk} = \frac{N}{q_i r_i}.$$

对上式取对数得到评分函数的分解式

$$f(B | D) = \sum_{i=1}^n \sum_{j=1}^{q_i} \left[\lg \frac{\Gamma(\alpha_{ij*})}{\Gamma(\alpha_{ij*} + m_{ij*})} + \sum_{k=1}^{r_i} \frac{\Gamma(\alpha_{ijk} + m_{ijk})}{\Gamma(\alpha_{ijk})} \right] = \sum_{i=1}^n \left\{ \sum_{j=1}^{q_i} \left[\lg \frac{\Gamma(\alpha_{ij*})}{\Gamma(\alpha_{ij*} + m_{ij*})} + \sum_{k=1}^{r_i} \frac{\Gamma(\alpha_{ijk} + m_{ijk})}{\Gamma(\alpha_{ijk})} \right] \right\} = \sum_{i=1}^n f(\{x_i, Pa(x_i)\} | D) \quad (3)$$

式中, $f(\{x_i, Pa(x_i)\} | D)$ 为家族评分函数, $\{x_i, Pa(x_i)\}$ 记为变量 x_i , 其父节点集 $Pa(x_i)$ 以及它们间的边形成的局部结构。

2 基于蚁群优化的结构学习算法

本文提出的贝叶斯网络学习算法,结合了局部学习、条件独立检验、搜索-评分等方法,首先,应用基于限制技术的 MMPC 来构建无向网络的框架,然后,执行基于蚁群优化算法的评分-搜索来修补搜索空间并确定网络结构的边及其方向。

2.1 MMPC 算法

MMPC 算法是一种局部因果发现算法,用来评估变量 x_i 的父子节点集合。给定数据集 D , 目标变量 T , MMPC 算法得到每个节点的候选父子节点集(candidate parents and children, CPC)。此算法的第一步利用最大-最小策略启发式使变量依次进入目标节点 T 的候选父子节点集 $CPC(T)$, 在第二步中移去 CPC 中第一步进入的本不该有的变量。算法具体描述^[8]如下:

算法 MMPC

输入:目标变量 T ,数据集 D

输出: T 的父子节点集合

- 1 CPC 为空
- 2 repeat
- 3 $\{F, assocF\} = \text{MaxMinHeuristic}(T, CPC)$
- 4 if $assocF \neq 0$
 - then $CPC = CPC \cup F$
- 5 end if
- 6 until CPC 不变
- 7 for all $x \in CPC$
- 8 do
 - if $\exists S \in CPC. s. t. \text{Ind}(X; T | S)$
 - then $CPC = CPC \setminus \{X\}$
- 9 end for
- 10 return CPC

子程序 MaxMinHeristic(T, CPC)

输入:目标变量 T, CPC 子集

输出:以 CPC 为条件集,与 T 最小依赖的最大变量

- 1 $assocF = \max_{X \in V} \min \text{Assoc}(X; Y | CPC)$
- 2 $F = \arg \max_{X \in V} \min \text{Assoc}(X; Y | CPC)$
- 3 Return $\{F, assocF\}$

然而,在第二阶段中移去变量时,由于独立统计信息量的取值不同,会造成候选集合不符合下一部分搜索的准确度要求。为使最终获得的贝叶斯网络尽量接近真实的结构,一般对独立性测试的置信度取值接近 1,来减少本不该有的变量,但同时就造成搜索空间的丢失。

2.2 基于蚁群优化算法的搜索

蚁群优化^[10-11]是模拟蚁群觅食行为的元启发式算法。由大量蚂蚁组成的集体觅食行为就表现出一种信息正反馈现象:某一路径越短,该路径上走过的蚂蚁就越多,所留下的信息素强度也就越大,后来者选择该路径的概率因此就越大。蚂蚁个体之间通过信息素交流来选择最短路径并达到搜索食物的目的。因此,蚁群优化算法被用于旅行商等一系列路径寻优问题,为了将其应用到贝叶斯网络结构学习中,本文对蚁群优化算法做了如下改进:

问题表述:本算法中的 MMPC 部分给出了结构框架,即无向图中的候选边集合,本文用集合来表示,用矩阵的形式表示图 G 。

约束:算法的唯一约束要求最后生成的图(贝叶斯网络)必须是无环图。

信息素:每一条有向边 E_{ij} 关联着一个信息素 τ_{ij} 。在蚂蚁移动过程中,每确定一条边的方向,都会立刻调用局部更新规则更新该边上的信息素含量

$$\begin{cases} \tau_{ij} \leftarrow (1 - \varepsilon) \tau_{ij} + \varepsilon \tau_0 \\ \tau_{ji} \leftarrow (1 - \varepsilon) \tau_{ji} \end{cases} \quad (4)$$

式中, ϵ 满足 $0 < \epsilon < 1$; τ_0 是信息素初始值。只有最优的蚂蚁(构建了迄今为止评分函数值最大的结构 G^+)被允许在每次迭代之后采用全局更新规则

$$\tau_{ij} \leftarrow (1 - \rho)\tau_{ij} + \rho\Delta\tau_{ij} \quad (5)$$

式中, ρ 满足

$$0 < \rho < 1, \Delta\tau_{ij} = \frac{1}{|f(G^+; D)|}$$

启发式信息: 有向边 E_{ij} 加入到部分解 x 后对评分函数的改进值, 定义如下

$$\eta_{ij} = f(x \cup E_{ij}) - f(x) \quad (6)$$

转移概率: 位于节点 i 的蚂蚁 k , 根据随机比例转移规则确定与节点 j 之间边的方向。转移规则如下

$$p = \frac{[\tau_{ij}]^\alpha [\eta_{ij}]^\beta}{[\tau_{ij}]^\alpha [\eta_{ij}]^\beta + [\tau_{ji}]^\alpha [\eta_{ji}]^\beta} \quad (7)$$

当边 (i, j) 不在 CPC 中时, 取 $\tau_{ij} = \tau_0$ 。

修补搜索空间: 位于节点 i 的蚂蚁 k , 以概率 q 将边 (i, j) (节点 j 属于节点 i 的父子节点集中的节点的父子节点集合)临时加入候选边集合, 其中 q 是均匀分布在 $[0, 1]$ 中的随机变量, 其中 $0 \leq q \leq q_0 \leq 1$, 调整 q_0 来加强“探索”的力度。只有在全局信息素更新时, 将最优蚂蚁修补的边最终加入到候选边集合, 并用于下次的迭代。

局部搜索: 由于蚁群优化算法的结构构建步骤使用了和局部搜索截然不同的搜索领域, 可以用于局部地优化蚂蚁构建的解, 本文采用爬山算法进行局部搜索。

算法描述如下:

算法 ACO-B

输入: CPC

输出: 贝叶斯网络结构 G_b

- 1 令 G^+, G_b 为空图
- 2 对每一个候选边, 令 $\tau_{ij} = \tau_0$
- 3 for 1 to T_{\max} (最大迭代次数)
- 4 随机放置 m 个蚂蚁到各个节点
- 5 for 1 to m
- ANT
- end for
- 6 $G^+ = \text{hillclimbing}(G_b)$
- 7 全局更新信息素, 用式(5)
- 8 将最优蚂蚁的修补边加入候选集合
- 9 if $f(G_b; D) \leq f(G^+; D)$
- then $G_b \leftarrow G^+$
- 10 end for
- 11 return G_b

子程序算法 ANT

输入: CPC

输出: 贝叶斯网络结构

- 1 令 G_k 为空图
- 2 while(候选边集合 S 不为空)
- 3 以概率 q 将边 (i, j) 加入 S
- 4 do 随机选一条与所在节点相连的边
- 5 蚂蚁 k 以概率 p 确定边的方向
- 6 计算启发式信息, 用式(6)
- 7 检查是否有环
- 8 信息素局部更新, 用式(4)
- 9 更新 G_k
- 10 候选边集合 $S \leftarrow S \setminus \{S_i\}$
- 11 end while
- 12 return G_k

3 仿真评估

为了检验本文提出的基于蚁群优化的贝叶斯网络结构学习算法, 我们选择评价结构学习算法的基准——逻辑报警还原机理网 (a logical alarm reduction mechanism, ALARM) 网络进行结构学习, 并与 MMHC 算法来比较。其中, ALARM 网络(如图 1 所示)具有 37 个节点和 46 条有向边。MMHC 算法是典型的混合学习算法, 大量实验表明 MMHC 算法与文献[1-7]相比, 在网络学习质量和学习效率上更优。在实验中, 我们采用样本容量分别为 500, 1 000, 3 000, 5 000, 10 000 的数据集合。

为评价学习得到的网络结构的质量, 本文采用较直观的汉明距离^[8]作为标准。汉明距离定义为 $H(G) = M(G) + A(G) + I(G)$, 即与原始网络相比丢失边 $M(G)$ 、多余边 $A(G)$ 和逆向边 $I(G)$ 的数量。通常一个好的搜索算法可以从数据集学习到的网络尽可能的与原始网络相似, 即丢失的边, 多余的边以及方向相反的边的数量最低。

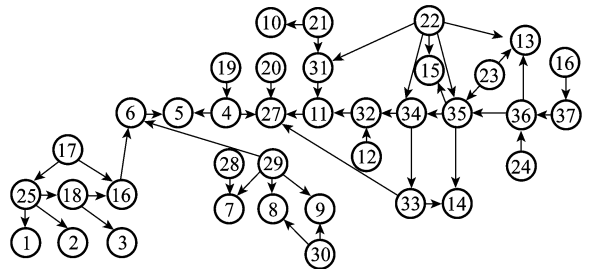


图1 标准 ALARM 网络

本实验中参数的设置如下: 在 MMPC 阶段独立性测试的置信度 $\alpha = 0.95$, ACO 阶段 $\epsilon = \rho = 0.1$, $\alpha = 1$, $\beta = 4$, $m = 30$, $T_{\max} = 100$, $q_0 = 0.1$ 。仿真结果如图 2 所示。

图 2 显示随着样本容量的增大, 汉明距离越来越小, 越接近真实的网络结构, 在样本容量为 1 000, 10 000 时, 本算法和 MMPC 算法取得的汉明距离相同, 分别为 9、3, 在其他样本容量的情况下, 距离都小于 MMPC 算法取得的结果。

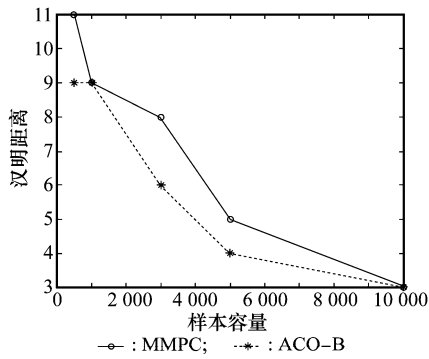


图 2 ALARM 网络学习结果

图 3 显示本算法丢失边的数量普遍少于 MMPC 算法。结果表明本算法在一定程度上修补了搜索空间,对结构学习确实取得了优良的效果。

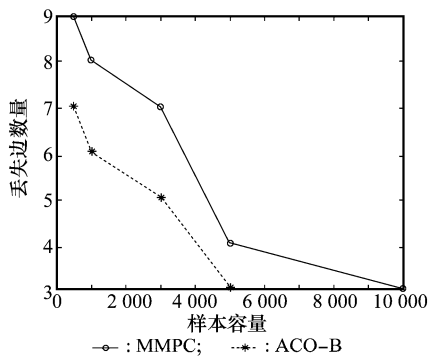


图 3 ALARM 网络学习修补结果

4 结 论

本文提出的混合算法,结合了评分-搜索与条件独立检验等方法的优点,首先利用条件独立检验构建网络的整体框架,然后利用蚁群优化算法进行搜索,很好地弥补了在 MMPC 阶段搜索空间的遗漏。在实际情况中,往往因为成

本等问题无法得到大量的样本数据,而仿真实验表明在小样本容量的情况下,本算法比之 MMHC 算法取得了更好的效果,在实际应用中会更具有竞争优势。

参考文献:

- [1] Campos L M. Independency relationships and learning algorithms for singly connected networks[J]. *Journal of Experimental and Theoretical Artificial Intelligence*, 1998,10(4):511 - 549.
- [2] Campos L M, Huete J F. A new approach for learning Bayesian networks using independence criteria[J]. *International Journal of Approximate Reasoning*, 2000,24(1):11 - 37.
- [3] Lam W, Bacchus F. Learning Bayesian networks: an approach based on the MDL principle[J]. *Computational Intelligence*, 1994,10(4):269 - 293.
- [4] Heckerman D, Geiger D, Chickering D M. Learning Bayesian networks: the combination of knowledge and statistical data[J]. *Machine Learning*, 1995,20(1):197 - 244.
- [5] Campos L M, Fernandez L J, Gamez J, et al. Ant colony optimization for learning Bayesian networks[J]. *International Journal of Approximate Reasoning*, 2002,31(21):291 - 311.
- [6] Acid S, Campos L M. A hybrid methodology for learning Bayesian networks: benedict[J]. *International Journal of Approximate Reasoning*, 2001,27(3):235 - 262.
- [7] Dash D, Druzdel M. A hybrid anytime algorithm for the construction of causal models from sparse data[C]// *Proc. of the 15th Conference on Uncertainty in Artificial Intelligence*, 1999:142 - 149.
- [8] Tsamardinos I, Brown L F, Aliferis C F. The max-min hill-climbing Bayesian network structure learning algorithm [J]. *Machine Learning*, 2006,65(1):31 - 78.
- [9] 张连文,郭海鹏. 贝叶斯网引论[M]. 北京:科学出版社,2006.
- [10] Marco D, Thomas S. *Ant colony optimization*[M]. The MIT Press, 2004.
- [11] Dorigo M, Maniezzo V, Colorni A. The ant system: optimization by a colony of cooperating agents[C]// *IEEE Trans. on Systems, Man and Cybernetics*, 1996, B(26):29 - 41.