

小数据集 BN 建模方法及其在威胁评估中的应用

邱若海,高晓光,郭志高

(西北工业大学电子信息学院,陕西西安 710129)

摘要: 贝叶斯网络是数据挖掘领域的主要工具之一. 在某些特定场合,如重大装备的故障诊断、地质灾害预测及作战决策等,希望用少量数据得到较好的结果. 因此,本文针对小数据集条件下的贝叶斯网络学习问题展开研究. 首先,建立基于连接概率分布的结构约束模型,提出 I-BD-BPSO(Improved-Bayesian Dirichlet-Binary Particle Swarm Optimization) 结构学习算法;其次,建立单调性参数约束模型,提出 MCE(Monotonicity Constraint Estimation) 参数学习算法;最后,应用所提算法构建威胁评估模型并应用变量消元法进行推理计算. 实验结果表明,在小数据集条件下,本文的结构学习算法优于经典的二值粒子群优化算法,参数学习算法优于最大似然估计、保序回归及凸优化算法,并能够构建有效的威胁评估模型.

关键词: 贝叶斯网络; 小数据集; 二值粒子群优化; 威胁评估

中图分类号: TP18 **文献标识码:** A **文章编号:** 0372-2112 (2016)06-1504-08

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2016.06.035

The Modeling Method with Bayesian Networks and Its Application in the Threat Assessment Under Small Data Sets

DI Ruo-hai, GAO Xiao-guang, GUO Zhi-gao

(School of Electronic and Information, Northwestern Polytechnical University, Xi'an, Shaanxi 710129, China)

Abstract: Bayesian network is one of the main tools for data mining. In such cases as large equipment fault diagnosis, geological disaster forecast, operational decision, etc, good results are expected to achieve based on small data sets. Therefore, this article focuses on the problem of learning Bayesian network from small data sets. Firstly, the structure constraint model based on the probability distribution of the connection was built. Then, the improved-Bayesian Dirichlet-binary particle swarm optimization algorithm was proposed. Secondly, the monotonicity parameter constraint model was defined and the monotonicity constraint estimation algorithm was proposed. Finally, the proposed algorithm was applied to construct the threat assessment model. Then, the model was used for reasoning with the variable elimination method. Experimental results reveal that the structure learning algorithm outperforms classical binary particle swarm optimization algorithm and the parameter learning method surpasses maximum likelihood estimation, isotonic regression and convex optimization method for small data sets. The threat assessment model is also proved to be effective.

Key words: Bayesian network; small data sets; binary particle swarm optimization; threat assessment

1 引言

针对小数据集条件下的贝叶斯网络(Bayesian Network, BN)结构学习,主要有两类方法:基于扩展数据的方法和基于先验约束的方法,本文属于后者. 基于先验约束的方法有两种思路,一种是将先验约束转化为狄利克雷分布中的等价样本量^[1]来约束评分函数,另一种则是通过定性知识限制贝叶斯网络的搜索过程. 基于等价样本量方法的优点在于计算简单,而缺点在于无法灵活的融入各种形式的专家知识,因为对于领域

专家来说,往往更倾向于给出定性知识而非精确的数值. 进而,一些定性的结构约束模型被提出,如节点序^[2,3]、因果关系^[4]、边的存在^[5,6]等各种形式. 但上述约束并未考虑约束的不确定性,虽然吴红等人^[7]用边的存在概率描述约束的不确定性,并将其以权重的形式加入到最小描述长度(Minimum Description Length, MDL)评分之中,进而采用模拟退火算法进行搜索. 但其提出的约束模型只是考虑到节点之间存在边的概率,并未考虑到不存在边和存在反向边的概率,并且其

收稿日期:2014-10-21;修回日期:2015-05-10;责任编辑:覃怀银

基金项目:国家自然科学基金(No. 60774064, No. 61305133);全国高校博士点基金(No. 20116102110026);中央高校基本科研业务费专项基金(No. 3102015KY0902, No. 3102015BJ(II)GH01)

修改 MDL 评分的方法是一种近似方法,缺少一定的理论支撑.针对此问题,本文提出一种新的结构约束,并将约束以先验概率的形式融入到评分函数中,进而结合二值粒子群优化算法学习 BN 结构.

针对小数据集条件下的 BN 参数学习,现有的方法基本可分为两类.第一类是利用约束将参数学习问题转化为似然函数或者熵函数的约束优化问题^[8-11],进而应用凸优化^[8,10]或梯度法^[9,11]进行求解.此类方法优点是参数学习提供了一个统一的框架,但缺点在于算法复杂度较高,且学习精度一般.第二类是将约束通过虚拟样本的形式转化为参数的先验信息^[13-15],进而结合贝叶斯估计进行求解.此类方法的缺点是针对不同的约束必须有不同的处理方法,但其优点在于针对某些特定约束其计算简单且精度往往较高,本文属于此类方法.此外,文献[12]将定性影响约束转化为序约束并应用保序回归算法调整参数大小,使得参数满足序约束.虽然其计算简单,但精度难以保证.通过分析现有的文献,我们发现所有的参数约束都是描述子节点相同父节点不同的这一类参数之间的大小关系.而对父节点相同而子节点不同这一类参数之间的关系并未提及.因此,本文提出一种单调性约束模型来描述此类约束,并给出基于此模型的参数学习方法.

2 基于连接概率分布的 BN 结构学习

2.1 连接概率分布的结构约束模型

贝叶斯网络是描述各个变量之间连接关系的有向无环图,先验概率指在这些变量集所组成的网络空间中,某一具体网络的存在概率.一方面,专家很难有这样的经验,另一方面,当变量数大于 3 时,专家很难在这么大的一个空间里给出某一具体网络的概率.专家一般只能对网络中的某些变量之间的连接概率给出一些经验.设 a 和 b 为贝叶斯网络中的两个变量,同时,引入连接变量 r ,在本文中 r 的取值为 a 指向 b ($a \rightarrow b$)、 b 指向 a ($b \rightarrow a$)、 a 和 b 之间无连接 ($a \cdots b$),本文中专家经验表示为连接变量的概率分布,由概率分布的定义可得以下结论,如式(1)所示.

$$P(a \rightarrow b) + P(b \rightarrow a) + P(a \cdots b) = 1 \quad (1)$$

假设在一个贝叶斯网络中,专家给出 $P(a \rightarrow b) = 0.8$,那么,我们即认为连接变量 r 剩下的取值服从均匀分布,如式(2)所示.

$$P(b \rightarrow a) = P(a \cdots b) = \frac{1 - P(a \rightarrow b)}{2} = 0.1 \quad (2)$$

依据式(1)和(2)可得到贝叶斯网络中所有连接变量的概率分布,在得到概率分布之后,进而根据此概率分布求出每个网络的先验概率.而在实际的问题中,专家不可能对所有的变量都给出约束,往往只是对其中

部分变量给出约束,所以如何利用部分约束获得每个网络的先验概率是首先要解决的问题.

2.2 基于改进先验分布的评分函数

设整个网络有 m 个变量, G 为其对应的有向无环图, D 为样本数据, r_1, r_2, \dots, r_n 为专家约束所对应的连接变量,令 $J = P(r_1, r_2, \dots, r_n)$, $R = (r_1, r_2, \dots, r_n)$, 如果 C 为连接变量联合分布的某一取值,则 $J_c = P(R = C)$, 在 m 个变量所对应的解空间中,每一个解都唯一决定 C 的取值.本文要求取的评分函数为 $P(G|D, J)$ 如式(3)所示.

$$\begin{aligned} P(G|D, J) &= \frac{P(D|G, J) \cdot P(G|J)}{P(D|J)} \\ &= \frac{P(D|G) \cdot P(G|J)}{P(D|J)} \end{aligned} \quad (3)$$

第一个等号由贝叶斯定理可以得到,第二个等号成立是因为当 G 给定时, J 与数据 D 是条件独立的,所以 $P(D|G, J) = P(D|G)$. 当专家约束和网络数据给定时, $P(D|J)$ 就是一个常数,所以,要使得整个网络的评分最高,只需使得分子 $P(D|G) \cdot P(G|J)$ 取最大即可.

$$\begin{aligned} P(G|J) &= P(G, C_c|J) = P(G|J, C_c) \cdot P(C_c|J) \\ &= P(G|C_c) \cdot P(C_c|J) \\ &= P(G|C_c) \cdot J_c \end{aligned} \quad (4)$$

其中 C_c 为网络空间中某一网络所对应的约束连接变量, C_c 的取值由网络结构 G 唯一确定,其分布由专家经验确定.由于 C_c 的取值由 G 唯一确定,所以第一个等号成立.根据贝叶斯定理和链式规则可得到第二个等号成立.当约束所对应的连接变量取值已知时,网络结构 G 和连接变量的分布条件独立,则第三个等号成立.当 C_c 已知条件下,某一网络结构 G 的存在概率实际上就是包含这一特定网络子结构的概率,如果能够得到在整个网络空间中含有这一特定子结构的网络个数 N_c ,那么可以得到:

$$P(G|C_c) = \frac{1}{N_c} \quad (5)$$

为了表述方便,这里令 $P(C_c|J) = J_c$,进而完整的评分函数则表示为:

$$\begin{aligned} P(G|D, J) &= \frac{P(D|G, J) \cdot P(G|J)}{P(D|J)} \\ &= \frac{P(D|G)}{P(D|J)} \cdot \frac{J_c}{N_c} \end{aligned} \quad (6)$$

由于 $P(D|J)$ 为一常量,进而将评分简化为:

$$\begin{aligned} \log P(G|D, J) &= \log P(D|G) + \log P(G|J) \\ &= \log P(D|G) + \log \frac{J_c}{N_c} \end{aligned} \quad (7)$$

上面的评分属于贝叶斯统计评分这一大类,与现有的评分函数(BD 评分)的区别在于结构先验分布不同, BD 评分一般认为结构的先验分布服从均匀分布,而本文

提及的评分函数中,贝叶斯网络的结构先验分布不一定服从均匀分布.其中,某一具体网络结构的先验概率而是由专家经验决定的.为了方便描述,本文中将此评分函数命名为评分 I-BD.具体表达式如(8)所示.

$$\begin{aligned} \text{Score}_{\text{I-BD}} &= \sum_{i=1}^n \sum_{j=1}^{q_i} \left[\log \frac{\Gamma(\alpha_{ij^*})}{\Gamma(\alpha_{ij^*} + m_{ij^*})} \right. \\ &\quad \left. + \sum_{k=1}^{r_i} \log \frac{\Gamma(\alpha_{ijk} + m_{ijk})}{\Gamma(\alpha_{ijk})} \right] + \log \frac{J_C}{N_C} \end{aligned} \quad (8)$$

其中, m_{ijk} 是数据中满足 $X_i = k$, 其父节点 $\pi(X_i) = j$ 的样本个数, α_{ijk} 为参数先验分布中的超参数.

J_C 为专家经验所对应连接变量的联合概率分布,假设专家给出 3 个变量 a, b, c 之间的约束模型如表 1 所示,依据表 1 中每个连接变量的分布,进而求得连接变量的联合分布如表 2 所示.

表 1 专家经验对应的约束模型

	→	←	...
r_{ab}	0.8	0.1	0.1
r_{bc}	0.2	0.2	0.6
r_{ac}	0.6	0.3	0.1

表 2 专家经验对应的 J_C

	r_{ab}	r_{bc}	r_{ac}	J_C
C_1	→	→	→	0.096
C_2	→	→	←	0
C_3	→	→	...	0.016
...
C_{13}	←	←	→	0
...
C_{26}	←	0.018
C_{27}	0.006

由表 1 和 2 可以看出 J_C 是通过将 3 个节点所组成的有向无环图进行列举,并且分别计算每个有向无环图存在的概率.

n 个变量组成有向无环图的个数由 Robinson 给出了解析表达式,经过查找文献,并未发现对含有特定结构约束的有向无环图个数求解的解析表达式,即本文的 N_C . 当网络的变量较少时(一般小于 4),我们可以通过枚举法将所有的有向无环图都列举出来,找出其中符合要求的网络结构,进而得到 N_C . 但往往枚举是行不通的,随着变量数的增加有向无环图的个数呈指数倍数增加.在这种情况下,精确的计算 N_C 是很困难的,所以,本文给出一种基于抽样的近似方法.

$$N_C = N \cdot \frac{S_C}{S} \quad (9)$$

$$\begin{cases} f(0) = 1, f(1) = 1 \\ N = f(n) = \sum_{i=1}^n (-1)^{i+1} \frac{n!}{(n-i)!i!} f(n-i), n > 1 \end{cases} \quad (10)$$

其中, N 为 n 个变量的有向无环图的个数,可由式(10)计算得到, S 为对 n 个变量所组成的网络结构空间的采样个数, S_C 为符合特定结构约束的样本个数.

2.3 I-BD-BPSO 算法实现

2.1 节将专家经验表示为连接概率分布,2.2 节利用连接概率分布对 BD 评分进行改进,进而得到 I-BD 评分,在此基础上,本节结合 BPSO 算法学习 BN 结构.算法中涉及到的粒子表示、更新及无效结构处理的方法可参考文献[16].具体流程如算法 1.

算法 1 I-BD-BPSO

Input: 数据 D , 专家经验 E , BPSO 算法相关参数
 1: 初始化 m 个粒子组成的初始种群 $Q(t)$, 进化代数 $t=0$;
 2: 对粒子进行无效结构处理;
 3: 根据专家经验 E , 利用式(9)和(10)计算 J_i 和 N_i ;
 4: while $t <$ 设定进化代数 do
 5: 计算每个粒子的 I-BD 评分;
 6: 更新局部最优和全局最优;
 7: 根据局部最优和全局最优计算粒子速度;
 8: 通过粒子速度来更新粒子的位置,进而得到新的粒子;
 9: 对粒子进行无效结构处理,进而获得新的种群;
 10: $t = t + 1$;
 11: end while;
Output: 获得新种群的全局最优值

3 基于单调性约束的贝叶斯网络参数学习算法

3.1 单调性约束模型

一个由 n 个变量 $X = \{X_1, X_2, \dots, X_n\}$ 组成的贝叶斯网 B , 不失一般性,设其中的节点 X_i 共有 r_i 个取值,其父节点 $\pi(X_i)$ 的取值共有 q_i 个组合. 对一固定的 j (即父节点的取值给定时)有:

$$k_1 > k_2 \Rightarrow P(X = k_1 | \pi(X) = j) > P(X = k_2 | \pi(X) = j) \quad (11)$$

假设贝叶斯网 B 的参数符合单调性约束,为了推导方便,这里令 $P(X = k_1 | \pi(X) = j) = z_1$, 那么 $P(X = k_n | \pi(X) = j) = z_n$, 于是由参数的规范性和单调性可得:

$$\begin{cases} 0 < z_1 < \frac{1}{n} \\ z_1 < z_2 < \frac{1-z_1}{n-1} \\ \dots \\ z_{n-2} < z_{n-1} < \frac{1-\sum_{i=1}^{n-2} z_i}{2} \\ z_n = 1 - \sum_{i=1}^{n-1} z_i \end{cases} \Rightarrow \begin{cases} z_1 + z_2 + \dots + z_n = 1 \\ z_1 < z_2 < \dots < z_n \end{cases} \quad (12)$$

3.2 先验分布中超参数的确定

考虑 θ 的贝叶斯估计,一般情况下,在计算 $P(D|\theta)$ 时需要对先验概率 $P(\theta)$ 做一些假设:

假设 1:关于不同变量 X_i 的参数相互独立.

假设 2:给定一个变量 X_i 对应于 $\pi(X_i)$ 的不同取值的参数相互独立.

假设 3: $P(\theta_{ij*})$ 是狄利克雷分布,其中 θ_{ij*} 由 $\theta_{j1}\theta_{j2}, \dots, \theta_{jr}$ 所组成. 所以,一般假设先验分布 $P(\theta)$ 是狄利克雷分布 $D(\alpha_1, \alpha_2, \dots, \alpha_r)$, 即

$$p(\theta) = \frac{\Gamma(\alpha)}{\prod_{i=1}^r \Gamma(\alpha_i)} \prod_{i=1}^r \theta_i^{\alpha_i-1} \quad (13)$$

注意狄利克雷分布 $D(\alpha_1, \alpha_2, \dots, \alpha_r)$ 的边缘分布是 $B(\alpha_1, \alpha_2)$, 即对于每个参数 θ_{ji} 或 θ_{jr} 服从 $B(\alpha_1, \alpha_2)$. 假设 $P(\theta)$ 为狄利克雷分布 $D(\alpha_1, \alpha_2, \dots, \alpha_r)$ 就等于假设关于 θ 的先验知识相当于 α 个虚拟数据样本,其中满足 $X = x_i$ 的样本数为 α_i . 由贝叶斯估计可得:

$$\theta_i^* = \int \theta_i p(\theta_i | D) d\theta_i = \frac{m_i + \alpha_i}{m + \alpha} \quad (14)$$

式(12)给出了参数的取值范围,一般可认为参数在区间内服从均匀分布,设参数 θ 服从 $U(\theta_1, \theta_2)$, 问题转化为用 $B(\alpha_1, \alpha_2)$ 去等效 $U(\theta_1, \theta_2)$, 即:

$$\min_{\alpha_1, \alpha_2} \int_0^1 (U(\theta_1, \theta_2) - B(\alpha_1, \alpha_2))^2 d\theta \quad (15)$$

为了解决上述问题,采用二阶矩法进行拟合:

$$m_1 = \int_0^1 U(\theta) \theta d\theta \quad (16)$$

$$m_2 = \int_0^1 U(\theta) \theta^2 d\theta$$

$$M_1 = \int_0^1 B(\theta) \theta d\theta = \frac{\alpha_1}{\alpha_1 + \alpha_2}$$

$$M_2 = \int_0^1 B(\theta) \theta^2 d\theta = \frac{\alpha_1 \alpha_2}{(\alpha_1 + \alpha_2)^2 (\alpha_1 + \alpha_2 + 1)} + M_1^2 \quad (17)$$

$$\begin{cases} M_1 = m_1 \\ M_2 = m_2 \end{cases} \Rightarrow \begin{cases} \alpha_1 = \frac{m_1^2 - m_1 m_2}{m_2 - m_1^2} \\ \alpha_2 = \frac{m_1 - m_2}{m_2 - m_1^2} (1 - m_1) \end{cases} \quad (18)$$

通过式(15)到(18)可以求得参数 α_1 和 α_2 , 并将其带入到式(14)中,从而获得 θ 的贝叶斯估计.

3.3 MCE 算法实现

假设贝叶斯网 B 的参数 θ_{ijk} 符合单调性约束, θ_{ijk} 中 k 有 n 种取值,应用 3.1 节的单调性约束模型,结合 3.2 节的先验分布超参数确定方法,以贝叶斯估计为算法基本框架,进而给出 MCE 算法流程. 如算法 2.

算法 2 MCE

Input: 数据 D , 专家经验 E

- 1: 根据专家经验 E 构建单调性约束模型, 令 $k = 1$;
- 2: while $k < n + 1$ do
- 3: 由式(12)得到参数的 θ_{ijk} 的约束区间;
- 4: 利用式(15~18)获取 θ_{ijk} 先验分布的超参数;
- 5: 利用式(14)计算 θ_{ijk} 的贝叶斯估计;
- 6: $k = k + 1$;
- 7: end while;

Output: 符合单调性约束的所有参数

4 仿真实验

本文中结构学习和参数学习过程是分开进行的. 首先,利用第 2 节提到的结构学习算法学习结构,其次,在结构已知的条件下,利用第 3 节的参数学习算法学习参数. 仿真实验部分由 3 部分组成:4.1 节验证结构算法仿真比较,4.2 节验证参数学习算法仿真比较,4.3 节利用结构和参数学习算法构建 UAV 威胁评估模型.

4.1 结构学习算法仿真

图 1 给出结构学习部分的仿真模型及专家约束模型:

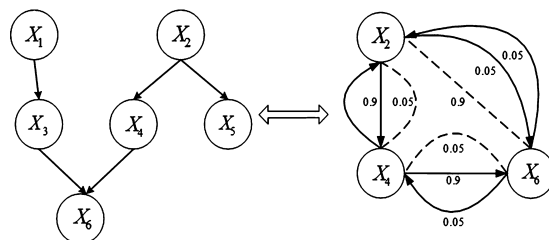


图 1 仿真实验模型及专家约束模型

为了证明在小数据集条件下,利用 I-BD 评分能够将专家经验引入到贝叶斯结构学习中,进而得到更为准确的网络结构,本文引入汉明距离和标准的 BD 评分进行仿真分析.

$$\text{汉明距离} = \text{丢失的边} + \text{反转的边} + \text{多余的边} \quad (9)$$

Score_{BD}

$$= \sum_{i=1}^n \sum_{j=1}^{q_i} \left[\log \frac{\Gamma(\alpha_{ij*})}{\Gamma(\alpha_{ij*} + m_{ij*})} + \sum_{k=1}^{r_i} \log \frac{\Gamma(\alpha_{ijk} + m_{ijk})}{\Gamma(\alpha_{ijk})} \right] \quad (20)$$

为了说明算法的有效性,在样本数为 50 的条件下,与经典的 BPSO 算法进行比较,分别利用 BD 评分和汉明距离对两种算法进行评估,仿真结果如下:

通过对图 2 的仿真结果进行分析可知,在小数据集条件下,利用 I-BD 评分学习得到的贝叶斯网络结构的汉明距离较小,BD 评分较高,说明网络更符合真实网

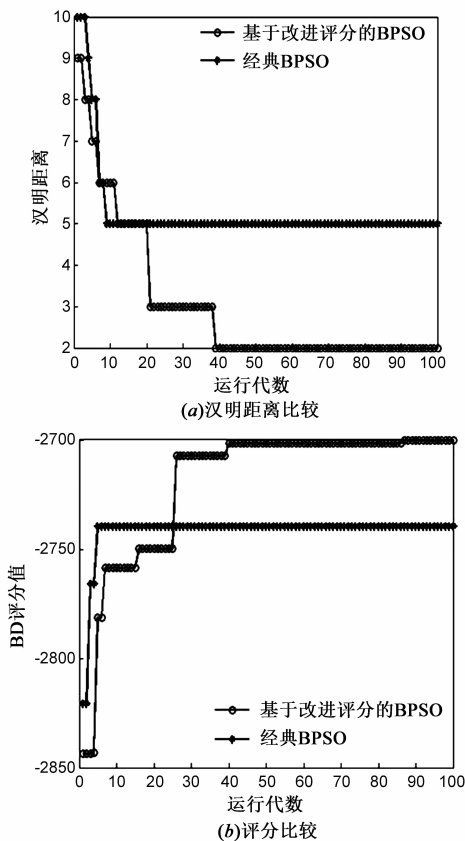


图2 运行100代两种算法仿真结果

络,进而证明 I-BD-BPSO 算法的有效性.

4.2 参数学习算法仿真

使用 BN 学习中经典的草坪湿润网络模型进行算法验证实验,我们通过专家经验给出网络中参数所服从的单调性约束,如式(21)所示.

$$\begin{cases} P(R=1|C=1) > P(R=0|C=1) \\ P(R=1|C=0) < P(R=0|C=0) \\ P(S=1|C=1) < P(S=0|C=1) \\ P(S=1|C=0) > P(S=0|C=0) \\ P(W=1|R=1,S=1) > P(W=0|R=1,S=1) \\ P(W=1|R=1,S=0) < P(W=0|R=1,S=0) \\ P(W=1|R=0,S=1) < P(W=0|R=0,S=1) \\ P(W=1|R=0,S=0) > P(W=0|R=0,S=0) \end{cases} \quad (21)$$

将 MCE 算法与最大似然估计(Maximum Likelihood Estimation, MLE)、保序回归(Isotonic Regression Estimation, IRE)和凸优化(Convex Optimization Estimation, COE)进行对比,采用 KL(Kullback-Leibler)距离反映参数学习精度.

$$KL(P, Q) = \sum_x P(X) \log \frac{P(X)}{Q(X)} \quad (22)$$

本文分别对 4 种算法所得参数分布的 KL 距离进

行对比仿真,分别将每种算法运行 10 次,求取其平均值,仿真结果如图 3 所示,图 3(b)是局部放大图.从图上可以看出,本文的 MCE 方法能够得到最小的 KL 距离,获得参数的准确性最高,保序回归算法和凸优化算法稍差一些,但都要优于 MLE 算法.在样本数据量小于 20 时,本文算法的优势尤为明显,而随着样本数的增加,保序回归算法和凸优化与本文算法的差距逐渐减小.

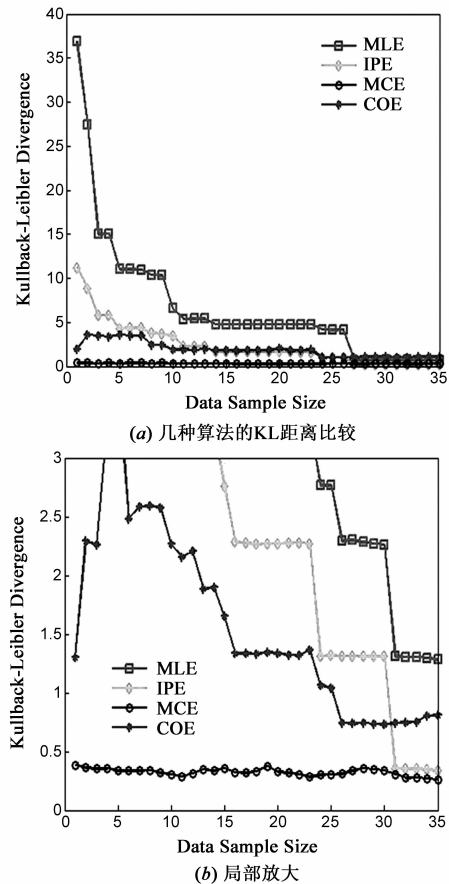


图3 几种算法的仿真结果比较

4.3 UAV 威胁评估建模仿真

4.3.1 任务想定

红方:有人作战飞机和无人机协同执行对蓝方的防空压制.

蓝方:由指挥站和若干防空导弹车组成防空系统.

作战描述:为提高作战效率,降低作战成本,红方采用有人作战飞机指挥控制无人机的作战模式对蓝方防空力量进行压制,当无人作战飞机进入蓝方防空区域一段时间后,受到蓝方强电磁干扰,为确保红方有人作战飞机的安全,有人作战飞机必须撤离,由无人机单独完成对敌防空压制任务,在执行防空压制任务时需要对方目标进行威胁评估,进而为攻击决策提供支持.

想定中考虑与威胁(T)相关的因素为目标的速度

(V)、距离(D)和目标与无人机进入角之和(A),下面对这几个威胁因素的取值进行解释.

(1) 目标的速度(V):速度特指速度的大小. 令无人机速度为 V_u , 目标的速度的取值为式(23)所示.

$$V = \begin{cases} 1, & V < 0.6V_u \\ 2, & 0.6V_u \leq V < 1.5V_u \\ 3, & V \geq 1.5V_u \end{cases} \quad (23)$$

(2) 目标的距离(D):目标与无人机之间的距离. 距离的取值为式(24)所示.

$$D = \begin{cases} 1, & 40\text{km} \leq D < 60\text{km} \\ 2, & 20\text{km} \leq D < 40\text{km} \\ 3, & D < 20\text{km} \end{cases} \quad (24)$$

(3) 目标和无人机的进入角之和(A):目标的速度方向与目标线的夹角和无人机的速度方向与目标线的夹角之和. 进入角之和的取值为式(25)所示.

$$A = \begin{cases} 1, & A \geq \pi \\ 2, & A < \pi \end{cases} \quad (25)$$

(4) 目标的威胁(T):威胁值大小是对以上 3 个因素的综合, 目标的威胁取值由式(26) ~ (30) 给出.

$$T = 0.5d + 0.3a + 0.2v \quad (26)$$

$$d = \begin{cases} 0, & d > 60\text{km} \\ 1 - \frac{d}{60}, & d \leq 60\text{km} \end{cases} \quad (27)$$

$$a = 1 - \frac{A}{2\pi} \quad (28)$$

$$v = \begin{cases} 0.105, & V < 0.6V_u \\ \frac{V}{V_u} - 0.5, & 0.6V_u \leq V < 1.5V_u \\ 0.986, & V \geq 1.5V_u \end{cases} \quad (29)$$

$$T = \begin{cases} 1, & T < 0.3 \\ 2, & 0.3 < T < 0.6 \\ 3, & T > 0.6 \end{cases} \quad (30)$$

4.3.2 仿真过程

在 4.3.1 节的任务想定下,利用本文提出 I-BD-BP-SO 算法与 MCE 算法,实现 UAV 在与有人机通信中断后的小数据集条件下的威胁评估模型构建,并通过剩余的样本数据进行推理分析,进一步说明本文提出的小数据集下的 BN 学习算法的有效性以及可行性. 下面给出具体的仿真条件和仿真方法:

(1) 本文仿真工作使用 Matlab 版本为 R2010a, Vc++ 版本为 6.0, 运行环境为操作系统 Window XP.

(2) 在通信中断前,有人机实时接收 UAV 传来的目标信息,进而进行威胁评估和攻击决策. 实验中首先对想定过程进行仿真,进而采集相关的实验数据,利用式(23) ~ (30) 对相关实验数据进行离散化,实验中共获得仿真数据 1200 组,用其中 50 组作为小数据集构建

BN 模型.

(3) 得到威胁评估模型之后应用变量消元法并结合剩余的仿真数据进行推理,进而说明小数据集条件下本文算法在 UAV 威胁评估中的可行性.

4.3.3 仿真结果

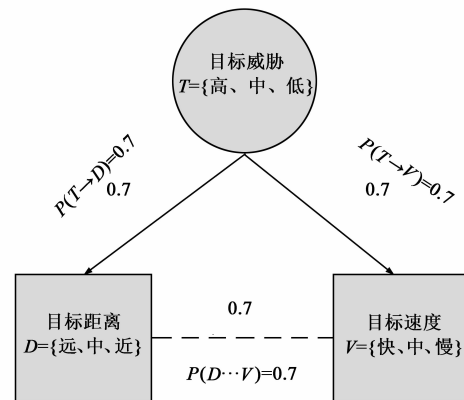
首先,在仿真之前给出专家经验所对应的结构约束如图 4(a),然后,在样本数据为 50 的条件下构建威胁评估网络,结果如图 4(b)所示. 实验中令目标威胁节点为 1 号节点,目标距离、目标与无人机的进入角之和、目标速度节点分别为 2、3、4 号节点.

当获得威胁评估网络的结构之后,需要确定网络的参数,下面给出参数约束:

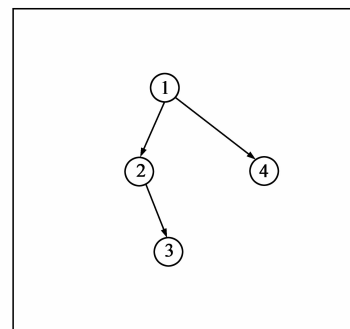
$$\begin{cases} P(D=1|T=1) > P(D=2|T=1) > P(D=3|T=1) \\ P(D=2|T=2) > P(D=1|T=2) > P(D=3|T=2) \\ P(D=3|T=3) > P(D=2|T=3) > P(D=1|T=3) \end{cases} \quad (31)$$

$$\begin{cases} P(V=1|T=1) > P(V=2|T=1) > P(V=3|T=1) \\ P(V=2|T=2) > P(V=1|T=2) > P(V=3|T=2) \\ P(V=3|T=3) > P(V=2|T=3) > P(V=1|T=3) \end{cases} \quad (32)$$

$$\begin{cases} P(A=1|D=1) < P(A=2|D=1) \\ P(A=1|D=2) < P(A=2|D=2) \\ P(A=1|D=3) > P(A=2|D=3) \end{cases} \quad (33)$$



(a) 威胁评估网络的约束模型



(b) 本文算法所构建的网络

图 4 约束模型与本文算法所得结构

式(31)~(33)中 $D = \{1, 2, 3\}$ 表示 $D = \{\text{远}, \text{中}, \text{近}\}$, $V = \{1, 2, 3\}$ 表示 $V = \{\text{慢}, \text{中}, \text{快}\}$, $A = \{1, 2\}$ 表示 $A = \{\text{大}, \text{小}\}$, $T = \{1, 2, 3\}$ 表示 $T = \{\text{低}, \text{中}, \text{高}\}$. 进而得到参数学习结果如表 3 所示.

表 3 样本数量为 50 时本文算法所得网络参数

目标威胁程度 (T1:低 T2:中 T3:高)	目标距离 (D1:远 D2:中 D3:近)	进入角之和 (A1:大 A2:小)	目标速度 (V1:慢 V2:中 V3:快)
T1(低)	(0.5854, 0.2776, 0.1369)	(0.1529, 0.8471)	(0.5587, 0.3032, 0.1381)
T2(中)	(0.3748, 0.5097, 0.1154)	(0.2759, 0.7241)	(0.4353, 0.3135, 0.2511)
T3(高)	(0.1574, 0.2863, 0.5563)	(0.7708, 0.2292)	(0.1587, 0.2851, 0.5562)

为了进一步证明算法的有效性和可行性,将所得模型用于威胁评估,通过剩余的 1150 条样本数据来检验 BN 威胁评估模型的推理准确度,推理算法选择变量消元法,推理的准确度具体计算过程如下:

$$T = \arg \max_{i \in \{1, 2, 3\}} (p(T = i | D, A, V)) = \arg \max_{i \in \{1, 2, 3\}} (p(T = i | D, V)) \quad (34)$$

$$N(j) = \begin{cases} 1, & T(j) = \text{Data}_T(j) \\ 0, & \text{其他情况} \end{cases} \quad (35)$$

$$\eta = \frac{\sum_{j=1}^n N(j)}{n} \quad (36)$$

其中 T 表示威胁变量, n 为用于检验的样本个数,取值为 1150, $\text{Data}_T(j)$ 表示第 j 条数据中的威胁变量的取值, $T(j)$ 表示推理算法得到的威胁变量取值. 利用式(34)~(36)得 $\eta = 0.7035$. 由此可见,小数据条件下的 BN 威胁评估模型能够很好的对威胁数据进行建模.

5 总结

本文提出了一种小数据集条件下的 BN 建模方法,并构建威胁评估模型. 仿真结果表明,在小数据集条件下,本文提出的建模方法通过结合正确的专家经验,能够得到较好的结构和参数,进而保证模型的推理精度,为小数据集条件下的建模问题提供了一种有效的方法. 由于专家经验的正确与否会直接影响到模型的构建,下一步的主要研究当专家提供的约束不完全正确时,如何在约束和数据之间取得一个折中,使得算法具有一定的自适应性.

参考文献

[1] SILANDER T, KONTKANEN P, MYLLYMAKI P. On sensitivity of the map Bayesian network structure to the equivalent sample size parameter [A]. David H. Proceedings of the 23rd International Conference on Uncertainty in Artificial Intelligence [C]. Vancouver: Springer, 2008. 360 - 367.
 [2] 朱明敏,刘三阳,汪春峰. 基于先验节点序学习贝叶斯网络结构的优化方法[J]. 自动化学报,2011,37(12):1514

- 1519.
 ZHU Ming-min, LIU San-yang, WANG Chun-feng. An optimization approach for structural learning Bayesian networks based on prior node ordering [J]. Acta Automatica Sinica, 2011, 37(12): 1154 - 1519. (in Chinese)
 [3] COOPER G F, HERSKOVITS E. A Bayesian method for the induction of probabilistic networks from data [J]. Machine Learning, 1992, 9(4): 309 - 347.
 [4] BORBOUDAKIS G, TSAMARDINOS I. Incorporating causal prior knowledge as path-constraints in Bayesian networks and maximal ancestral graphs [A]. Andrew M. Proceedings of the 29th International Conference on Machine Learning [C]. Edinburgh: ACM, 2012. 28 - 36.
 [5] COMPOSC P, ZENG Z, JI Q. Structure learning of Bayesian networks using constraints [A]. Dan R. Proceedings of the 26th International Conference on Machine Learning [C]. Montreal: ACM, 2009. 280 - 287.
 [6] CAMPOSC P, JI Q. Efficient structure learning of Bayesian networks using constraints [J]. Journal of Machine Learning Research, 2011, 12(3): 663 - 689.
 [7] 吴红,王维平,杨峰. 融合先验信息的贝叶斯网络结构学习方法 [J]. 系统工程与电子技术, 2012, 34(12): 2585 - 2591.
 WU Hong, WANG Wei-ping, YANG Feng. Structure learning method of Bayesian network with prior information [J]. Systems Engineering and Electronics, 2012, 34(12): 2585 - 2591. (in Chinese)
 [8] ALTENDORF E, RESTICAR A. Learning from sparse data by exploiting monotonicity constraints [A]. Max C. Proceedings of the 21st International Conference on Uncertainty in Artificial Intelligence [C]. Edinburgh: Springer, 2005. 18 - 26.
 [9] CAMPOSC C P, JI Q. Improving Bayesian network parameter learning using constraints [A]. Ejiri M. Proceedings of the 19th International Conference on Pattern Recognition [C]. Florida: IEEE, 2008. 1 - 4.
 [10] NICULESCU R S, MITCHELL T M. Bayesian network learning with parameter constraints [J]. Journal of Machine Learning Research, 2006, 7(1): 1357 - 1383.
 [11] LIAO W H, QIANG J. Learning Bayesian network param-

- eters under incomplete data with domain knowledge[J]. Pattern Recognition, 2009, 42(11): 3046 – 3056.
- [12] FEELDER A D, VANDERGAGA L. Learning Bayesian networks parameters under order constraints[J]. Journal of Approximate Reasoning, 2006, 42(1): 37 – 53.
- [13] 任佳, 高晓光, 茹伟. 数据缺失的小样本条件下 BN 参数学习[J]. 系统工程理论与实践, 2011, 31(1): 172 – 177.
REN Jia, GAO Xiao-guang, RU Wei. Parameters learning of BN in small sample on data missing[J]. Systems Engineering—Theory & Practice, 2011, 31(1): 172 – 177. (in Chinese)
- [14] RUI C, SHOEMAKER R, WANG W. Anovel knowledge-driven systems biology approach for phenotype prediction upon genetic intervention[J]. IEEE Transactions on Computational Biology and Bioinformatics, 2011, 1(8): 1170 – 1181.
- [15] YUN Z, FENTON N, NEIL M. Bayesian network approach to multinomial parameter learning using data and expert judgments[J]. Journal of Approximate Reasoning, 2014, 55(2): 1252 – 1268.
- [16] 高晓光, 邸若海, 郭志高. 基于改进粒子群优化算法的贝叶斯网络结构学习[J]. 西北工业大学学报, 2014, 32(5): 749 – 754.
GAO Xiao-guang, DI Ruo-hai, GUO Zhi-gao. Structure learning of BN based on improved particle swarm optimization[J]. Journal of Northwest Polytechnical University, 2014, 32(5): 749 – 754. (in Chinese)

作者简介



邸若海 男, 1986 年出生于陕西西安. 现为西北工业大学博士研究生. 主要研究方向为复杂系统建模和贝叶斯网络学习.
E-mail: xfwtdrh@163.com



高晓光 (通信作者) 女, 1957 年出生于辽宁沈阳. 现为西北工业大学教授, 博士生导师. 主要研究方向为复杂系统建模及效能评估.
E-mail: cxg2012@nwpu.edu.cn