

# 变结构动态贝叶斯网络的机制研究

高晓光<sup>1</sup> 陈海洋<sup>2</sup> 史建国<sup>3</sup>

**摘要** 传统的动态贝叶斯网络 (Dynamic Bayesian networks, DBNs) 描述的是一个稳态过程, 而处理非稳态过程, 变结构动态贝叶斯网络更适用、更灵活、更有效。为了克服现有变结构离散动态贝叶斯网络推理算法只能处理硬证据的缺陷, 本文在深入分析变结构动态贝叶斯网络机制及其特征的基础上, 提出了变结构离散动态贝叶斯网络的快速推理算法。此外, 对变结构动态贝叶斯网络的特例, 即数据缺失动态贝叶斯网络进行了定义并构建了相应的模型。仿真实验验证了变结构离散动态贝叶斯网络快速推理算法的有效性及其计算效率。

**关键词** 动态贝叶斯网络, 推理, 软证据, 复杂度

**DOI** 10.3724/SP.J.1004.2011.01435

## Study on the Mechanism of Structure-variable Dynamic Bayesian Networks

GAO Xiao-Guang<sup>1</sup> CHEN Hai-Yang<sup>2</sup> SHI Jian-Guo<sup>3</sup>

**Abstract** Traditional dynamic Bayesian networks (DBNs) are essentially models that describe a variety of stable processes. To deal with unstable processes, structure-variable dynamic Bayesian networks are more applicable, flexible, and effective. Currently, however, the various inference algorithms under consideration for structure-variable discrete dynamic Bayesian networks (DDBNs) can only handle hard evidence. In this paper, an in-depth and theoretical analysis is given for the mechanism and key characteristics of structure-variable dynamic Bayesian networks, and on this basis, a fast inference algorithm is proposed. Furthermore, a special class of structure-variable dynamic Bayesian networks, dynamic Bayesian networks with missing data, is defined rigorously along with associated network topology and parameter settings of such networks. Several experimental simulations have shown the effectiveness and efficiency of our fast inference algorithm.

**Key words** Dynamic Bayesian networks (DBNs), inference, soft evidences, complexity

贝叶斯网络<sup>[1-2]</sup> 是一种图形化的模型, 能够图形化地表示一组变量间的联合概率分布函数。贝叶斯网络分为静态贝叶斯网络和动态贝叶斯网络 (Dynamic Bayesian networks, DBNs)。动态贝叶斯网络<sup>[3-4]</sup> 是静态贝叶斯网络的进一步发展, 它除了继承静态贝叶斯网络的优点外, 还具有静态贝叶斯网络所没有的优点: 在分析问题的过程中考虑了时间因素, 在事件推理过程中具有前后连续性, 使得这种推理方法更符合客观世界, 同时采用概率方法结合专家知识进行描述, 并结合历史信息和证据库, 因此, 具有信息时间积累能力, 能更有效地降低不同层次信息综合推理过程中的不确定性。基于此, 利用动态贝叶斯网络解决不确定性问题越来越受到众多研究人员的青睐<sup>[5-9]</sup>。动态贝叶斯网络具

有多种表现方式, 其中, 隐马尔科夫模型 (Hidden Markov models, HMM)<sup>[10-11]</sup> 与卡尔曼滤波模型 (Kalman filter models, KFM)<sup>[11-13]</sup> 是 DBNs 的特例, 同时也是目前研究和应用最广泛的模型。在这两个基本模型的基础上, 发展了其他结构更为复杂的动态贝叶斯网<sup>[5-9]</sup>, 它们都遵循两个基本假设<sup>[14]</sup>: 1) 一阶马尔科夫假设, 即各个节点之间的边或者位于同一时间片内, 或者位于相邻时间片之间, 不能跨越时间片; 2) 时齐性或齐次性, 即  $B_t$  中的参数不随时间片变化。在本文中, 我们把遵循这两个基本假设的动态贝叶斯网络称之为传统的动态贝叶斯网络, 它针对环境状态的变化是由一个稳态过程引起的, 而稳态过程意味着变化的过程是由本身不随时间变化的规律支配的。在稳态假设下, 每个时间片的网络结构相同, 只需要为某个“代表性的”时间片中的变量指定条件概率分布就可以了。

对于非稳态过程, 如图 1 所示, 它的整个过程可以看作是由若干个不同的稳态过程组成的, 对于每一个稳态过程, 我们可以构建传统的动态贝叶斯网络, 如  $DBN_1, DBN_2, \dots, DBN_n$ 。从一个稳态

收稿日期 2011-01-11 录用日期 2011-06-14  
Manuscript received January 11, 2011; accepted June 14, 2011  
国家自然科学基金 (60774064) 资助  
Supported by National Natural Science Foundation of China (60774064)  
1. 西北工业大学电子信息学院 西安 710129 2. 西安工程大学电子信息学院 西安 710048 3. 海军航空工程学院 7 系 烟台 264001  
1. School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710129 2. School of Electronic Information, Xi'an Polytechnic University, Xi'an 710048 3. No. 7 Department of Navy Aeronautical and Astronautical University, Yantai 264001

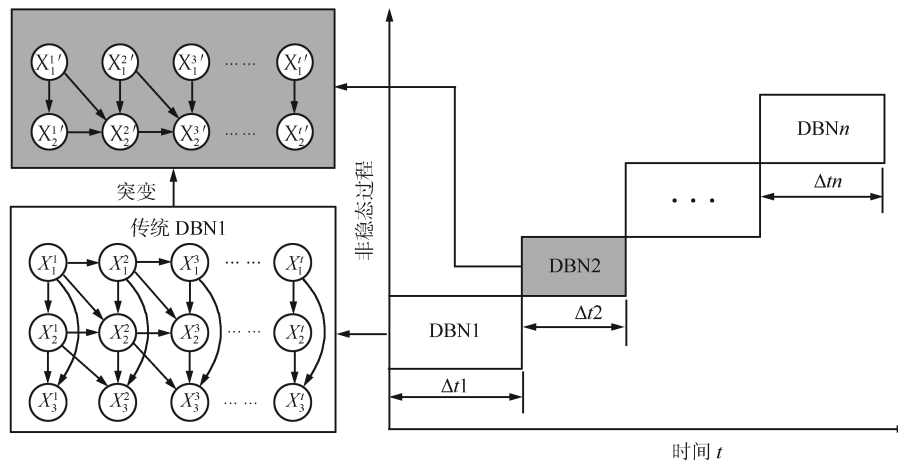


图 1 非稳态过程是由若干个稳态过程组成, 在每个稳态过程可以构建传统 DBNs

Fig. 1 An unstable process consists of several stable processes, and the traditional DBNs can be constructed on each of the stable processes

过程到另一个稳态过程, 反映了环境的状态发生了突变<sup>[15-16]</sup>. 对于整个非稳态过程, 每个时间片的网络结构可能不相同, 不存在一个“代表性的”时间片, 因此, 传统的动态贝叶斯网络的假设前提就得不到完全满足, 难以对环境突变过程进行建模和推理. 为了解决这类推理问题, 文献 [16] 从定义上对传统的 DDBNs (Discrete dynamic Bayesian networks) 进行了拓展, 提出了变结构 DDBNs 的概念, 允许各个时间片上的网络结构和参数可以不相同. 在此基础上, 本文从数学的角度对变结构动态贝叶斯网络进行了定义, 给出了它的性质, 并分析了变结构动态贝叶斯网络的数据结构. 如果单从结构上将传统的动态贝叶斯网络扩展到变结构动态贝叶斯网络并不难, 然而传统动态贝叶斯网络的推理算法是基于图形变换, 所以将传统的动态贝叶斯网络的推理算法直接应用在这种模型之上就存在问题.

以基于图形变换的接口算法为例<sup>[11]</sup>, 当每一个时间片的静态贝叶斯网络相同时, 实现起来还比较容易, 但当每一个时间片上的静态贝叶斯网络不同时, 将带来难以估计的计算量. 因此, 文献 [16] 从变结构动态贝叶斯网络的特点出发, 提出了变结构 DDBNs 的推理算法, 无需图形变换, 就能进行推理. 不过该算法只能处理硬证据, 即只能处理那些观测变量能确切地告诉我们处于它的哪一种状态的证据. 然而在实际的应用中, 往往获得的是软证据, 即不能确切地告诉我们观测变量处于它的哪一个状态, 只能给出观测变量属于它的某一个状态的概率. 为了克服该算法的这一缺陷, 本文借鉴文献 [17-18] 处理软证据的基本思想, 提出了改进的变结构 DDBNs 的推理算法, 通过对该算法的复杂度分析表明, 处

理软证据会增大算法的复杂度, 进而提出了变结构 DDBNs 的快速推理算法. 另外, 我们还研究了变结构动态贝叶斯网络的一个特例, 即在传统的动态贝叶斯网络上, 一些时间片上的某些观测节点的证据信息发生缺失而造成了网络结构的变化, 由此得到的变结构动态贝叶斯网络, 在本文中被义为数据缺失动态贝叶斯网络. 这类网络从本质上来讲不是真正的变结构动态贝叶斯网络, 因为它反映的是稳态随机过程, 而真正的变结构动态贝叶斯网络反映的是非稳态随机过程, 不过它具有变结构动态贝叶斯网络的特点, 因而也把它归为变结构动态贝叶斯网络. 传统的动态贝叶斯网络观测数据的缺失可能由多种原因引起的, 其中, 数据发生随机缺失的情况, 可以经过数据修补将变结构 DDBNs 恢复成传统的动态贝叶斯网络. 文献 [19-20] 已给出这类网络缺失数据的修补算法.

## 1 变结构动态贝叶斯网络

### 1.1 定义及其性质

**定义 1.** 具有  $T$  个时间片的变结构动态贝叶斯网络定义为  $(B^0, B^1, B^2, \dots, B^T)$ , 其中,  $B^0$  是一个初始贝叶斯网络, 定义了初始时刻的概率分布  $P(Z_0)$ .  $B^t$  是一个包含了两个相邻时间片的贝叶斯网络 (第  $t$  个时间片和第  $t-1$  个时间片), 定义了两个相邻时间片各变量之间的条件分布, 即  $P(Z^t|Z^{t-1}) = \prod_{i=1}^{n_t} P(Z_i^t|Pa(Z_i^t))$ , 其中,  $Z_i^t$  是表示第  $t$  个时间片的第  $i$  个节点,  $Pa(Z_i^t)$  是  $Z_i^t$  的父节点,  $n_t$  表示第  $t$  个时间片的节点数目.

变结构动态贝叶斯网络定义的  $B^t$  与传统动态

贝叶斯网络定义的  $B_{-}$  的不同点在于:

- 1)  $B_{-}^t$  所包含的两个时间片的贝叶斯网络结构可能随着  $t$  的变化而变化;
- 2) 放宽了对动态贝叶斯网络关于时齐性的前提假设, 即不再要求  $B_{-}^t$  中的参数相同.

图 2 所示是一个简单的变结构动态贝叶斯网络, 其中图 2(a) 是  $B^0$ , 定义了初始时刻的分布  $P(Z_0)$ . 图 2(b) 是  $B_{-}^t$ , 定义了某一结构发生变化的时刻与前一时刻之间的条件概率  $P(Z^t|Z^{t-1})$ , 其中, 该时间片中的节点数增加了 1 个.

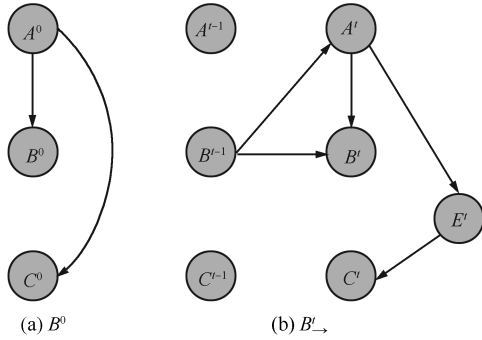


图 2 一个简单的变结构动态贝叶斯网络  
Fig. 2 A simple structure-variable DBN

根据初始分布和相邻时间片之间的条件分布, 可以将变结构动态贝叶斯网络展开到第  $T$  个时间片, 结果得到一个跨越多个时间片的联合概率分布, 即

$$P(Z^{0:T}) = \prod_{t=0}^{T-1} \prod_{i=1}^{n_t} P(Z_i^t | Pa(Z_i^t)) \quad (1)$$

变结构动态贝叶斯网络具有如下性质:

- 1) 非平稳性, 即  $B_{-}^t$  中的参数不是恒定不变的;
- 2) 一阶马尔科夫性, 即各节点之间的边, 或者位于同一个时间片内, 或者位于相邻时间片间, 但不能跨越时间片.

从数据结构上看, 变结构动态贝叶斯网络与传统动态贝叶斯网络相比, 差异可能有: 1) 结构相似, 但是变量的状态数发生变化; 2) 各个时间片内的网络结构发生变化, 即节点个数发生变化或变量之间的关系发生变化; 3) 相邻时间片间的变量之间的依赖关系发生变化; 4) 前三种的某几种变化的组合.

因此, 要完全表示一个  $T$  个时间片的变结构 DDBNs, 需要如下的数据结构:

- 1) 表示各个时间片的离散静态贝叶斯网络的  $T$  个有向无环图  $G_1, G_2, \dots, G_T$ , 和各个时间片的离散静态贝叶斯网络的条件概率表  $IntraCPT_1, IntraCPT_2, \dots, IntraCPT_T$ ;
- 2) 表示下一个时间片的离散静态贝叶斯网络与上一个时间片的离散静态贝叶斯网络之间的依赖关系的  $T-1$  张状态转移概率表  $InterCPT_1, InterCPT_2, \dots, InterCPT_T$ .

## 1.2 改进的变结构 DDBNs 的推理算法

### 1.2.1 变结构 DDBNs 的推理算法

不失一般性, 假定一个变结构 DDBNs 共有  $T$  个时间片, 第  $t$  个时间片对应的离散静态贝叶斯网络的结构为  $BNT_t$  ( $t = 1, \dots, T$ ), 对应的条件概率表为  $IntraCPT_t$ . 具有  $n_t$  个隐藏节点和  $m_t$  个观测节点, 分别记为  $X_i^t$  ( $t = 1, \dots, T, i = 1, \dots, n_t$ ) 和  $Y_j^t$  ( $t = 1, \dots, T, j = 1, \dots, m_t$ ), 其中, 上标表示变量所在的时间片, 下标表示变量在隐藏变量集合或观测变量集合中的序号. 前后两个时间片间的状态转移概率表为  $InterCPT_t$  ( $t = 1, \dots, T-1$ ), 对这个变结构 DDBNs 进行推理, 本质上就是计算所有的观测变量处于其观测状态下隐藏节点的联合分布

$$\begin{aligned} P(X_1^1 = x_1^1, \dots, X_{n_1}^1 = x_{n_1}^1, \dots, X_1^T = x_1^T, \dots, \\ X_{n_T}^T = x_{n_T}^T | Y_1^1 = y_1^1, \dots, Y_{m_1}^1 = y_{m_1}^1, \dots, \\ Y_{m_T}^T = y_{m_T}^T) \end{aligned} \quad (2)$$

其中,  $\{Y_1^1 = y_1^1, \dots, Y_{m_1}^1 = y_{m_1}^1, \dots, Y_{m_T}^T = y_{m_T}^T\}$  是所有观测变量的某一组合状态.  $\{X_1^1 = x_1^1, \dots, X_{n_1}^1 = x_{n_1}^1, \dots, X_1^T = x_1^T, \dots, X_{n_T}^T = x_{n_T}^T\}$  是隐藏变量的一个组合状态.

根据贝叶斯公式

$$P(X|Y) = \frac{P(X, Y)}{\sum_X P(X, Y)} \quad (3)$$

其中,  $X, Y$  为变量. 可得:

$$\begin{aligned} P(X_1^1 = x_1^1, \dots, X_{n_1}^1 = x_{n_1}^1, \dots, X_{n_T}^T = x_{n_T}^T | Y_1^1 = y_1^1, \dots, Y_{m_1}^1 = y_{m_1}^1, \dots, Y_{m_T}^T = y_{m_T}^T) = \\ \frac{P(X_1^1 = x_1^1, \dots, X_{n_1}^1 = x_{n_1}^1, \dots, X_{n_T}^T = x_{n_T}^T, Y_1^1 = y_1^1, \dots, Y_{m_1}^1 = y_{m_1}^1, \dots, Y_{m_T}^T = y_{m_T}^T)}{\sum_{X_1^1, \dots, X_{n_1}^1, \dots, X_{n_T}^T} P(X_1^1 = x_1^1, \dots, X_{n_1}^1 = x_{n_1}^1, \dots, X_{n_T}^T = x_{n_T}^T, Y_1^1 = y_1^1, \dots, Y_{m_1}^1 = y_{m_1}^1, \dots, Y_{m_T}^T = y_{m_T}^T)} \end{aligned} \quad (4)$$

由贝叶斯网络的条件独立性可知, 所有变量的联合概率分布等于所有条件概率的乘积. 因此有:

$$P(X_1^1 = x_1^1, \dots, X_{n_1}^1 = x_{n_1}^1, \dots, X_{n_T}^T = x_{n_T}^T, Y_1^1 = y_1^1, \dots, Y_{m_1}^1 = y_{m_1}^1, \dots, Y_{m_T}^T = y_{m_T}^T) = \prod_{i,j} P(Y_j^i = y_j^i | Pa(Y_j^i = y_j^i)) \times \prod_{k,l} P(X_l^k = x_l^k | Pa(X_l^k = x_l^k)) \quad (5)$$

其中,  $i = 1, \dots, T, j = 1, \dots, m_i, k = 1, \dots, T, l = 1, \dots, n_k$ . 抽象地说, 一个联合分布是一个多变量函数, 采用消元法, 消去分母中的变量  $\{X_1^1, \dots, X_{n_1}^1, \dots, X_{n_T}^T\}$ , 则有式 (6) (见本页底部). 式 (6) 中,  $i = 1, \dots, T, j = 1, \dots, m_i, k = 1, \dots, T, l = 1, \dots, n_k$ , 且  $i, j, k, l$  都是整数.

式 (6) 的分子是观测变量和隐藏变量处于某一组合状态的联合分布, 分母是通过对所有组合状态的联合概率分布求和, 消去隐藏变量.

上述的变结构 DDBNs 仅能处理硬证据信息, 为了克服该推理算法的不足, 提出了改进的变结构离散动态贝叶斯网路的推理算法.

### 1.2.2 改进的变结构 DDBNs 的推理算法

如果观测到的证据信息是软证据, 用  $y_{l_s}^k$  表示第  $k$  个时间片的第  $l$  个观测节点处于它的第  $s$  个状态. 对软证据信息, 可以借鉴文献 [17-18] 处理多证据的思路, 通过加权求和对多证据进行综合. 因此可对式 (6) 修正为式 (7) (见本页底部). 式中,  $y_{1:m_{i_0}}^t = \{y_{1_0}^t, y_{2_0}^t, \dots, y_{m_{i_0}}^t\}$ ,  $y_{k_0}^t$  表示观测变量  $Y_k^t$  所处的状态,  $i = 1, \dots, T, j = 1, \dots, n_i$ ,

$k = 1, \dots, T, l = 1, \dots, m_k$ , 且  $i, j, k, l$  都是整数.

如果要求某一个隐藏变量的后验概率, 则对式 (7) 进行加和运算:

$$P(X_j^i = x_j^i | y_{1_0}^1, y_{2_0}^1, \dots, y_{m_{1_0}}^1, \dots, y_{1_0}^T, \dots, y_{m_{T_0}}^T) = \sum_{x_{1:n_1}^1, x_{1:n_2}^2, \dots, x_{1:n_T}^T \setminus x_j^i} P(x_1^1, x_2^1, \dots, x_{n_1}^1, \dots, x_1^T, \dots, x_{n_T}^T | y_{1_0}^1, y_{2_0}^1, \dots, y_{m_{1_0}}^1, \dots, y_{1_0}^T, \dots, y_{m_{T_0}}^T) \quad (8)$$

改进的变结构 DDBNs 的推理算法的实现步骤:

**步骤 1.** 求观测变量和隐藏变量处于每一种组合状态的联合概率, 并对联合概率分布进行加权;

**步骤 2.** 对步骤 1 中隐藏变量的组合相同的联合概率求和, 得到式 (7) 的分子;

**步骤 3.** 对步骤 2 中隐藏变量所有组合状态求和, 得到式 (7) 的分母, 即得到归一化因子;

**步骤 4.** 在步骤 2 和步骤 3 的计算结果基础上, 利用式 (7) 就可以计算得到在观测证据条件下隐藏变量的后验概率;

**步骤 5.** 在步骤 4 的基础上, 利用式 (8) 求出每个隐藏变量的后验概率.

### 1.2.3 复杂度分析

假设变结构 DDBNs 的单个时间片上的观测节点不超过  $m$  个, 隐藏节点数不超过  $n$  个, 节点的最大状态数为  $N$ , 共获得了  $T$  个时间片, 则隐藏变量的状态组合数有  $N^{nT}$  种. 若观测证据为硬证据, 则计算隐藏变量的一种组合状态的联合分布需要  $(m+n)T$  次运算, 遍历隐藏变量的所有组合状态需

$$P(X_1^1 = x_1^1, \dots, X_{n_1}^1 = x_{n_1}^1, \dots, X_{n_T}^T = x_{n_T}^T | Y_1^1 = y_1^1, \dots, Y_{m_1}^1 = y_{m_1}^1, \dots, Y_{m_T}^T = y_{m_T}^T) = \frac{\prod_{i,j} P(Y_j^i = y_j^i | Pa(Y_j^i = y_j^i)) \prod_{k,l} P(X_l^k = x_l^k | Pa(X_l^k = x_l^k))}{\sum_{x_1^1, \dots, x_{n_1}^1, \dots, x_{n_T}^T} \prod_{i,j} P(Y_j^i = y_j^i | Pa(Y_j^i = y_j^i)) \prod_{k,l} P(X_l^k = x_l^k | Pa(X_l^k = x_l^k))} \quad (6)$$

$$P(x_1^1, x_2^1, \dots, x_{n_1}^1, \dots, x_1^T, \dots, x_{n_T}^T | y_{1_0}^1, y_{2_0}^1, \dots, y_{m_{1_0}}^1, \dots, y_{1_0}^T, \dots, y_{m_{T_0}}^T) = \frac{P(x_{1:n_1}^1, \dots, x_{1:n_T}^T, y_{1:m_{1_0}}^1, \dots, y_{1:m_{T_0}}^T)}{\sum_{x_{1:n_1}^1, x_{1:n_2}^2, \dots, x_{1:n_T}^T} P(x_{1:n_1}^1, \dots, x_{1:n_T}^T, y_{1:m_{1_0}}^1, \dots, y_{1:m_{T_0}}^T)} = \frac{\prod_{i,j} P(x_j^i | Pa(X_j^i)) \prod_{k,l} [P(Y_l^k = y_{l_s}^k | Pa(Y_l^k)) P(Y_l^k = y_{l_s}^k)]}{\sum_{x_{1:n_1}^1, x_{1:n_2}^2, \dots, x_{1:n_T}^T} \prod_{i,j} P(x_j^i | Pa(X_j^i)) \prod_{k,l} [P(Y_l^k = y_{l_s}^k | Pa(Y_l^k)) P(Y_l^k = y_{l_s}^k)]} \quad (7)$$

要  $(m+n)TN^{nT}$  次运算, 则式 (6) 的复杂度为  $O((n+m)TN^{nT})$ . 若观测证据为软证据, 则计算式 (7) 分子的一个联合分布需要  $(n+2m)TN^{mT}$  次运算, 遍历隐藏变量的所有状态需要  $(n+2m)TN^{(m+n)T}$  次运算, 则式 (7) 的复杂度为  $O((n+2m)TN^{(n+m)T})$ . 由式 (6) 与式 (7) 的复杂度对比可知, 对于同一个网络, 处理多状态证据比处理单状态证据的复杂度要高得多, 因此降低算法的复杂度就显得非常必要. 从式 (7) 的复杂度分析可知, 算法的复杂度主要是由计算分子的复杂度决定的, 如果能够降低分子计算的复杂度, 就能够降低推理算法的复杂度. 为了克服改进的变结构 DDBNs 算法计算量大的缺陷, 本文在对该推理算法的分子计算方式分析的基础上, 提出了变结构 DDBNs 的快速推理算法.

## 2 变结构 DDBNs 的快速推理算法

### 2.1 算法

处理硬证据的变结构 DDBNs 推理算法利用链式乘积规则和条件独立性, 将联合概率分解为一系列参数化的条件概率的乘积. 而对于能处理软证据的改进的变结构 DDBNs 的推理算法, 增加了期望求和的过程, 这样就显著增大了算法的计算量. 为了减少算法的计算量, 需对式 (7) 计算方式进行改变, 即交换求和与乘积的次序, 则有式 (9) (见本页底部). 式中,  $S_l^k$  表示第  $k$  个时间片上的第  $l$  个观测节点共有  $S_l^k$  个状态,  $i = 1, \dots, T, j = 1, \dots, n_i, k = 1, \dots, T, l = 1, \dots, m_k$ , 且  $i, j, k, l$  都是整数.

利用式 (9) 的计算结果, 就可以根据式 (8) 求出某一个隐藏变量的后验概率.

变结构 DDBNs 快速算法的实现步骤:

**步骤 1.** 在软证据条件下, 计算联合概率分布;

**步骤 2.** 对步骤 1 的所有计算结果求和, 得到归一化因子;

**步骤 3.** 利用式 (9), 计算隐藏变量的后验概率;

**步骤 4.** 把式 (9) 的计算结果代入式 (8) 中, 就可以计算出每个隐藏变量的后验概率.

### 2.2 复杂度分析

假设变结构 DDBNs 的一个时间片最多有  $n$  个隐藏节点、 $m$  个观测节点, 节点的最大状态数为  $N$ , 共观测了  $T$  个时间片. 计算式 (9) 分子的一个联合概率分布需要  $(2Nm+n)T$  次运算, 遍历隐藏变量的所有状态需要  $(2Nm+n)TN^{nT}$  次运算, 则式 (9) 的复杂度为  $O((2Nm+n)TN^{nT})$ .

由式 (7) 与式 (9) 的复杂度对比可知, 指数由  $(n+m)T$  变为  $nT$ , 计算量减少了. 对于较复杂的变结构 DDBNs,  $(n+m)T$  要远远大于  $nT$ , 计算量减少会更加显著, 故而提出的变结构 DDBNs 的快速推理算法是非常有效的.

变结构 DDBNs 的快速推理算法与改进的变结构 DDBNs 的推理算法本质区别在于: 快速推理算法把联合分布进行了分解, 使得运算局部化, 因此, 可以降低推理的复杂度.

从以上分析可知, 变结构 DDBNs 的快速推理算法的优点是: 1) 能实现软证据条件下的变结构 DDBNs 的推理; 2) 提高推理效率.

## 3 缺失数据动态贝叶斯网络

在动态贝叶斯网络的实际应用中, 观测数据由于各种原因会出现数据缺失现象, 例如在某些时间片上的证据信息观测不到、剔除了观测数据中的奇异值、数据发生了随机缺失等. 如果仅利用现有数据仍能进行推理, 因为贝叶斯网络反映的是整个数

$$\begin{aligned}
 & P(x_1^1, x_2^1, \dots, x_{n_1}^1, \dots, x_1^T, \dots, x_{n_T}^T | y_{1o}^1, y_{2o}^1, \dots, y_{m_1o}^1, \dots, y_{1o}^T, \dots, y_{m_To}^T) = \\
 & \frac{P(x_{1:n_1}^1, \dots, x_{1:n_T}^T, y_{1:m_1o}^1, \dots, y_{1:m_To}^T)}{\sum_{x_{1:n_1}^1, x_{1:n_2}^2, \dots, x_{1:n_T}^T} P(x_{1:n_1}^1, \dots, x_{1:n_T}^T, y_{1:m_1o}^1, \dots, y_{1:m_To}^T)} = \\
 & \frac{\prod_{i,j} P(X_j^i = x_j^i | Pa(X_j^i)) \prod_{k,l} \prod_{p=1}^{s_l^k} P(Y_l^k = y_{lp}^k | Pa(Y_l^k)) P(Y_l^k = y_{lp}^k)}{\sum_{x_{1:n_1}^1, x_{1:n_2}^2, \dots, x_{1:n_T}^T} \prod_{i,j} P(X_j^i = x_j^i | Pa(X_j^i)) \prod_{k,l} \prod_{p=1}^{s_l^k} P(Y_l^k = y_{lp}^k | Pa(Y_l^k)) P(Y_l^k = y_{lp}^k)} \quad (9)
 \end{aligned}$$

据域中数据间的概率关系, 缺失数据的节点就不会对网络的推理起作用, 相当于这个节点在该时间片上不存在, 因而网络的结构就会发生改变, 如图 3 所示. 这种变结构动态贝叶斯网络是变结构动态贝叶斯网络的一个特例, 下面我们从定义出发, 详述这种变结构动态贝叶斯网络的模型及其特点.

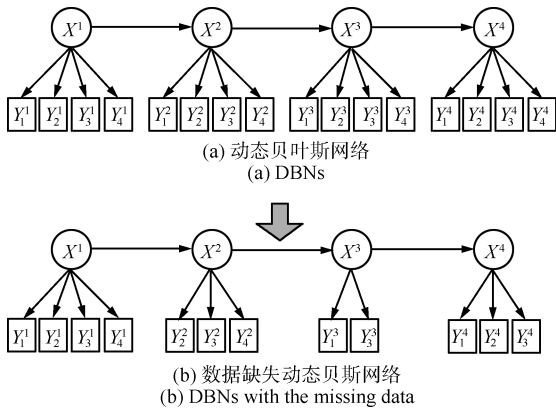


图 3 数据缺失动态贝叶斯网络模型

Fig. 3 The DBN model with the missing data

对于一个动态贝叶斯网络, 如果它在某些时间片上因某些观测变量的证据信息的缺失造成网络结构发生变化, 且该网络仍具有如下特征:

- 1) 相邻时间片变量之间的依赖关系保持不变;
- 2) 对于各时间片上存在的变量状态和变量间的依赖关系保持稳定;
- 3) 网络的传感器模型和转移模型的条件分布保持不变. 则把具有以上特点的变结构动态贝叶斯网络称为数据缺失动态贝叶斯网络.

对于数据缺失动态贝叶斯网络数据来说, 数据缺失的节点仅仅是未能获取到对应属性的观测证据, 而非对应属性的消失. 数据缺失 DDBNs 描述的是环境状态的变化是由一个稳态过程引起的, 而不是非稳态过程引起的, 而稳态过程意味着变化的过程是由本身不随时间变化的规律支配的, 那么网络的参数就不会随时间的变化而变化. 从这个角度上说, 这种变结构动态贝叶斯网络不是真正意义上的变结构, 仅仅是具有变结构的形式, 而真正意义上的变结构反映的是非稳态过程, 所以数据缺失动态贝叶斯网络的参数不会因为网络的数据缺失而发生变化. 由于数据缺失动态贝叶斯网络在形式上具有变结构动态贝叶斯网络的特点, 所以本文把这类动态贝叶斯网络归为变结构动态贝叶斯网络. 如果这类网络的数据缺失属于随机缺失, 那么可以用数据修补的方法把缺失数据补齐<sup>[19-20]</sup>, 这样变结

构动态贝叶斯网络就恢复成传统的动态贝叶斯网络.

在观测数据缺失率增大时, 单个时间片可利用的证据信息减少, 会直接影响该时间片的推理结果, 使得该时间片推理结果的置信度降低, 同时, 由于变结构动态贝叶斯网络的各个时间片之间也是相互影响的, 各个时间片之间的相互影响减弱. 因此, 随着观测数据缺失率的增大, 会对变结构动态贝叶斯网络的推理结果产生显著影响.

### 4 仿真实验

#### 4.1 构建变结构离散动态贝叶斯网络模型及参数设定

图 4 是识别空中飞机的单个时间片的贝叶斯网络, 目标类型为隐藏节点, 它分为 4 个状态: 战斗机、轰炸机、电子战飞机和预警机. 目标特征是观测节点, 主要包括飞行速度 ( $Y_1$ )、雷达波束 ( $Y_2$ )、高度 ( $Y_3$ ) 和目标雷达反射截面 (Radar cross-section, RCS) ( $Y_4$ ), 雷达波束 ( $Y_5$ ). 按照时间轴展开, 就可以构建离散动态贝叶斯网络模型, 如果数据不发生缺失, 构建的动态贝叶斯网络模型如图 3(a) 所示; 如果观测数据发生缺失, 构建的变结构动态贝叶斯网络模型如图 3(b) 所示. 模型的参数设定是基于武器专家知识及情报建立起来的, 其中, 状态转移概率如表 1 所示, 条件概率如表 2 所示, 目标类型的先验概率为  $P(X = F, B, J, A) = (0.40, 0.40, 0.15, 0.05)$ . 由于识别空中飞机类型的过程是一个稳态过程, 因此状态转移概率和条件概率在识别过程中保持不变. 在仿真实验中, 设定有 1 架轰炸机在初始时刻匀速突防飞行, 在发现我方拦截飞机后, 轰炸机转弯并加速逃跑, 共取得 5 个时间片的观测值, 观测结果如表 3 所示. 在第 4 和第 5 个时间片上, 由于轰炸机的雷达关机, 故观测不到雷达信号, 数据缺失了 16%.

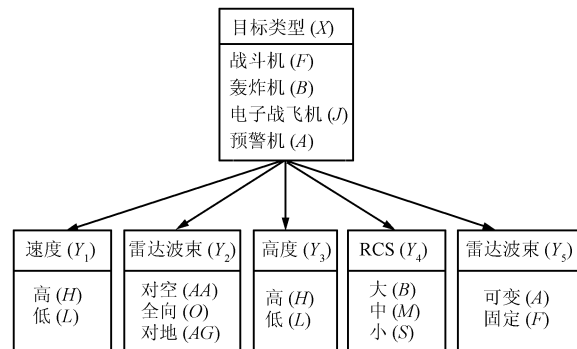


图 4 单个时间片贝叶斯网络

Fig. 4 Bayesian networks on a single time slice

表 1 状态转移概率

Table 1 Stable transition probabilities

$X_i \setminus X_{i+1}$	$F$	$B$	$J$	$A$
$F$	0.70	0.10	0.10	0.10
$B$	0.15	0.65	0.10	0.10
$J$	0.10	0.10	0.70	0.10
$A$	0.05	0.10	0.15	0.70

表 2 条件概率

Table 2 Conditional probability

$X$	$P(Y_1 X)$	$P(Y_2 X)$	$P(Y_3 X)$	$P(Y_4 X)$	$P(Y_5 X)$
	$H, L$	$AA, O, AG$	$H, L$	$B, M, S$	$A, F$
$F$	0.80,0.20	0.90,0.05,0.05	0.70,0.30	0.00,0.60,0.40	0.10,0.90
$B$	0.30,0.70	0.10,0.10,0.80	0.30,0.70	0.65,0.30,0.05	0.10,0.90
$J$	0.30,0.70	0.15,0.70,0.15	0.50,0.50	0.30,0.55,0.15	0.10,0.90
$A$	0.30,0.70	0.10,0.80,0.10	0.40,0.60	0.80,0.10,0.10	0.90,0.10

4.2 仿真实验结果及分析

根据表 3 的观测数据, 利用改进的变结构 DDBNs 的推理算法与变结构 DDBNs 的快速推理算法分别得到表 4 的左列推理结果和右列推理结果. 图 5 给出了改进的变结构 DDBNs 的推理算法和变结构 DDBNs 的快速推理算法在变结构网络的时间片数不同情况下的运行时间对比图.

表 3 观测数据 (数据缺失率为 16%)

Table 3 Observation data (Data loss rate: 16%)

时间片	$Y_1$	$Y_2$	$Y_3$	$Y_4$	$Y_5$
1	0.30,0.70	0.20,0.30,0.50	0.40,0.60	0.70,0.30,0.00	0.15,0.85
2	0.25,0.75	0.25,0.30,0.45	0.35,0.65	0.75,0.25,0.00	0.10,0.90
3	0.50,0.50	0.25,0.25,0.50	0.35,0.65	0.80,0.20,0.00	0.15,0.85
4	0.35,0.65	—	0.30,0.70	0.65,0.35,0.00	—
5	0.30,0.70	—	0.25,0.75	0.60,0.40,0.00	—

根据表 3 的观测数据条件, 表 5 给出了静态贝叶斯网络与变结构 DDBNs 的推理结果对比, 根据表 6 的观测数据 (数据缺失率为 28%), 表 7 给出了静态贝叶斯网络与变结构 DDBNs 的推理结果对比, 根据表 8 的观测数据 (数据缺失率为 40%), 表 9 给出了静态贝叶斯网络与变结构 DDBNs 的推理结果对比.

表 4 两种变结构 DDBNs 推理算法的推理结果对比 (数据缺失率为 16%)

Table 4 Comparison of inference results given by the two structure-variable DDBNs inference algorithms (Data loss rate: 16%)

时间片	改进的变结构 DDBNs 推理算法的推理结果: $F, B, J, A$ (%)	变结构 DDBNs 的快速推理算法的推理结果: $F, B, J, A$ (%)
1	1.61, 96.6, 1.22, 0.56	1.61, 96.6, 1.22, 0.56
2	0.56, 97.3, 0.86, 1.25	0.56, 97.3, 0.86, 1.25
3	0.92, 93.7, 1.64, 3.72	0.92, 93.7, 1.64, 3.72
4	2.0, 75.6, 7.31, 15.1	2.0, 75.6, 7.31, 15.1
5	3.89, 61.9, 12.9, 21.3	3.89, 61.9, 12.9, 21.3

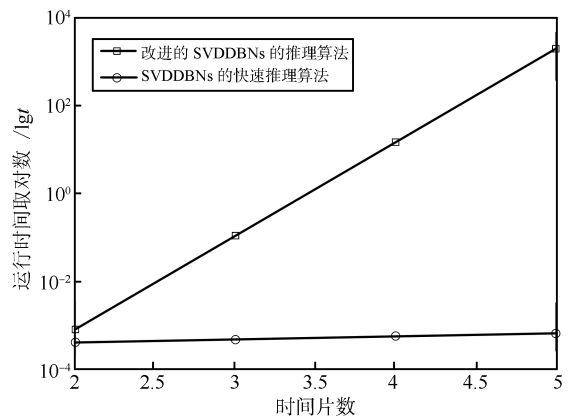


图 5 改进的变结构 DDBNs 的推理算法与变结构 DDBNs 的快速推理算法的运行时间对比

Fig. 5 Comparison of time costs of the improved structure-variable DDBNs inference algorithm and structure-variable DDBNs fast inference algorithm

从表 4 的推理结果可以看出, 改进的变结构 DDBNs 的推理算法与变结构 DDBNs (Structure-variable DDBNs, SVDDBNs) 的快速推理算法的推理结果在各个时间片上是相同的, 因此, 尽管这两种算法的推理机制有所不同, 但这两种算法的推理结果是等价的.

图 5 的纵坐标是对运行时间取对数, 从曲线走势来看, 两种算法的运行时间的对数与时间片数近似成线性关系. 改进的变结构 DDBNs 推理算法的斜率比较大, 而变结构 DDBNs 的快速推理算法的斜率非常小. 因此, 采用变结构 DDBNs 的快速推理算法能显著减少算法的运行时间, 提高了推理效率.

在表 5、表 7、表 9 的两列数据中, 每一组数据从左到右依次为战斗机、轰炸机、电子战飞机和预

警机的概率值, 我们取这 4 个概率值中的最大值所对应的飞机类型作为贝叶斯网络在该时间片上识别出的目标。

通过对表 5、表 7、表 9 左列与右列的数据对比可知, 尽管静态贝叶斯网络和变结构 DDBNs 都能正确识别出目标类型是轰炸机, 但变结构 DDBNs 比静态贝叶斯网络在对应的时间片上识别出目标的概率更高。

根据表 5、表 7、表 9 中的左列数据, 图 6 给出了数据缺失率分别为 16%、28%、40%的情况下, 静态贝叶斯网络在 5 个时间片上识别出目标为轰炸机的概率对比图。

根据表 5、表 7、表 9 中的右列数据, 图 7 给出了数据缺失率分别为 16%、28%、40%的情况下, 变结构 DDBNs 在 5 个时间片上识别出目标为轰炸机的概率对比图。

表 5 静态贝叶斯网络与变结构 DDBNs 的推理结果对比 (数据缺失率为 16%)

Table 5 Comparison of the inference results of static Bayesian networks and structure-variable DDBNs (Data loss rate: 16%)

时间片	静态贝叶斯网络的推理结果: $F, B, J, A$ (%)	变结构 DDBNs 的推理结果: $F, B, J, A$ (%)
1	6.18, 68.6, 11.4, 13.9	1.61, 96.6, 1.22, 0.56
2	5.57, 72.0, 9.98, 12.4	0.56, 97.3, 0.86, 1.25
3	6.73, 71.0, 9.09, 13.2	0.92, 93.7, 1.64, 3.72
4	7.47, 35.4, 22.4, 34.7	2.0, 75.6, 7.31, 15.1
5	7.36, 35.8, 23.4, 33.5	3.89, 61.9, 12.9, 21.3

表 6 观测数据 (数据缺失率为 28%)

Table 6 Observation data (Data loss rate: 28%)

时间片	$Y_1$	$Y_2$	$Y_3$	$Y_4$	$Y_5$
1	0.30,0.70	—	0.40,0.60	0.70,0.30,0.00	0.15,0.85
2	0.25,0.75	0.25,0.30,0.45	—	0.75,0.25,0.00	0.10,0.90
3	0.50,0.50	0.25,0.25,0.50	0.35,0.65	—	0.15,0.85
4	0.35,0.65	—	0.30,0.70	0.65,0.35,0.00	—
5	0.30,0.70	—	0.25,0.75	0.60,0.40,0.00	—

从图 6 的第 4、5 时间片可以看出, 在时间片的数据缺失不变情况下, 识别出目标的概率是保持不变的, 这是因为静态贝叶斯网络的各个时间片之间是互不影响的。随着观测数据缺失的增大, 识别结果的置信度越来越低。在数据缺失达到 40%时, 在 4、5 时间片仅能勉强识别出目标。

表 7 静态贝叶斯网络与变结构 DDBNs 的推理结果对比 (数据缺失率为 28%)

Table 7 Comparison of the inference results of static Bayesian networks and structure-variable DDBNs (Data loss rate: 28%)

时间片	静态贝叶斯网络的推理结果: $F, B, J, A$ (%)	变结构 DDBNs 的推理结果: $F, B, J, A$ (%)
1	10.7, 58.3, 13.8, 17.1	3.55, 93.8, 1.83, 0.80
2	6.85, 69.7, 10.8, 12.7	6.85, 69.7, 10.8, 12.7
3	25.0, 54.5, 11.6, 8.90	4.53, 89.4, 2.77, 3.27
4	7.47, 35.4, 22.4, 34.7	2.73, 73.5, 8.28, 15.5
5	7.36, 35.8, 23.4, 33.5	4.11, 60.6, 13.6, 21.7

表 8 观测数据 (数据缺失率为 40%)

Table 8 Observation data (Data loss rate: 40%)

时间片	$Y_1$	$Y_2$	$Y_3$	$Y_4$	$Y_5$
1	0.30,0.70	—	0.40,0.60	—	0.15,0.85
2	—	0.25,0.30,0.45	—	0.75,0.25,0.00	0.10,0.90
3	0.50,0.50	0.25,0.25,0.50	—	—	0.15,0.85
4	0.35,0.65	—	0.30,0.70	0.65,0.35,0.00	—
5	0.30,0.70	—	0.25,0.75	0.60,0.40,0.00	—

表 9 静态贝叶斯网络与变结构 DDBNs 的推理结果对比 (数据缺失率为 40%)

Table 9 Comparison of the inference results of static Bayesian networks and structure-variable DDBNs (Data loss rate: 40%)

时间片	静态贝叶斯网络的推理结果: $F, B, J, A$ (%)	变结构 DDBNs 的推理结果: $F, B, J, A$ (%)
1	25.7, 46.0, 15.8, 12.5	11.4, 85.4, 2.50, 0.69
2	11.2, 66.4, 10.3, 12.1	4.90, 91.7, 1.77, 1.65
3	29.3, 50.2, 11.9, 8.65	7.29, 85.8, 3.31, 3.63
4	7.47, 35.4, 22.4, 34.7	3.28, 71.5, 8.85, 16.3
5	7.36, 35.8, 23.4, 33.5	4.26, 59.4, 14.0, 22.3

从图 7 的曲线走势可以看出, 随着观测数据缺失率的不断增大, 变结构 DDBNs 在 5 个时间片上识别出目标概率不断降低。这是因为变结构 DDBNs 不仅可以利用单个时间片上的证据信息, 而且能够利用其他时间片上的证据信息, 各个时间片之间是相互影响的。随着观测数据缺失率的增大, 变结构 DDBNs 可利用的证据信息越来越少, 因



此, 在对应的时间片上, 识别出目标的概率也越来越低.

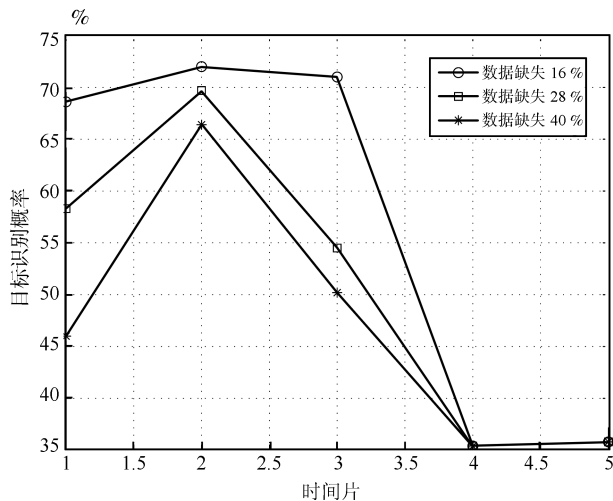


图 6 在数据缺失分别为 16%、28%、40% 时, 静态贝叶斯网络识别出目标的概率对比

Fig. 6 Comparison of the target recognition probabilities given by static Bayesian networks with data loss rates 16%, 28%, and 40%, respectively

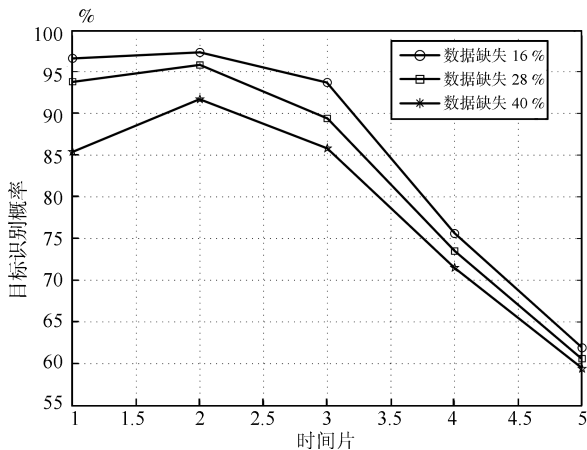


图 7 在数据缺失分别为 16%、28%、40% 时, 变结构 DDBNs 识别出目标的概率对比

Fig. 7 Comparison of the target recognition probabilities given by structure-variable DDBNs with data loss rates 16%, 28%, and 40%, respectively

## 5 结论

变结构动态贝叶斯网络是广义的动态贝叶斯网络. 在这个意义上, 目前研究的传统动态贝叶斯网络是变结构动态贝叶斯网络的特例. 变结构动态贝叶斯网络不仅允许网络的参数发生变化, 还允许变量的数目和网络结构发生变化. 因此, 有了变结构动态贝叶斯网络这一工具, 就可以对各种突变过程进

行建模和推理. 本文从变结构动态贝叶斯网络定义、性质、对算法适用范围的拓展、提高算法的计算效率及数据缺失动态贝叶斯网络等方面对变结构动态贝叶斯网络的机制进行了研究.

## References

- 1 Jensen F V. *An Introduction to Bayesian Networks*. New York: Springer, 1996
- 2 Jensen F V. *Bayesian Networks and Decision Graphs*. New York: Springer, 2001
- 3 Dean T, Kanazawa K. A model for reasoning about persistence and causation. *Computational Intelligence*, 1989, **5**(3): 142–150
- 4 Russell S, Norving P. *Artificial Intelligence: A Modern Approach (Second Edition)*. New Jersey: Prentice Hall, 2003. 559–580
- 5 Suandi S A, Enokida S, Ejima T. Face pose estimation from video sequence using dynamic Bayesian network. In: *Proceedings of the IEEE Workshop on Motion and Video Computing*. Copper Mountain, USA: IEEE, 2008. 1–8
- 6 Han P X, Mu R J, Cui N G. Active and dynamic multi-sensor information fusion method based on dynamic Bayesian networks. In: *Proceedings of the International Conference on Mechatronics and Automation*. Changchun, China: IEEE, 2009. 3076–3080
- 7 Du You-Tian, Chen Feng, Xu Wen-Li. Approach to human activity multi-scale analysis and recognition based on multi-layer dynamic Bayesian network. *Acta Automatica Sinica*, 2009, **35**(3): 225–232  
(杜友田, 陈峰, 徐文立. 基于多层动态贝叶斯网络的人的行为多尺度分析及识别方法. *自动化学报*, 2009, **35**(3): 225–232)
- 8 Peter Hearty, Norman Fenton, David Marquez, Martin Neil. Predicting project velocity in XP using a learning dynamic Bayesian network model. *IEEE Transactions on Software Engineering*, 2009, **35**(1): 124–137
- 9 Jin Nai-Gao, Yin Fu-Liang, Chen Zhe. Audio-visual speaker tracking based on dynamic Bayesian network. *Acta Automatica Sinica*, 2008, **34**(9): 1083–1089  
(金乃高, 殷福亮, 陈喆. 基于动态贝叶斯网络的音视频联合说话人跟踪. *自动化学报*, 2008, **34**(9): 1083–1089)
- 10 Ghahramani Z. An introduction to hidden Markov models and Bayesian networks. *International Journal of Pattern Recognition and Artificial Intelligence*, 2001, **15**(1): 9–42
- 11 Murphy K. *Dynamic Bayesian Networks: Representation, Inference and Learning* [Ph.D. dissertation], University of California, USA, 2002
- 12 Kalman R E. A new approach to linear filtering and prediction problems. *Transactions of the ASME-Journal of Basic Engineering*, 1960, **82**(Series D): 35–45
- 13 Roweis S, Ghahramani Z. A unifying review of linear Gaussian models. *Neural Computation*, **11**(2): 305–345
- 14 Zhang Lian-Wen, Guo Hai-Peng. *Introduction to Bayesian Networks*. Beijing: Science Press, 2006. 62–64  
(张连文, 郭海鹏. 贝叶斯网引论. 北京: 科学出版社, 2006. 62–64)

- 15 Chen Hai-Yang, Gao Xiao-Guang, Fan Hao. Inference algorithm of variable structure DDBNs and multi-target recognition. *Acta Aeronautica et Astronautica Sinica*, 2010, **31**(11): 2222–2227  
(陈海洋, 高晓光, 樊昊. 变结构 DDBNs 的推理算法与多目标识别. 航空学报, 2010, **31**(11): 2222–2227)
- 16 Gao Xiao-Guang, Shi Jian-Guo. Structure varied discrete dynamic Bayesian network and its inference algorithm. *Journal of Systems Engineering*, 2007, **22**(1): 9–14  
(高晓光, 史建国. 变结构离散动态贝叶斯网络及其推理算法. 系统工程学报, 2007, **22**(1): 9–14)
- 17 Chen H Y, Gao X G. Ship recognition based on improved forwards-backwards algorithm. In: Proceedings of the 6th International Conference on Fuzzy Systems and Knowledge Discovery. Tianjin, China: IEEE, 2009. 509–513
- 18 Shi Jian-Guo, Gao Xiao-Guang. Direct calculation inference algorithm for discrete dynamic Bayesian network. *Systems Engineering and Electronics*, 2005, **27**(9): 1626–1630  
(史建国, 高晓光. 离散动态贝叶斯网络的直接计算推理算法. 系统工程与电子技术, 2005, **27**(9): 1626–1630)
- 19 Chen H Y, Gao X G, Zheng J S. A kind of data repairing for missing data of discrete dynamic Bayesian networks. In: Proceedings of the 5th International Conference on Natural Computation. Tianjin, China: IEEE, 2009. 47–51
- 20 Chen H Y, Gao X G. Forwards-backwards information repairing algorithm and appliance on discrete dynamic Bayesian networks. In: Proceedings of the International Conference on Intelligent Human-Machine Systems and Cybernetics. Hangzhou, China: IEEE, 2009. 76–80



高晓光 西北工业大学电子信息学院教授. 主要研究方向为计算智能, 航空火力控制. E-mail: xggao@nwpu.edu.cn  
(**GAO Xiao-Guang** Professor at the School of Electronics and Information, Northwestern Polytechnical University. Her research interest covers computational intelligence and airborne fire control.)



陈海洋 博士, 西安工程大学电子信息学院讲师. 主要研究方向为先进控制理论及应用. 本文通信作者.  
E-mail: chy\_00@163.com  
(**CHEN Hai-Yang** Ph. D., lecturer at the School of Electronic Information, Xi'an Polytechnic University. His research interest covers advanced control theory and application. Corresponding author of this paper.)



史建国 海军航空工程学院 7 系副教授. 主要研究方向为先进火力控制理论, 智能火力控制技术.  
E-mail: cheeryapple@163.com  
(**SHI Jian-Guo** Associate professor in the No. 7 Department of Navy Aeronautical and Astronautical University. His research interest covers advanced theory of fire control and theory of intelligent fire control.)