

# 目标数据缺失下离散动态贝叶斯网络的参数学习

任 佳, 高晓光, 茹 伟

(西北工业大学电子信息学院, 陕西 西安 710129)

**摘 要:** 离散动态贝叶斯网络参数学习的难点在于: 隐藏节点的片间转移概率获得及观测数据发生不同程度缺失。针对上述问题, 提出基于目标缺失数据估计的前向递归参数学习算法。该算法利用离散动态贝叶斯网络中各观测变量与隐藏变量之间的对应关系, 采用支持向量机建立观测变量间的非线性函数关系完成缺失数据估计, 此基础上利用完整数据集和前向递归算法完成片内和片间参数更新。以空中目标识别为仿真背景, 通过与期望最大算法对比, 验证了该算法的学习效率和精度两个方面的优势。

**关键词:** 参数学习; 离散动态贝叶斯网络; 数据缺失; 前向递归

**中图分类号:** TP 181

**文献标志码:** A

**DOI:** 10.3969/j.issn.1001-506X.2011.08.39

## Parameter learning of discrete dynamic Bayesian network with missing target data

REN Jia, GAO Xiao-guang, RU Wei

(Department of Electronic Engineering, Northwestern Polytechnical University, Xi'an 710129, China)

**Abstract:** The difficulty of discrete dynamic Bayesian network parameter learning lies in: obtaining the transition probability of hidden nodes between slices, lack of observational data in varying degrees. Focusing on the above problems, the forward recursive parameters learning algorithm based on target data missing estimation is proposed. The algorithm uses the correspondent relation between the observed variables and hidden variables in discrete dynamic Bayesian network, using support vector machine to establish a nonlinear function between observed variables for completing the missing data estimation. A complete data set and the forward recursive algorithm are applied to complete parameters updating in inter-slice and in-slice. On the background of aerial target recognition, the advantages of the proposed method at efficiency and accuracy are illustrated compared with the expectative maximization method.

**Keywords:** parameter learning; discrete dynamic Bayesian network; data missing; forward recursion

## 0 引 言

动态贝叶斯网络(dynamic Bayesian networks, DBN)是以概率网络为基础, 将静态网络结构与时间信息相结合, 形成具有处理时序数据能力的随机模型。目前 DBN 理论研究主要集中在网络时变特性<sup>[1]</sup>、约束条件下学习<sup>[2]</sup>、数据缺失下的推理方法<sup>[3]</sup>等几个方面。对于 DBN 的应用主要是根据背景或想定, 学习合理的网络结构, 由专家给出相关网络参数, 通过概率推理得到评估或决策结果。其在基因表达<sup>[4]</sup>、语音识别<sup>[5]</sup>、软件风险测试<sup>[6]</sup>和图像处理<sup>[7]</sup>等方面已得到应用。但在上述应用中, 网络参数往往由领域专家给出, 忽视了参数学习过程, 因此无法保证推理结果的合理性与准确性。

目前, 贝叶斯网络参数学习主要在静态网络模型中展开研究。静态网络模型中应用最广泛的参数学习方法是最大似然估计, 它是通过对一类样本数据出现频次的统计来获得网络参数。受此思想的启发, 文献[8]通过模型似然函数求解梯度上升似然值来估计关系贝叶斯网络参数。文献[9]利用不同的领域知识约束参数估计过程, 以提高 BN 参数精度。文献[10]通过约束条件将优化变量限定在凸区域里, 产生一个优化问题(这个优化问题存在先验参数约束、后验参数约束、狄利克雷约束)使得每一个先验参数和各自的后验参数相关联, 在样本数据量较小的情况下能够获得精度较高的网络参数。文献[11-12]针对静态网络参数学习开展参数平滑和变结构下参数自适应研究, 拓展了静态网络参数学习方法, 为动态网络参数学习提供了借鉴。

收稿日期: 2010-09-22; 修回日期: 2011-04-02。

基金项目: 国家自然科学基金(60774064)资助课题

作者简介: 任佳(1981-), 男, 博士研究生, 主要研究方向为人工智能与机器学习。E-mail: wwwrenjia@gmail.com

基于样本统计的静态网络参数学习方法在动态系统中并不完全适用。主要的难点在于隐藏变量片间转移概率的获取,尤其在观测数据缺失情况下学习网络参数更加困难。文献[13]将期望最大(expectation maximization, EM)算法与模型选择相结合,实现了贝叶斯网络中缺失项估计和静态结构学习。但在数据缺失条件下 EM 算法效率不高,计算量随缺失数据集大小线性增加。通过引入贝叶斯主要成分分析归类(Bayesian principal component analysis imputation, BPCAI)算法、LSI (least square impute)算法、K 邻域(K-nearest neighbour)等方法可以显著提高参数估计收敛速度,但存在较严重的估计误差问题<sup>[14-15]</sup>。

针对离散动态贝叶斯网络(discrete dynamic Bayesian network, DDBN)参数学习问题,需要根据 DDBN 的特点将静态系统中成熟参数学习算法进行扩展。本文通过构建前向递归(forward recursion, FR)算法来获得 DDBN 片间和片内的转移概率,利用 DDBN 中隐藏变量与观测变量间的关联性,在数据缺失的条件下采用支持向量机(support vector machines, SVM)算法建立观测数据间的非线性函数关系,完成对缺失数据集的修补,为 FR 算法提供完整的数据集。

## 1 离散动态贝叶斯网络参数学习

### 1.1 前向递归参数学习算法

DDBN 参数学习的目的是得到最有可能的网络参数使观测样本出现概率最大。在拥有完整数据集的情况下,可以将已有的前向算法应用到含有多个节点的离散动态网络参数学习当中。

首先,建立一个包含隐藏节点的 DDBN 模型如图 1 所示。为了简化推导和便于叙述,这里只给出包含单个隐藏节点的 DDBN 模型。隐藏节点用无色的方形表示,观测节点用灰色的方形表示。用  $X_t, Y_t$  表示  $t$  时刻隐藏序列和观测序列。 $x_t$  表示  $t$  时刻隐藏变量的值, $x_t$  存在  $n$  种状态,即  $x_t \in \{1, 2, 3, \dots, n\}$ 。 $y_t^k$  表示  $t$  时刻第  $k$  个离散观测变量的观测值。

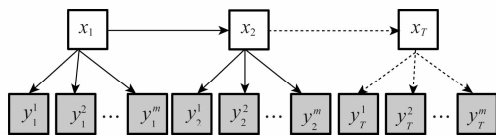


图 1 含有隐藏节点的 DDBN 示意图

其次,对 DDBN 模型参数进行定义。DDBN 参数主要包括三部分:

(1) 初始隐状态分布矩阵,即先验信息  $\pi = (\pi_i)_{1 \times n}$ ,其中  $\pi(i) = P(x_1 = i), i = 1, 2, \dots, n$ ;

(2) 隐变量间转移矩阵  $A = (a_{ij}^t)_{n \times n}$ ,其中  $a_{ij}^t = P(x_t = j | x_{t-1} = i)$ ;

(3) 隐藏变量与观测变量转移矩阵  $B$ ,若将模型输出参数定义为  $P(Y_t | x_t)$ ,则一个有限的参数集可以用矩阵  $B =$

$(b_i^k)_{1 \times n}$ 来表示,即  $b_i^k(k) = P(y_t^k = c | x_t = i)$ , $b_i^k(k)$ 表示  $t$  时刻隐藏变量状态为  $i$  时与第  $k$  个观测节点间的转移概率。

下面给出 FR 参数学习算法,算法流程如图 2 所示。

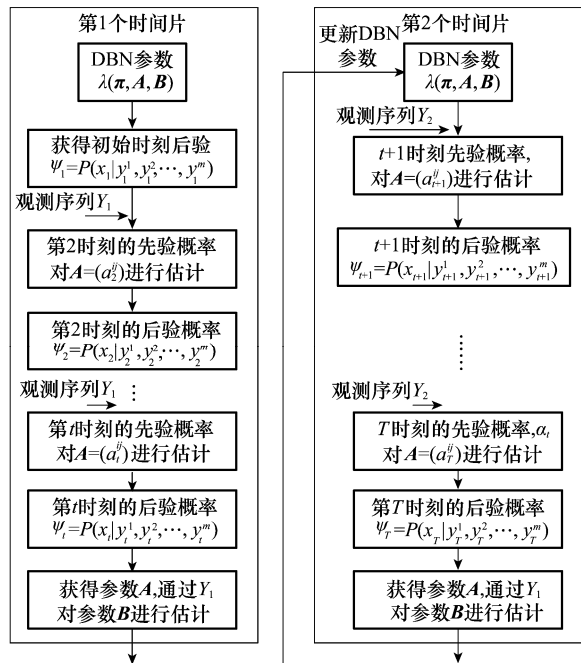


图 2 参数学习流程

(1) 根据初始时刻的先验信息  $\pi$  及该时刻的观测序列  $Y_1$  获得初始时刻的后验概率  $\psi_1$ :

$$\psi_1 = P(x_1 | y_1^1, \dots, y_1^m) =$$

$$\frac{\pi \prod_{k=1}^m P(y_1^k | x_1)}{\sum_{x_1} \pi \prod_{k=1}^m P(y_1^k | x_1)} = \frac{\pi \prod_{k=1}^m b_i^k(k)}{\sum_{x_1} \pi \prod_{k=1}^m b_i^k(k)} \quad (1)$$

(2) 依据前向推理算法的思想,利用当前时刻观测信息对下一时刻隐藏变量的状态进行估计,定义前向传播算子为  $\alpha$ ,在  $t-1$  时刻观测信息下, $t$  时刻隐藏变量状态  $j$  的概率为

$$\alpha_t(j) = P(x_t = j | y_{t-1}^1, \dots, y_{t-1}^m) = \sum_{i=1}^n P(x_t = j | x_{t-1} = i) \cdot$$

$$P(x_{t-1} = i | y_{t-1}^1, \dots, y_{t-1}^m) = \sum_{i=1}^n \psi_{t-1}(i) a_{ij}^{t-1} \quad (2)$$

式中, $\psi_{t-1}$ 为  $t-1$  时刻隐藏变量的后验概率; $a_{ij}^{t-1}$ 表示  $t-1$  时刻隐藏变量状态  $i$  到  $t$  时刻隐藏变量状态  $j$  的转移概率。这一步主要是获得  $t$  时刻的先验概率。

(3) 根据  $t$  时刻的观测数据对转移矩阵  $A$  中的元素  $a_{ij}^t$  更新,则  $t-1$  时刻隐藏变量状态  $i$  到  $t$  时刻隐藏变量状态  $j$  的转移概率为

$$\xi_{ij}^t = P(x_t = j | x_{t-1} = i, y_t^1, \dots, y_t^m) = \frac{\psi_{t-1}(i) a_{ij}^{t-1} \prod_{k=1}^m b_i^k(k)}{\sum_{j=1}^n \psi_{t-1}(i) a_{ij}^{t-1} \prod_{k=1}^m b_i^k(k)} \quad (3)$$

(4) 求取  $1 \rightarrow t$  时刻参数  $\mathbf{A}$  的均值,进行归一化处理。隐藏变量  $1 \rightarrow t$  时刻状态  $i$  到状态  $j$  的期望转移概率

$$E_t = \sum_{M=1}^t \xi_M^i / t, \text{ 归一化因子为 } 1/C, \text{ 则参数 } \mathbf{A} \text{ 更新为}$$
$$\hat{\mathbf{A}} = (\hat{a}_{1,t}^i)_{n \times n} = \frac{E_t}{C} \tag{4}$$

(5) 在  $1 \rightarrow t$  时刻下,参数  $\mathbf{B}$  是通过统计某个观测值出现时其所对应隐藏状态的数量而获得,即

$$\hat{\mathbf{B}} = \hat{b}_{1,t}^i(k) = \frac{\sum_{M=1}^t \sum_{j=1}^n \xi_M^j}{\sum_{M=1}^t \sum_{j=1}^n \xi_M^j} \tag{5}$$

FR 算法是通过观测信息与先验信息对网络参数进行估计。当某一时刻观测信息缺失时,FR 算法能够根据 DDBN 结构,利用隐藏变量先验信息  $\pi$  和观测变量转移矩阵  $\mathbf{B}$  对缺失数据进行概率推理,获得完整的观测样本。但是当连续几个时刻观测数据发生缺失时,FR 算法的推理精度下降,无法对缺失数据进行准确估计。因此,需要结合一种快速有效的缺失数据估计方法,来弥补这一缺陷。

1.2 数据缺失情况下的 DDBN 参数学习

在观测样本完整的情况下,DDBN 参数学习是对某一类观测数据或状态出现频次的统计过程。但在实际工程应用中,样本数据量小,经常出现样本数据丢失的情况,参数学习难以进行。目前主要有以下 3 种方法来解决这一问题:

- (1) 抛弃缺失数据,这样会造成两个问题,一是会引起数据集变小,可能造成统计样本不够;二是可能这些缺失项的出现不是随机的,抛弃这些数据项会使得数据集无法正确反映实际;
- (2) 给缺失的值赋一个特定的值,如“无”,但这样有可能改变底层的统计关系;
- (3) 对缺失数据进行估计。

在图 1 的 DDBN 中,观测变量能够反映目标的特征,它与反映目标状态的隐藏变量之间具有依赖关系。因此,在某一隐藏状态下,各观测变量的数据之间同样存在一定的非线性对应关系。当 DDBN 中某一观测节点发生数据缺失时,其估计过程可以看作是各观测节点数据对这种非线性函数关系的逼近。而 SVM 算法适合 DDBN 这一特点,它能够通过训练已有的完整数据集建立观测变量之间(对于观测节点间拥有父子关系的网络模型同样适用)非线性函数关系。因此,相比较经典估计理论,基于 SVM 估计方法最大程度地利用了所有数据信息,而不是仅采用单一变量的前后观测完成缺失数据的估计,适合于贝叶斯网络多输入的特点。

2 基于 SVM 的 DDBN 缺失数据估计

2.1 支持向量机缺失数据估计原理

在 DDBN 中观测数据是某个隐藏状态的外在表现形式。因此,DDBN 模型中任意节点的观测数据并不是随机出现的,必须与目标状态保持一致,这也使得每个时刻下的

观测数据间存非线性关系。针对这个特点,将 DDBN 缺失数据估计通过 SVM 训练的非线性函数来逼近<sup>[16]</sup>。

首先,假设一个包含  $l$  组样本  $\{x_i, y_1^i, \dots, y_n^i\}_{i=1}^l$  的数据集,其中  $x_i$  表示输出项(对应 DDBN 中某个含有缺失数据的观测变量), $Y = \{y_1^i, \dots, y_n^i\}_{i=1}^l$  表示输入项(对应 DDBN 中完整数据观测变量)。这里的  $x$  和  $y$ ,并不限定取 +1 或 -1,而可以取任意实数。寻找一个实值函数  $f(Y)$ ,以便用  $x = f(Y)$  来推断  $Y$  所对应的  $x$  值,使得该函数能够表达  $x$  对  $Y$  的关联性。

其次,使用一个非线性映射  $\varphi$  将含有完整数据的训练集  $D = \{y_1^i, \dots, y_n^i\}_{j < l}$  映射到一个高维空间  $H$ 。此时  $x_j (j = 1, \dots, m < l)$  的数据未发生缺失。在高维空间进行训练,可用下式表示:

$$\begin{cases} f(D) = \mathbf{w}^T \varphi(D) + b \\ \mathbf{w} \in H, b \in R, \varphi: R^n \rightarrow H \end{cases} \tag{6}$$

然后,根据结构最小化原理,将函数分类问题转化为求极小化风险函数  $f(D)$ :

$$R_{eq}[f] = \|\mathbf{w}\|^2 / 2 + C \cdot R_{emp} \tag{7}$$

式中,  $\|\mathbf{w}\|^2 / 2$  描述函数的复杂程度; $C$  为预先指定的常数(通常为惩罚因子),它在经验风险  $R_{emp}$  和函数复杂度之间取一折衷。

最后,将非线性分类问题转化为求下面的优化问题

$$\begin{aligned} \min \quad & \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{j=1}^m \xi_j \\ \text{s. t. } \quad & x_j (\mathbf{w}^T \varphi(D) + b) \geq 1 - \xi_j \\ & \xi_j \geq 0, j = 1, \dots, m, m < l \end{aligned} \tag{8}$$

式中,  $\xi_j$  为松弛变量。上述方程可通过对偶形式的 Lagrange 多项式求解。在优化过程中涉及到高维空间中的内积运算,本文采用 RBF 核函数。

至此通过对完整数据集训练构建了  $x$  与  $Y$  间的非线性函数关系,当输出项  $x$  下一时刻数据缺失时,通过输入该时刻下的一组完整数据  $Y$  获得对  $x$  的估计。

2.2 基于 SVM 的 DDBN 缺失数据估计实现

使用文献[17]中的空中目标识别 DDBN 模型检验 SVM 分类算法估计 DDBN 缺失数据的能力,如图 3 所示。这里将观测节点称之为特征变量,其获得的观测值叫做特征数据。

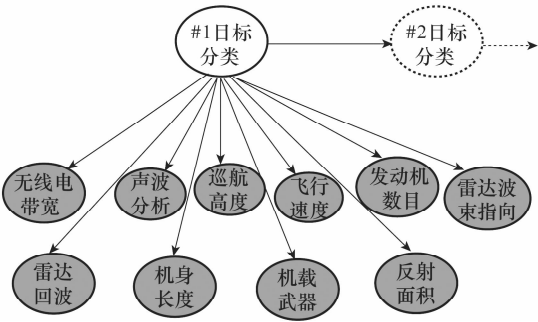


图 3 空中目标识别的 DDBN 模型

该 DDBN 模型的目标识别种类为歼击机和轰炸机。为了获得目标特征数据,根据专家经验构造一组参数假设作为网络模型的真实转移概率。在该参数下对网络的特征变量进行采样,将各时刻下的采样值作为目标特征数据。假设从时刻 21 到时刻 30 中有目标高度数据丢失。根据 SVM 算法将有缺失数据的变量作为 SVM 的输出项,其余的观测变量作为 SVM 的输入项,将这些变量在时刻 21 前的数据分别代入输入和输出项进行训练,获得分类函数。在此基础上将时刻 21 下已获得的特征数据代入该函数完

成对缺失数据的估计。

在核函数中,选用 RBF 核,通过 10 重交叉验证获得惩罚因子  $C=1.452\ 2$ ,核函数宽度  $\sigma=0.017\ 5$ 。表 1 给出特征变量巡航高度  $H$  从 21 时刻到 30 时刻数据发生缺失时的估计结果。在表 1 中,数据 1,2,⋯分别表示特征变量的某一状态,例如机身长度  $L$  的 3 种状态(短、中、长)分别对应特征数据 1,2,3;反射面积  $R$  的 3 种状态(小、中、大)对应特征数据 1,2,3。表 1 中用[]表示该时刻下的缺失数据,使用 SVM 算法的缺失数据估计结果在[]后给出。

表 1 部分缺失属性及基于 SVM 的补全结果

| 组数 | 无线电带宽 $W$ | 雷达回波 $A$ | 声波分析 $S$ | 机身长度 $L$ | 巡航高度 $H$ | 飞行速度 $S$ | 发动机数目 $E$ | 反射面积 $R$ | 雷达波束 $B$ | 机载武器 $P$ |
|----|-----------|----------|----------|----------|----------|----------|-----------|----------|----------|----------|
| 1  | 1         | 1        | 1        | 1        | 3        | 1        | 1         | 2        | 2        | 2        |
| 2  | 2         | 2        | 1        | 3        | 4        | 1        | 2         | 3        | 1        | 1        |
| 3  | 1         | 1        | 1        | 3        | 2        | 1        | 1         | 3        | 1        | 1        |
| 4  | 2         | 1        | 1        | 3        | 2        | 1        | 1         | 3        | 1        | 1        |
| 5  | 2         | 1        | 1        | 3        | 3        | 1        | 1         | 2        | 2        | 1        |
| ⋮  | ⋮         | ⋮        | ⋮        | ⋮        | ⋮        | ⋮        | ⋮         | ⋮        | ⋮        | ⋮        |
| 21 | 1         | 1        | 1        | 2        | [2]2     | 1        | 1         | 2        | 1        | 1        |
| 22 | 2         | 1        | 2        | 3        | [3]3     | 1        | 1         | 3        | 2        | 1        |
| 23 | 1         | 1        | 1        | 1        | [3]3     | 1        | 1         | 2        | 1        | 1        |
| 24 | 2         | 1        | 1        | 2        | [3]3     | 1        | 1         | 3        | 2        | 1        |
| 25 | 1         | 1        | 2        | 3        | [4]3     | 1        | 1         | 3        | 2        | 1        |
| 26 | 1         | 1        | 2        | 2        | [3]3     | 1        | 2         | 3        | 1        | 1        |
| 27 | 1         | 2        | 1        | 3        | [3]2     | 1        | 1         | 3        | 2        | 1        |
| 28 | 2         | 1        | 1        | 2        | [2]2     | 1        | 1         | 3        | 1        | 1        |
| 29 | 2         | 1        | 1        | 3        | [1]2     | 1        | 1         | 3        | 1        | 1        |
| 30 | 2         | 1        | 2        | 3        | [3]3     | 1        | 1         | 3        | 1        | 1        |

3 数据缺失下的 DDBN 参数学习实现

数据缺失下的 DDBN 参数学习流程如图 4 所示<sup>[16]</sup>。

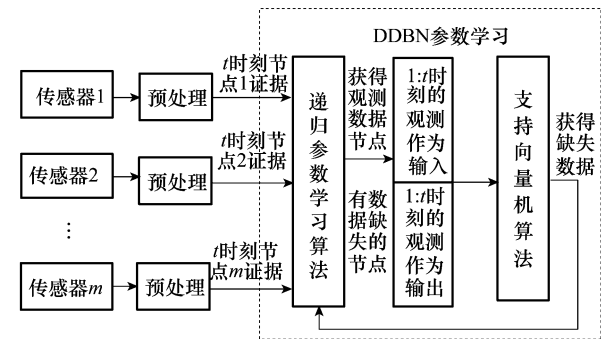


图 4 数据缺失下的 DDBN 参数学习流程

- (1) 通过  $m$  个传感器对未知事物进行感知,将其作为隐藏变量的外在表现节点,构造 DDBN;将获得观测信息经过预处理(例如坐标转换、数据对准、模糊处理等),作为特征数据导入 DDBN 所对应的  $m$  个观测节点;
- (2)  $t$  时刻下第  $k$  个观测节点丢失数据,将该节点从时

- 刻 1 到时刻  $t-1$  的观测数据作为 SVM 的输出,其他节点从时刻 1 到时刻  $t-1$  的观测数据作为输入,将输入和输出集映射到一个高维空间  $H$  中,进行分类函数的训练,寻找 SVM 的最优超平面,建立输入与输出间的映射关系;
- (3) 在获得分类函数的基础上,将  $t$  时刻没有缺失的各节点观测数据作为该分类函数的输入,估计该时刻下缺失的观测数据;
- (4) 在获得完整数据的基础上,采用 FR 算法,完成当前时刻下的参数学习。随着 DBN 模型不断扩展,获得的网络参数趋于稳定,最终获得全局最优的参数值。

4 仿真试验

- 在 Pentium M 处理器、CPU 主频 1 GHz、内存为 512 MB 试验平台上采用 Matlab 软件进行仿真,给出下列初始条件:
- (1) 使用图 3 空中目标识别模型,根据专家经验构造一组参数假设作为网络模型的真实转移概率,观测节点采样都是在该参数下进行;
- (2) 特征变量中机身长度  $L$ 、反射面积  $R$ 、雷达波束  $B$  包含 3 种状态,巡航高度  $H$  包含 4 种状态,其他变量都为 2 种状态;

- (3) 目标识别模型扩展为 4 个时间片,按照每次采集 10 组、20 组、50 组的采样模式对模型特征变量进行采样,即每种采样模式下对 4 个时间片特征变量依次采样;
- (4) 每个时间片采样数据按 10%、20%进行缺失处理;
- (5) 采用的对比算法为 EM 算法<sup>[13]</sup>;
- (6) SVM-FR 与 EM 初始参数  $\{\pi_0, A_0, B_0\}$  相同,EM 算法迭代次数为 10 次。

采用 SVM-FR 与 EM 两种算法在含有缺失数据的 3 种采样模式下进行参数估计,两种算法运算时间如表 2 所示。

| 表 2 两种算法估计数据运行时间 s |        |        |        |
|--------------------|--------|--------|--------|
|                    | 10 组/片 | 20 组/片 | 50 组/片 |
| SVM-FR 10%         | 0.3    | 0.9    | 1.1    |
| EM 10%             | 6.1    | 12.5   | 29.5   |
| SVM-FR 20%         | 0.4    | 1      | 1.3    |
| EM 20%             | 6.3    | 12.8   | 29.7   |

下面给出采样次数为 20 组/片,含有 20%缺失数据的参数学习结果,如表 3 所示。表 3 中给出识别模型中部分参数。

| 表 3 缺失 20%下的部分参数学习结果 |             |              |             |             |
|----------------------|-------------|--------------|-------------|-------------|
|                      | T1(战斗机)     | T2(轰炸机)      | SVM-FR      | EM          |
| 片间                   | T1/T1/0.665 | T2/T1/ 0.339 | T1/T1/0.442 | T2/T1/ 0.5  |
| 转移                   | T1/T2/0.335 | T2/T2/ 0.661 | T1/T2/0.558 | T2/T2/ 0.5  |
| 先验                   | 0.487       | 0.513        | 0.5         | 0.5         |
| W                    | T1/W1/ 0.4  | T2/W1/0.992  | T1/W1/.368  | T2/W1/0.833 |
|                      | T1/W2/ 0.6  | T2/W2/0.008  | T1/W2/.632  | T2/W2/0.167 |
| H                    | T1/H1/0.688 | T2/H1/0.001  | T1/H1/0.684 | T2/H1/0     |
|                      | T1/H2/0.3   | T2/H2/0.02   | T1/H2/0.29  | T2/H2/0.142 |
|                      | T1/H3/0.01  | T2/H3/0.859  | T1/H3/0.026 | T2/H3/0.762 |
|                      | T1/H4/0.002 | T2/H4/0.12   | T1/H4/0     | T2/H4/0.096 |
| S                    | T1/S1/0.6   | T2/S1/0.3    | T1/S1/0.815 | T2/S1/0.238 |
|                      | T1/S2/0.4   | T2/S2/0.7    | T1/S2/0.185 | T2/S2/0.762 |
| B                    | T1/ B1/0.1  | T2/ B1/0.9   | T1/B1/0.132 | T2/B1/0.785 |
|                      | T1/ B2/0.9  | T2/ B2/0.1   | T1/B2/0.868 | T2/B2/0.215 |

在 10% 数据缺失条件下,采用 KL 分离 (Kullback-Leibler divergence,  $KL(\theta|\hat{\theta})=\sum_x P_{\theta}(X)\ln\frac{P_{\theta}(X)}{P_{\hat{\theta}}(X)}$ ) 测试获得的片间转移概率与给定参数的差距。KL 分离的平均数在表 4 中给出。

| 表 4 KL 分离的平均数 |         |         |         |         |
|---------------|---------|---------|---------|---------|
|               | T=4     |         |         |         |
|               | 10 组    | 20 组    | 40 组    | 50 组    |
| SVM-FR        | 0.109 8 | 0.068 7 | 0.012 7 | 0.008 6 |
| EM            | 0.124 3 | 0.169 3 | 0.047 9 | 0.009 8 |

如表 4 所示,在有缺失的小样本条件下,通过 SVM-FR 获得的片间转移概率精度高于 EM 算法。随着时间片不断扩展,两种算法学习精度不断提高,同时可以看到 SVM-FR 在 20 组采样条件下与 EM 在 40 组采样条件下参数学习的结果精度相当,表明提出的 SVM-FR 参数学习算法在一定

程度上减少了对大数据量的依赖,对小样本达到大数据量参数精度是有帮助的。

在 20% 数据缺失,采样次数为 20 组/片的条件下,将 SVM-FR 和 EM 两种算法获得参数分别代入识别模型,对目标类型进行推理。随机采样 30 组数据作为证据输入识别模型,使用连接树算法进行推理。两种算法在各时间片下目标识别结果如图 5 所示。

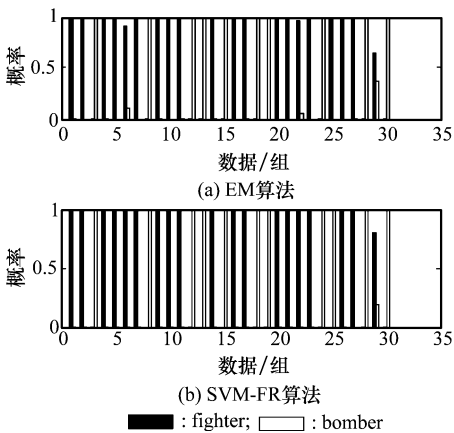


图 5 两种算法在各时间片下目标识别结果

由图 5 可以看出,除去第 25 时间片,两种算法目标识别结果一致。造成第 25 片推理结果差异的原因是由 SVM 对缺失数据估计误差造成的。每个时间片下 SVM-FR 获得的目标识别概率高于 EM 算法,说明 SVM-FR 算法获得的参数能够满足目标识别模型推理要求。

5 结 论

本文围绕数据缺失条件下 DDBN 参数学习问题展开研究,根据 DDBN 的特点提出 FR 参数学习算法,充分利用 DDBN 概率传递的特点获得片间与片内的转移概率。当特征数据发生缺失时引入 SVM 算法进行缺失数据估计,在此基础上实现数据缺失下 DDBN 参数学习。通过空中目标识别模型的仿真发现 SVM-FR 算法具有以下优势:

- (1) 参数学习时间短,整个学习过程所用时间远远小于 EM 算法;
- (2) 参数估计精度较高,在两种算法获得的网络参数下,识别模型推理结果精度相当;
- (3) 适应小样本条件下参数学习。

同时 SVM-FR 算法还存在以下问题:

- (1) 参数估计精度依赖于先验信息的获取;
- (2) SVM 在对缺失数据估计过程,需要多种特征来完成分类函数的训练,因此构建的网络模型考虑的因素增多,增加了建模的复杂性。

针对这些问题,下一步准备进行约束条件下参数学习,将领域知识引入动态网络参数学习当中,进一步提高网络参数学习的精度与运算效率。

## 参考文献:

- [1] Song L, Kolar M, Xing E. Time-varying dynamic Bayesian networks[C]// *Proc. of the 23rd Neural Information Processing Systems*, 2009:1732–1740.
- [2] Campos C P, Zeng Z, Ji Q. Structure learning of Bayesian networks using constraints [C]// *Proc. of the 26th Annual International Conference on Machine Learning*, 2009:113–120.
- [3] Chen H, Gao X. Forwards-backwards information repairing algorithm and appliance on discrete dynamic Bayesian networks[C]// *Proc. of the International Conference on Intelligent Human-Machine Systems and Cybernetics*, 2009:76–80.
- [4] Dojer N, Gambin A, Mizera A, et al. Applying dynamic Bayesian networks to perturbed gene expression data [J]. *BMC Bioinformatics*, 2006, 7(1):249.
- [5] Saenko K, Livescu K, Glass J, et al. Multistream articulatory feature-based models for visual speech recognition [J]. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2009, 31(9):1700–1707.
- [6] Hearty P, Fenton N, Marquez D, et al. Predicting project velocity in XP using a learning dynamic Bayesian network model[J]. *IEEE Trans. on Software Engineering*, 2009, 35(1):124–137.
- [7] Rajapakse J C, Wang Y, Zheng X, et al. Probabilistic framework for brain connectivity from functional MR images [J]. *IEEE Trans. on Medical Image*, 2008, 27(6):825–833.
- [8] Jaeger M. Parameter learning for relational Bayesian networks [C]// *Proc. of the 24th International Conference on Machine Learning*, 2007:369–376.
- [9] Niculescu R S, Mitchell T M, Rao R B. A theoretical framework for learning Bayesian networks with parameter inequality Constraints[C]// *Proc. of the 20th International Joint Conference on Artificial intelligence*, 2007:155–160.
- [10] Campos C P, Ji Q. Improving Bayesian network parameter learning using constraints[C]// *Proc. of the 19th International Conference on Pattern Recognition*, 2008:113–120.
- [11] 王双成, 冷翠平, 杜瑞杰. 贝叶斯网络参数学习中的噪声平滑[J]. 系统仿真学报, 2009, 21(16):5053–5060. (Wang S C, Leng C P, Du R J. Noise smoothing in learning parameters of Bayesian network[J]. *Journal of System Simulation*, 2009, 21(16):5053–5060.)
- [12] 王双成, 冷翠平. 贝叶斯网络适应性学习[J]. 小型微型计算机系统, 2009, 30(4):706–709. (Wang S C, Leng C P. Adaptability learning for Bayesian network[J]. *Journal of Chinese Computer Systems*, 2009, 30(4):706–709.)
- [13] Friedman N. The Bayesian structural EM algorithm[C]// *Proc. of the Fourteenth Annual Conference on Uncertainty in Artificial Intelligence*, 1998:125–133.
- [14] Sehgal M. Collateral missing value imputation: a new robust missing value estimation algorithm for microarray data [J]. *Bioinformatics*, 2005, 21:2417–2423.
- [15] Oba S, Masaaki S. A Bayesian missing value estimation method for gene expression profile data [J]. *Bioinformatics*, 2003, 19:2088–2096.
- [16] 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 J, Gao X G, Ru W. Parameters learning of BN in small sample base on data missing [J]. *Systems Engineering-Theory and Practice*, 2011, 31(1):172–177.)
- [17] 史建国, 高晓光, 李相民. 离散模糊动态贝叶斯网络用于无人作战飞机目标识别[J]. 西北工业大学学报, 2006, 24(1):45–49. (Shi J G, Gao X G, Li X M. Improving identification capability of UCAV(unmanned combating air vehicle) [J]. *Journal of Northwestern Polytechnical University*, 2006, 24(1):45–49.)