

文章编号: 1001-0920(2010)11-1737-05

具有传递变量的动态贝叶斯网络结构学习

王双成¹, 王 辉², 许广林^{1a}

(1. 上海立信会计学院 a. 数学与信息学院, b. 开放经济与贸易研究中心, 上海 201620; 2. 中央民族大学信息工程学院, 北京 100081)

摘 要: 针对现有学习方法对完全时间不对称数据的动态贝叶斯网络学习不具有实用性, 提出一种借助传递变量进行完全时间不对称数据的动态贝叶斯网络结构学习方法. 首先进行相邻时间片间的传递变量序列学习; 然后, 基于节点排序和局部打分-搜索, 进行动态贝叶斯网络局部结构学习; 最后通过时序扩展得到整个动态贝叶斯网络结构.

关键词: 动态贝叶斯网络; 完全时间不对称数据; 传递变量; 结构学习

中图分类号: TP181

文献标识码: A

Learning dynamic Bayesian network structure with transfer variables

WANG Shuang-cheng¹, WANG Hui², XU Guang-lin^{1a}

(1a. School of Mathematics and Information, 1b. Opening Economy and Trade Research Center, Shanghai Lixin University of Commerce, Shanghai 201620, China; 2. School of Information Engineering, The Central University for Nationalities, Beijing 100081, China. Correspondent: WANG Shuang-cheng, E-mail: wangsc@lixin.edu.cn)

Abstract: At present, there are not the methods of learning dynamic Bayesian network structure from no time symmetry data. Therefore, a method of learning dynamic Bayesian network structure from no time symmetry data is developed by combining transfer variables. In this method, first transfer variable series between two adjacent time slices are found. Then dynamic Bayesian network part structure can be learned based on sorting nodes and local search and scoring method. Finally, a whole dynamic Bayesian network structure can be presented by extending along time series.

Key words: Dynamic Bayesian network; No time symmetry data; Transfer variable; Structure learning

1 引 言

动态贝叶斯网络 (DBN)^[1] 是贝叶斯网络与时间序列的结合, 在继承贝叶斯网络优势的同时, 又增加了时间特征功能, 可将传统的不同时间片间时序因果联系与现代的时间片内因果关系融为一体, 并通过量化推理进行动态分析、预测和控制, 已在语音识别、图像处理、视频分析、生物信息分析和系统控制等方面得到了广泛的应用^[2-4]. DBN 同贝叶斯网络一样, 也由结构和参数两部分构成, 根据结构和数据可估计出参数, 因此结构学习同样是 DBN 学习的核心.

基于 DBN 解决实际问题的基础是动态贝叶斯网络学习, 而一般 DBN 学习非常复杂, 往往没有实际意义. 为简化问题, Friedman^[1] 首先给出了受平稳性和马尔科夫性两个假设约束的 DBN, 这种约束 DBN 可转化为先验网和转换网, 并使用相邻时间片完全时

间对称数据和整体打分-搜索方法进行先验网和转换网结构学习. 目前已发展了许多 DBN 结构学习方法^[5,6], 但这些方法主要是基于主观领域专家知识或 Friedman 方法在打分函数和搜索策略方面的改进算法, 对普遍存在的相邻时间片数据完全不时间对称 (同一种情况在两个相邻时间片数据集中的观测记录不对应) 情况下的 DBN 结构学习还缺少针对性研究.

本文通过引入传递变量来建立不同时间片变量之间的时序联系, 在基于节点排序和局部打分-搜索方法进行 DBN 结构局部学习的基础上, 结合时序扩展便能实现 DBN 结构的整体学习. 实验结果显示了使用该方法进行相邻时间片完全时间不对称数据 DBN 结构学习的有效性.

收稿日期: 2009-09-24; 修回日期: 2010-02-17.

基金项目: 国家自然科学基金项目(60675036); 上海市教委重点学科建设项目(J51702); 上海市教委科研创新重点项目(09zz202).

作者简介: 王双成(1958-), 男, 吉林蛟河人, 教授, 博士, 从事人工智能、机器学习等研究; 王辉(1961-), 男, 吉林榆树人, 教授, 硕士, 从事数据挖掘和应用等研究.

2 传递变量学习

本文用 $\mathbf{X}[1], \dots, \mathbf{X}[T]$ 表示随机向量序列, $\mathbf{X}[t] = \{X_1[t], \dots, X_n[t]\}, 0 \leq t \leq T; \mathbf{x}[t] = \{x_1[t], \dots, x_n[t]\}$ 为具体的配置; $D[0], D[1], \dots, D[T]$ 表示时间片数据集; $N[0], N[1], \dots, N[T]$ 是对应数据集中记录的数量; 在概率模式中的变量和表示概率模式的图形模式中的节点有时不加区分.

由不同时间片中对应的传递变量形成马尔科夫链, 传递变量学习需要按照时序依次进行, 其学习顺序为 $H[0, 1], H[0, 1, 2], \dots, H[T-2, T-1, T]$. 第 1 个传递变量称为初始传递变量, 其他传递变量称为后继传递变量. 下面给出传递变量学习方法.

2.1 初始传递变量学习

在变量集 $\{X_1[0], \dots, X_n[0]\}$ 和 $\{X_1[1], \dots, X_n[1]\}$ 中插入传递变量 $H[0, 1]$, 将 $X_1[0], \dots, X_n[0]$ 和 $X_1[1], \dots, X_n[1]$ 的时间片数据集 $D[0]$ 和 $D[1]$ 叠加得到 $D[0, 1]$. 在 $D[0, 1]$ 中增加传递变量值后的数据集用 $\mathbf{DH}[0, 1]$ 表示, $N[0, 1] = N[0] + N[1]$ 表示记录数量, 叠加后数据集的变量用 $X_1[0, 1], \dots, X_n[0, 1]$ 表示 (只是一种形式表示). 传递变量 $H[0, 1]$ 和两组变量之间选择普遍采用的星形结构, 用 $S[0, 1]$ 表示.

传递变量学习包括发现传递变量的最优维数和取值. 对于传递变量 $H[0, 1]$ 和指定的维数, 根据星形结构和 Gibbs 抽样^[7] 确定传递变量的值, 在此基础上, 以最优解释能力为评价标准, 通过对各种可能维数打分比较来发现最优传递变量的维数.

2.1.1 确定传递变量的值

对于选择的维数, 在数据集 $\mathbf{DH}[0, 1]$ 中随机初始化 $H[0, 1]$ 的值, 并进行迭代修正, 直到满足终止条件时结束迭代. 在每一次迭代中, 按数据集中记录的顺序依次对传递变量 $H[0, 1]$ 的值进行修正, 并调整相关参数, 直到修正完所有传递变量的值, 实现一次迭代. 一次迭代的时间复杂度是 $O(N[0, 1])$.

1) 修正传递变量 $H[0, 1]$ 的值

设 $H[0, 1]$ 在第 $m (1 \leq m \leq N[0, 1])$ 个记录具有待修正值 $h_m[0, 1]$, $\hat{h}_m[0, 1]$ 表示修正后的值, 则 $H[0, 1]$ 的可能取值为 $h^1[0, 1], \dots, h^{r(0,1)}[0, 1]$. 用 $\mathbf{DH}^{(k)}[0, 1]$ 表示在 $\mathbf{DH}[0, 1]$ 基础上第 $k+1$ 次迭代修正前的数据集, $\mathbf{DH}_m^{(k)}[0, 1]$ 表示第 $k+1$ 次迭代修正中对 $h_m[0, 1]$ 修正后的数据集, $\mathbf{DH}^{(k+1)}[0, 1]$ 表示第 $k+1$ 次迭代修正后的数据集.

基于星形结构, 可得

$$p(h[0, 1] | x_{m1}[0, 1], \dots, x_{mn}[0, 1],$$

$$\begin{aligned} & \mathbf{DH}_{m-1}^{(k)}[0, 1], S[0, 1]) = \\ & \alpha[0, 1] p(h[0, 1] | \mathbf{DH}_{m-1}^{(k)}[0, 1], S[0, 1]) \times \\ & \prod_{i=1}^n p(x_{mi}[0, 1] | h[0, 1], \mathbf{DH}_{m-1}^{(k)}[0, 1], S[0, 1]). \end{aligned} \quad (1)$$

其中 $\alpha[0, 1]$ 为与变量 $H[0, 1]$ 无关的量. 对

$$p(h[0, 1] | \mathbf{DH}_{m-1}^{(k)}[0, 1], S[0, 1]) \times \prod_{i=1}^n p(x_{mi}[0, 1] | h[0, 1], \mathbf{DH}_{m-1}^{(k)}[0, 1], S[0, 1])$$

进行归一化处理, 记

$$w(u) = \frac{p(h^u[0, 1] | \mathbf{DH}_{m-1}^{(k)}[0, 1], S[0, 1])}{\sum_{j=1}^{r(0,1)} p(h^j[0, 1] | \mathbf{DH}_{m-1}^{(k)}[0, 1], S[0, 1])} \times \frac{\prod_{i=1}^n p(x_{mi}[0, 1] | h^u[0, 1], \mathbf{DH}_{m-1}^{(k)}[0, 1], S[0, 1])}{\prod_{i=1}^n p(x_{mi}[0, 1] | h^j[0, 1], \mathbf{DH}_{m-1}^{(k)}[0, 1], S[0, 1])},$$

$$u \in \{1, \dots, r(0,1)\}.$$

对于生成的随机数 λ , 传递变量 $H[0, 1]$ 的取值为

$$\hat{h}[0, 1] = \begin{cases} h^1[0, 1], & 0 < \lambda < w(1); \\ \vdots & \\ h^u[0, 1], & \sum_{k=1}^{u-1} w(k) < \lambda \leq \sum_{k=1}^u w(k); \\ \vdots & \\ h^{r(0,1)}[0, 1], & \lambda > \sum_{k=1}^{r(0,1)-1} w(k). \end{cases} \quad (2)$$

2) 调整参数

如果 $h_m[0, 1] \neq \hat{h}_m[0, 1]$, 则需调整对应的参数, 然后利用调整后的参数进行抽样以修正下一个待修正的数据. 参数调整方法可参见文献 [8].

2.1.2 迭代终止检验

采用相邻传递变量值序列的一致性检验进行终止迭代判断. 设相邻两次迭代所得到的传递变量值序列分别为

$$h_1^{(k)}[0, 1], h_2^{(k)}[0, 1], \dots, h_{N[0,1]}^{(k)}[0, 1],$$

$$h_1^{(k+1)}[0, 1], h_2^{(k+1)}[0, 1], \dots, h_{N[0,1]}^{(k+1)}[0, 1].$$

则有

$$\begin{aligned} & \text{sig}(h_i^{(k)}[0, 1], h_i^{(k+1)}[0, 1]) = \\ & \begin{cases} 0, & h_i^{(k)}[0, 1] = h_i^{(k+1)}[0, 1]; \\ 1, & h_i^{(k)}[0, 1] \neq h_i^{(k+1)}[0, 1]. \end{cases} \end{aligned} \quad (3)$$

对于给定的阈值 $\delta \geq 0$, 如果

$$\frac{1}{N[0, 1]} \sum_{i=1}^{N[0, 1]} \text{sig}(h_i^{(k)}[0, 1], h_i^{(k+1)}[0, 1]) \leq \delta,$$

则结束迭代.

2.2 确定传递变量的维数

基于 DBN 解决实际问题的方式是推理, 而最优解释是 DBN 的主要推理方式之一. 因其他推理都与最优解释推理密切相关, 故选择变量之间的最优解释能力作为确定最优维数的打分标准, 会使引入传递变量 DBN 的推理更加有效和可靠.

定义 1 对于变量集 $\{X_1[0, 1], \dots, X_n[0, 1], H[0, 1]\}$, 用 $\text{OEA}(h[0, 1] | x_1[0, 1], \dots, x_n[0, 1], S[0, 1])$ 表示 $X_1[0, 1], \dots, X_n[0, 1]$ 对变量 $H[0, 1]$ 的最优解释能力, 即

$$\begin{aligned} \text{OEA}(h[0, 1] | x_1[0, 1], \dots, x_n[0, 1], S[0, 1]) &= \\ & \frac{1}{N(x_1[0, 1], \dots, x_n[0, 1])} \times \\ & \sum_{x_1[0, 1], \dots, x_n[0, 1]} \max_{h[0, 1](x_1[0, 1], \dots, x_n[0, 1])} \{p(h[0, 1] | \\ & x_1[0, 1], \dots, x_n[0, 1], S[0, 1])\} = \\ & \frac{1}{N(x_1[0, 1], \dots, x_n[0, 1])} \times \\ & \sum_{x_1[0, 1], \dots, x_n[0, 1]} \max_{h[0, 1](x_1[0, 1], \dots, x_n[0, 1])} p(h[0, 1] | \\ & S[0, 1]) \prod_{i=1}^n p(x_i[0, 1] | h[0, 1], S[0, 1]), \end{aligned} \quad (4)$$

其中 $N(x_1[0, 1], \dots, x_n[0, 1])$ 是 $X_1[0, 1], \dots, X_n[0, 1]$ 的配置情况数量.

设 $M[0, 1]$ 是 $H[0, 1]$ 最大可能的维数(可根据先验知识确定), 依据星形结构和 Gibbs 抽样得到数据集序列 $D_2[0, 1], \dots, D_{M[0, 1]}[0, 1]$, 再对每一个数据集分别进行最优解释能力打分. 选择

$$k_0 = \max_{M[0, 1] \geq k \geq 2} \{\text{OEA}(h[0, 1] | x_1[0, 1], \dots, x_n[0, 1], D_k[0, 1], S[0, 1])\}$$

作为传递变量的最优维数, $D_{k_0}[0, 1]$ 是具有传递变量最后取值的数据集.

2.3 后继传递变量学习

当在变量集 $\{X_1[t-1], \dots, X_n[t-1]\}$ 和 $\{X_1[t], \dots, X_n[t]\}$ 中插入传递变量时, 因还受前一个传递变量的影响, 故采用 $H[t-2, t-1, t]$ 来表示. 利用上述方法确定 $H[t-2, t-1, t]$ 的值需要两个阶段: 首先依据 $\text{DH}[t-2, t-1]$ 中包含 $D[t-1]$ 的数据进行传递变量 $H[t-2, t-1, t]$ 与 $D[t-1]$ 对应部分的学习, 这样得到 $\text{DH}[t-1, t]$ 中传递变量 $H[t-2, t-1, t]$ 与 $D[t-1]$ 相对应的值; 然后再学习 $\text{DH}[t-1, t]$ 中与 $D[t]$ 相对应部

分的值, 从而实现传递变量 $H[t-2, t-1, t]$ 的学习, 并获得完整的数据集 $\text{DH}[t-1, t]$.

在 $H[t-2, t-1, t]$ 学习中, 两个阶段用于抽样的分解式需改写成

$$p(h[t-2, t-1, t] | h[t-3, t-2, t-1], \text{DH}_{m-1}^{(k)}[t-1, t], S[t-1, t]) \prod_{i=1}^n p(x_{mi}[t-1, t] |$$

$$h[t-2, t-1, t], \text{DH}_{m-1}^{(k)}[t-1, t], S[t-1, t])$$

和

$$p(h[t-2, t-1, t] | \text{DH}_{m-1}^{(k)}[t-1, t], S[t-1, t]) \times \prod_{i=1}^n p(x_{mi}[t-1, t] | h[t-2, t-1, t], \text{DH}_{m-1}^{(k)}[t-1, t], S[t-1, t]).$$

在第 1 阶段确定 $H[t-2, t-1, t]$ 的最优维数和部分取值; 第 2 阶段在已发现最优维数和部分取值的基础上确定剩余部分的值.

3 具有传递变量的 DBN 结构学习

具有传递变量的 DBN 结构局部学习涉及两种情况, 分别是首尾局部结构和中间局部结构学习. 下面分别给出学习方法.

3.1 首尾局部结构学习

由于首尾局部结构学习基本相同, 这里只给出首局部结构学习方法. 用 $\text{DH}[t](0 \leq t \leq T)$ 表示在 $D[t]$ 中增加了传递变量值后的数据集. 依据数据集 $\text{DH}[0]$ 建立传递变量 $H[0, 1]$ 约束下 $X_1[0], \dots, X_n[0]$ 的完全有向无环图, 并基于完全有向无环图排序节点(这种节点排序方法称为完全有向无环图法); 然后在节点顺序的基础上, 采用局部打分-搜索方法进行具有传递变量的首局部结构学习.

3.1.1 节点排序

$R(X_j[0] \rightarrow X_i[0] | H[0, 1])$ 表示在 $H[0, 1]$ 约束下, $X_j[0]$ 对 $X_i[0]$ 的条件相对平均熵, 即

$$R(X_j[0] \rightarrow X_i[0] | H[0, 1]) = \frac{\text{CE}(X_i[0] | X_j[0], H[0, 1])}{\text{CE}(X_i[0] | H[0, 1]) \cdot |X_i[0]|}. \quad (5)$$

其中

$$\begin{aligned} \text{CE}(X_i[0] | X_j[0], H[0, 1]) &= \\ & - \sum_{x_i[0], x_j[0], h[0, 1]} p(x_i[0], x_j[0], h[0, 1]) \times \\ & \log p(x_i[0] | x_j[0], h[0, 1]) \end{aligned}$$

和

$$\text{CE}(X_i[0] | H[0, 1]) =$$

$$- \sum_{x_i[0], h[0,1]} p(x_i[0], h[0,1]) \log p(x_i[0]|h[0,1])$$

是条件熵。

对于任意的 $X_i[0], X_j[0]$, 如果 $R(X_j[0] \rightarrow X_i[0]|H[0,1], DH[0]) > R(X_i[0] \rightarrow X_j[0]|H[0,1], DH[0])$, 则定向为 $X_j[0] \rightarrow X_i[0]$; 否则, 定向为 $X_i[0] \rightarrow X_j[0]$. 可以证明, 采用这样的定向方法所建立的完全有向图不存在环路, 使用所建立的完全有向无环图排序节点, 最终获得的节点序列为 $X_{d_1}[0], \dots, X_{d_n}[0]$.

3.1.2 局部结构学习

在节点排序的基础上, 使用打分-搜索方法进行 DBN 结构局部学习, 打分函数和搜索策略分别选择普遍采用的 MDL (Minimal description length)^[9] 打分和贪婪搜索. 在最坏情况下 (学习得到的是每两个时间片节点之间都存在一条弧的有向完全图), 算法的时间复杂度是 $O(n^3)$.

3.2 中间局部结构学习

中间局部结构学习与首局部结构学习类似, 差别仅在于首局部结构学习中 $X_1[0], \dots, X_n[0]$ 只受 $H[0,1]$ 的约束, 而在关于 $X_1[t-1], \dots, X_n[t-1]$ ($2 \leq t \leq T-1$) 的中间局部结构中受到 $H[t-3, t-2, t-1]$ 和 $H[t-2, t-1, t]$ 的同时约束. 将局部学习过程进行时序延展便可得到整个动态贝叶斯网络结构.

4 实验与分析

使用模拟和真实数据, 分别从传递变量学习, DBN 局部结构学习和引入传递变量的有效性 3 个方面进行实验与分析.

4.1 数据集

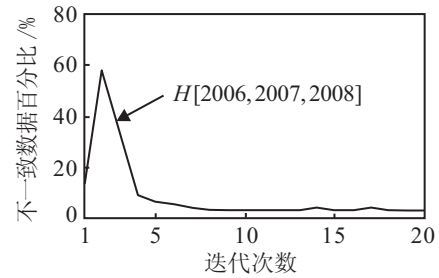
对于真实数据, 选择与企业综合风险 (包括财务风险和经营风险) 相关的 26 个指标变量, 并从相关数据源获取这些变量的时间片 (面板) 数据, 再根据专家知识对连续变量进行离散化, 得到所需的真实实验数据. 对于模拟数据, 将网站 <http://www.norsys.com> 提供的 ALARM 网, Boerlage 网, Car-Diagnosis 网和 Asia 网与时间序列结合, 建立 DBN, 从而生成时间片模拟数据用于实验.

4.2 传递变量学习

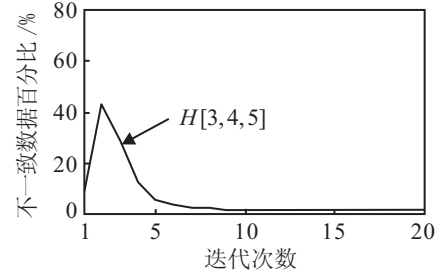
分别使用真实数据和 ALARM 网模拟数据进行确定传递变量值迭代的收敛性和获取传递变量的最优维数实验.

4.2.1 确定传递变量值迭代的收敛性

设置阈值 $\delta = 5\%$, 传递变量 $H[2006, 2007, 2008]$ (真实数据传递变量) 和 $H[3, 4, 5]$ (ALARM 网模拟数据传递变量) 的维数分别取 6 和 5. 确定传递变量值迭代的收敛情况如图 1 所示.



(a) 真实数据



(b) 模拟数据

图 1 确定传递变量值迭代的收敛情况

从图 1 可以看出, 确定 $H[2006, 2007, 2008]$ 的值经过 7 次迭代后收敛; 而确定 $H[3, 4, 5]$ 的值经过 6 次迭代便收敛, 显示出了较高的效率, 而且没有出现大的波动现象, 表明传递变量的引入与其他变量是相容的.

4.2.2 获取传递变量的最优维数

利用变量之间最优解释能力标准评价传递变量维数的各种情况, 传递变量的维数范围均选为 2~10. 具体打分情况如图 2 所示.

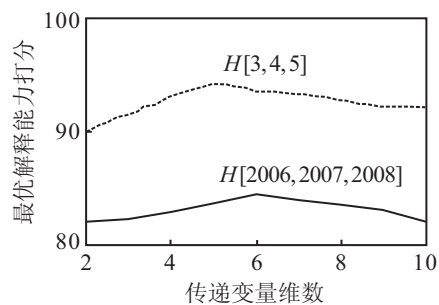


图 2 获取传递变量的最优维数

从图 2 可以看出, 传递变量 $H[2006, 2007, 2008]$ 和 $H[3, 4, 5]$ 的最优维数分别为 6 和 5.

4.3 DBN 结构学习

分别为数据集 $D[2003], \dots, D[2007]$ 和 $D[2008]$ 增加传递变量的值, 使用完全有向无环图法排序节点, 得到节点序列. 在此基础上, 基于局部打分-搜索方法进行 DBN 结构局部学习, 学习过程沿时序延展便可得到整个 DBN 结构, 如图 3 所示. 其中: 数字表示变量的编号, 传递变量连接的左右两侧分别是两个数据集学习获得的 DBN 局部结构.

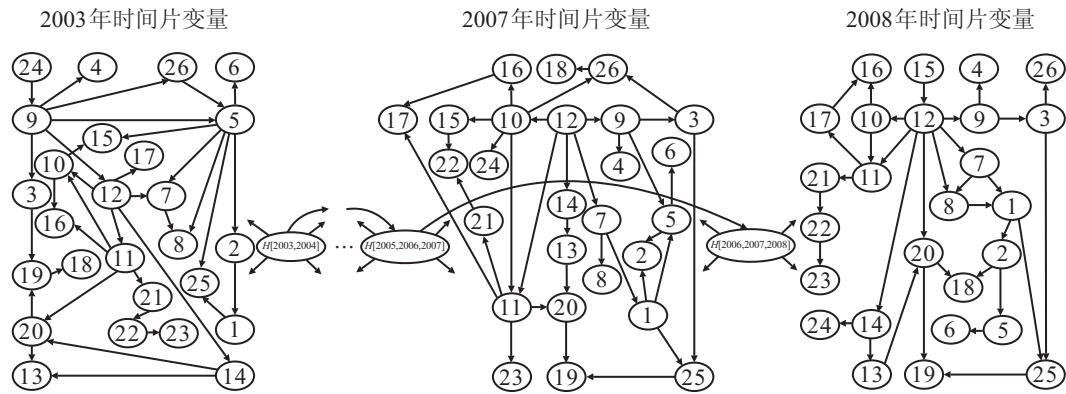


图 3 学习得到的 DBN 结构

4.4 引入传递变量的有效性

变量之间的最优解释能力是约束下(除被考察变量外,其他变量均实例化)所有变量的不确定性程度的平均值,最优解释能力越高,表示变量之间解释能力越强,约束下的不确定程度越低.因此,相对于似然

或 MDL 标准等,该指标能更有效地刻画引入传递变量的有效性.使用真实和模拟数据,为时间片数据集引入相关传递变量的值,并对增加传递变量前后原有时间片变量之间最优解释能力打分进行比较.具体情况如表 1 和表 2 所示.

表 1 真实数据集传递变量传递信息的有效性

网络情况	DH[2003]	DH[2004]	DH[2005]	DH[2006]	DH[2007]	DH[2008]
无传递变量	74.68	74.12	70.71	73.26	71.65	72.93
有传递变量	76.85	76.49	73.22	75.38	74.87	74.58
提高解释能力/%	2.91	3.19	3.54	2.89	4.49	2.26

表 2 模拟数据集传递变量传递信息的有效性

模拟数据集	DH[0]	DH[1]	DH[2]	DH[3]	DH[4]	DH[5]
ALARM	2.65	2.92	3.71	3.14	3.33	2.86
Car Diagnosis	3.84	4.31	4.05	4.47	4.45	3.96
Boerlage	2.53	3.13	3.21	3.35	3.18	2.84
Asia	2.45	2.85	2.92	2.76	3.24	2.37

从表 1 和表 2 可以看出,引入传递变量确实能够有效提高时间片变量之间的最优解释能力,即传递变量可实现相邻时间片之间的信息传递,从而提高基于 DBN 推理的可靠性.另外,通过与领域专家的交流,上面学习得到的具有传递变量的动态贝叶斯网络结构既反映了经验中的因果联系,又展示了许多专家经验中所没有但在数据中蕴含的因果关系,切实起到了客观因果知识对主观因果经验的修正和补充的作用.

5 结 论

本文建立了一种相邻时间片完全时间不对称数据的 DBN 学习方法.该方法通过引入传递变量来建立相邻时间片变量之间的联系和信息传递通道,并基于节点排序和局部打分-搜索进行 DBN 局部学习,再沿时序延展来建立整个 DBN 结构.如此学习得到的 DBN 结构不但实现了相邻时间片之间的信息时序传递,而且反映了网络随时间变化而演化的情况.

参考文献(References)

- [1] Friedman N, Murphy K, Russell S. Learning the structure of dynamic probabilistic networks[C]. Proc of the 14th Int Conf on Uncertainty in Artificial Intelligence. Madison, 1998: 139-147.
- [2] 吕国云,蒋冬梅,张艳宁,等.基于多流多状态动态贝叶斯网络的音视频连续语音识别[J].电子与信息学报,2008,30(12): 2906-2911.
(Lv G Y, Jiang D M, Zhang Y N, et al. DBN based multi-stream multi-states model for continue audio-visual speech recognition[J]. J of Electronics and Information Technology, 2008, 30(12): 2906-2911.)
- [3] Laurence L S, Marc S. Recognition of degraded characters using dynamic Bayesian networks[J]. Pattern Recognition, 2008, 41(10): 3092-3103.
- [4] 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[J].自动化学报,2009,35(3): 225-232.
(Du Y T, Chen F, Xu W L. Approach to human activity multi-scale analysis and recognition based on multi-layer dynamic Bayesian network[J]. Acta Automatica Sinica, 2009, 35(3): 225-232.)

(下转第1746页)