小样本贝叶斯网络结构学习的KDE-CGA算法

许建锐 李战武 徐 安

(空军工程大学航空航天工程学院 西安 710038)

摘 要 针对小样本数据条件下的贝叶斯网络结构学习,首先利用核密度估计(Kernel Density Estimation,KDE)对小规模样本数据进行拓展,然后引用云遗传算法(Cloud Theory-based Genetic Algotithm,CGA)对贝叶斯网络结构进行学习。通过优化改进核密度函数及其窗宽提高数据拓展效果;通过将云理论引入遗传算法中,自适应地改变交叉率和变异率,避免了算法局部寻优问题。仿真结果验证了该算法的有效性。

关键词 小样本,贝叶斯网络,结构学习,核密度估计,云遗传算法

中图法分类号 TP181 **文献标识码** A

KDE-CGA Algorithm of Structure Learning for Small Sample Data Bayesian Network

XU Jian-rui LI Zhan-wu XU An

(Aeronautics and Astronautics Engineering College, Air Force Engineering University, Xi'an 710038, China)

Abstract In view of learning the Bayesian network under the condition of the small sample data, this paper firstly made use of kernel density estimation to expand the small scale sample data, then adopted the cloud theory-based genetic algotithm to learn the structure of Bayesian network. In order to improve the effect of data expanding, the paper discussed the way of improving the density function and its window breadth. At the same time, the cloud theory was combined with genetic algotithm. We changed crosses rate and variation rate properly, avoided the problem of looking an excellent answer in a part. Simulation results show that the algorithm is effective and practical.

Keywords Small sample data, Bayesian network, Structure learning, Kernel density estimation, Cloud theory-based genetic algorithm

1 引言

贝叶斯网络在数据挖掘方面有着重要应用,它能够用节点表示实际问题中的随机变量,将变量之间的因果关系直观地体现出来,在解决不确定性复杂问题方面有着独特的优势。

从数据中挖掘形成贝叶斯网络主要分为结构学习和参数学习[1]。结构学习是结合先验知识,基于训练样本,寻找有最大后验概率的贝叶斯网络结构的过程[2];参数学习是在建立贝叶斯网络结构学习的基础上,学习得到各节点的条件概率。其中,学习贝叶斯网络的训练样本一般都是完备的大规模的数据集,但在很多情况下,尤其是在战场态势瞬息万变的空战条件下,需要快速进行态势感知,做出决策,这时可以利用的样本规模较小,样本中所表达的信息不够完整,导致难以保证贝叶斯网络结构学习的准确性和可靠性。因此给出小样本条件下贝叶斯网络结构学习的方法研究。

数据不完整使得学习贝叶斯网络更加复杂且困难。缺失部分数据会导致以下两个方面问题:1)用于评价网络的积分函数不再具有可分解形式,因此不能进行局部搜索;2)由于部分统计因子之间的因果关系更加难以确定,导致学习效率低下。目前,主要有两类方法对小样本条件下的贝叶斯网络结构进行学习,一类是设计改进新的算法直接进行学习,如Friedmen率先提出的SEM算法[3],其思路是在每一次迭代中只对结构变化的局部进行评分,再从当前条件下选取最优的网络结构进行下一轮迭代,直到结构趋于收敛;另一类是通过拓展数据集来增大样本规模[4],提高结构学习的可靠性。拓展数据的方式一般是Bootstrap抽样[5-6],但这种方式是从数据集中进行重复抽样,未增加额外的信息,难以提高学习效果。本文对文献[7]中采用的密度核估计进行改进,使得拓展数据更加符合样本数据的信息,而且采用云遗传算法(Cloud Theory-based Genetic Algotithm,CGA)进行结构学习,避免了K2算法需要先确定变量顺序的问题。

2 核密度函数的改进

假设从一维总体 X 中抽出独立同分布样本 $X_1, X_2, \cdots, X_n$,但 X 的密度函数 $f(x)$ 未知,$x \in R$,则 $f(x)$ 的密度核估计为:

$$f_n(x)=\frac{1}{nh_n}\sum_{i=1}^{n}K\left(\frac{x-X_i}{h_n}\right) \tag{1}$$

其中,$K(\ )$ 为 R 上的核函数或Borel可测函数的窗,h_n 是窗宽,其为正数并且与 n 有关。给定样本数据后,核函数和窗宽的选择直接影响着核估计性能的好坏,其中,核函数影响着核估计的精度,很多文献[8-10]都对核函数加以限制,使得估计量

许建锐(1994—),男,硕士生,主要研究方向为火力指挥控制与评估工程;李战武(1978—),男,副教授,硕士生导师,主要研究方向为航空武器火力控制原理与技术、作战决策;徐 安(1982—),男,讲师,主要研究方向为航空武器火力控制原理与技术、作战决策。

与被估计量的偏差尽可能变小，从而改进估计的精度。一般要求核函数 $K(\ )$ 是非负的、连续的、有界的、对称的等。窗宽 h_n 随 n 的增大而变小，即随着 n 趋于无穷大而逐渐趋于 0；若 h_n 太大，会使得经过 $(x-X_i)/h_n$ 压缩变换之后的 x 的平均化作用突出，从而将密度的细节部分掩盖；反之，若 h_n 太小，则会导致随机性的影响增加，可能掩盖了 $f(x)$ 的部分重要特性，因此选择一个恰当的 h_n 十分重要。

2.1 核函数的迭代

常用的核函数有很多，如 Parzen 核、Epanechikov 核和 Exponent 核等。在解决实际问题时，往往先利用以上多个核函数逐一进行估计，然后根据结果再决定选取的核函数，但目前构造核函数时并没有一个完善的方法，本文在选取核函数时采用迭代算法。

为了使迭代顺利进行，首先要求核函数必须是偶函数，$x\geqslant 0$ 时，$K(x)$ 是非增函数；其次核函数必须是概率密度函数，即 $\int K(t)\mathrm{d}t=1$。由于式(1)满足非负性、在 $(-\infty,+\infty)$ 上有界和在 $(-\infty,+\infty)$ 的积分有界等条件，可以将密度核估计值式(1)作为新的核函数进行迭代，得到下一步的估计值。但 $f_n(x)$ 并非偶函数，因此需要通过适当处理将 $f_n(x)$ 延拓成偶函数[11]。

假定选取均匀核作为密度核估计的核函，即：

$$K(x)=\begin{cases}0.5, & x\in[-1,1]\\ 0, & x\notin[-1,1]\end{cases} \tag{2}$$

则

$$\begin{aligned}E(x)&=\int_{-\infty}^{+\infty}xf_n(x)\mathrm{d}x\\&=\int_{-\infty}^{+\infty}x\frac{1}{nh_n}\sum_{i=1}^{n}(\frac{x-X_i}{h_n})\mathrm{d}x\\&=\frac{1}{nh_n}\sum_{i=1}^{n}\int_{X_i-h_n}^{X_i+h_n}x\cdot\frac{1}{2}\mathrm{d}x+\frac{1}{nh_n}\sum_{i=1}^{n}[\int_{-\infty}^{X_i-h_n}+\int_{X_i+h_n}^{+\infty}]\\&\quad x\cdot 0\mathrm{d}x\\&=\frac{1}{2nh_n}\sum_{i=1}^{n}\frac{1}{2}X^2\Big|_{X_i-h_n}^{X_i+h_n}\\&=\frac{1}{4nh_n}\sum_{i=1}^{n}4X_ih_n\\&=\bar{X}\end{aligned} \tag{3}$$

由上式可得，将样本值平移 $\bar{X}$ 便可使得 $f_n(x)$ 变换成偶函数。用 $f_{n1}(x)$ 表示，则：

$$f_{n1}(x)=\frac{1}{nh_n}\sum_{i=1}^{n}K(\frac{x-(X_i-\bar{X})}{h_n}) \tag{4}$$

此时，$f_{n1}(x)$ 为偶函数，便可利用其作为新的核函数进行迭代，迭代后的 $\hat{f}_{n1}(x)$ 为：

$$\hat{f}_{n1}(x)=\frac{1}{n^2h_n^2}\sum_{j=1}^{n}\sum_{i=1}^{n}K(\frac{(x-X_j)+h_n(\bar{X}-X_i)}{h_n^2}) \tag{5}$$

式(5)的意义为：当从 n 个样本中任意选取一个样本时，上式便在该样本附近有规则地产生 n 个新的样本，原来的 n 个样本则拓展成了 n^2 个，之后得到 n^2 个估计值的平均值。上述迭代算法可以认为是在每个样本附近都模拟总体的概率密度进行估计，得到更多的再生数据求平均值，因此估计值应比式(1)更加准确。由于常用密度估计值的拟合效果较好，而且迭代所需要的使用时间和占用的电脑内存较大，因此实际中只做一次迭代即可。

2.2 最优窗宽的选择

理论上来讲，最优窗宽的寻优是利用了密度估计值和真实密度之间的差值[12]，但在这过程中用到了总体的真实密度函数。假如总体分布的密度函数已知，那么对其进行估计就是多此一举了。窗宽的选择应考虑样本数据的密集程度，其计算方法有很多，这里采用递归方法展开探索。

首先根据经验，选择一个恰当的窗宽值 h_1，利用式(5)计算出样本总体的密度核估计：

$$\hat{\hat{f}}_{n1}(x)=\frac{1}{n^2h_1^2}\sum_{j=1}^{n}\sum_{i=1}^{n}K(\frac{(x-X_j)+h_1(\bar{X}-X_i)}{h_1^2}) \tag{6}$$

假定式(6)中的 $\hat{\hat{f}}_{n1}(x)$ 是样本的真实密度，且 $\hat{\hat{f}}_{n1}(x)''$ 在 $(-\infty,+\infty)$ 上处处连续且有界，则计算样本的积分均方误差 $MISE$ 为：

$$MISE(h)=E[\int(\hat{f}_{n1}(x)-\hat{\hat{f}}_{n1}(x))^2\mathrm{d}x] \tag{7}$$

为了使式(7)达到最小，可以通过计算 $MISE$ 的无偏估计进行求解。显然，$ISE(h)=\int(\hat{f}_{n1}(x)-\hat{\hat{f}}_{n1}(x))^2\mathrm{d}x$ 是 $MISE$ 的一个无偏估计：

$$\begin{aligned}ISE(h)&=\int(\hat{f}_{n1}(x)-\hat{\hat{f}}_{n1}(x))^2\mathrm{d}x\\&=\int\hat{f}_{n1}^2(x)\mathrm{d}x-2E[\hat{f}_{n1}(x)]+\int\hat{\hat{f}}_{n1}^2(x)\mathrm{d}x\end{aligned} \tag{8}$$

其中，$\int\hat{\hat{f}}_{n1}^2(x)\mathrm{d}x$ 与窗宽 h 无关，于是 $\int\hat{f}_{n1}^2(x)\mathrm{d}x-2E[\hat{f}_{n1}(x)]$ 取得最小时便可得到 h_0，再令 $h_2=h_0$ 代入式(5)，得到样本总体新的密度核估计：

$$\hat{\hat{f}}_{n1}(x)=\frac{1}{n^2h_2^2}\sum_{j=1}^{n}\sum_{i=1}^{n}K(\frac{(x-X_j)+h_2(\bar{X}-X_i)}{h_2^2}) \tag{9}$$

再假定式(9)中的 $\hat{\hat{f}}_{n1}(x)$ 是样本的真实密度，计算得到新的最优窗宽 h_3，经过如此多次的递归运算，h_n 将趋于一个稳定值，该值便可以作为样本总体的最优窗宽。

2.3 验证改进后的核密度估计

选用均匀核作为初始核，$h_n=0.6$，$n=120$ 时比较改进后的核密度估计与改进前的图形，可以看出迭代后的图像拟合得更好一些。

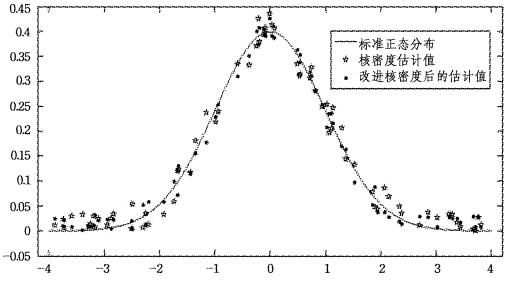


图 1 选用均匀核时核密度估计改进效果图

3 基于数据拓展的贝叶斯网络结构学习的算法步骤

贝叶斯网络结构学习可以看作是在巨大的搜索空间中寻优的过程，目的是找到与样本数据最匹配的网络结构，很多文

献[13-15]都对此做出了有益探索，其中遗传算法作为一种模拟生物进化过程的全局优化算法，得到了广泛关注。但由于一般的遗传算法使用的交叉率和变异率是不变的，对其在全局搜索的能力有了较大的限制，本文在遗传算法的基础上引入云模型理论，既保留了传统遗传算法在随机搜索方面的优势，又通过利用云模型的随机性和稳定倾向性的特点，适当改变遗传算法的交叉率和变异率，为更好地寻找最优贝叶斯网络结构提供了思路。

3.1　小样本的数据拓展

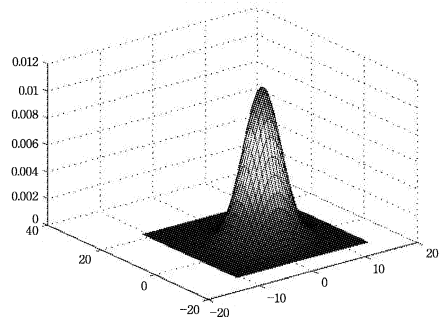
构建贝叶斯网络所需要的数据集一般是多维数据，网络节点决定着数据维度的大小；同时，由于贝叶斯网络的节点之间存在着一定的因果联系，因此各维度之间也有着一定的关联性。因此必须将样本看作多维自耦合的整体进行数据拓展，从而进行贝叶斯网络结构的学习。

假定需要学习的贝叶斯网络结构中包含的节点数为 N，则样本数据 X 总体是 N 维变量，设 $X_1, X_2, \cdots, X_n$ 为总体数据 X 中的样本，则根据上文所述可得到 X 的概率密度函数 $f(X)$ 的核估计为：

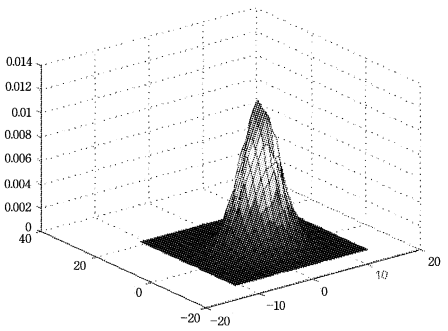
$$\hat{f}_{n1}(x)=\frac{1}{n^2 h_n^{2N}\det(S)^{1/2}}\sum_{j=1}^{n}\sum_{i=1}^{n} K\left(\frac{(X-X_j)^{\mathrm{T}}S^{-1}(X-X_j)+h_n^2(\bar{X}-X_i)^{\mathrm{T}}S^{-1}(\bar{X}-X_i)}{h_n^4}\right) \tag{10}$$

其中，$X=(x_1, x_2, \cdots, x_N)^{\mathrm{T}}$，$X_i=(x_{i1}, x_{i2}, \cdots, x_{iN})^{\mathrm{T}}$ $(i=1, 2, \cdots, n)$；n 为样本大小；h 为窗宽；S 是 $N\times N$ 维对称样本协方差矩阵[16]。

为验证多维数据核估计效果，本文采用二维数据样本，使用的核函数是标准高斯函数 $f(x)=\exp(-x^2/2)/\sqrt{2\pi}$，取均值矩阵 $M=\begin{bmatrix}2\\3\end{bmatrix}$ $S=\begin{bmatrix}10 & 2\\2 & 20\end{bmatrix}$，得到仿真结果如图 2 所示。



(a)理想概率密度函数图



(b)parsen 窗函数估计概率密度函数图

图 2　二维数据样本核估计效果

从图 2 可以看出，以二维数据样本为例的多维数据核估计取得了良好效果，概率密度函数比较贴近给定函数。

3.2　云遗传算法

3.2.1　定义

在解决实际问题对，全局最优解附近一般都会存在某个邻域，在该邻域内的其他解逐渐逼近最优解。假如当前解的适应度较小，则选择搜索的邻域就会较大，相反则搜索的邻域较小，进而逐步找到最优解的区域，最终求得最优解。

云遗传算法仍然采用一般遗传算法的交叉操作和变异操作，不同的是，由云模型的正态云生成算法实现变异操作，Y 条件云生成实现交叉操作。

3.2.2　云自适应交叉率与变异率

遗传算法中的变异操作反映了个体中某个基因在一定范围内的突变，而交叉操作实现了个体的整体变化。为避免一般的遗传算法易陷入局部最优和出现早熟收敛的情况，本文在计算出适应度的基础上，利用云模型的云发生器自适应地产生变异率和交叉率，既保留了一般遗传算法的快速寻优能力，又具有随机性，产生了避免陷入局部寻优的能力。自适应产生交叉率与变异率的方法如下。

(1)云自适应交叉率 p_c：

$$\begin{aligned}&Ex=f_{avg}\\&En=(f_{\max}-f_{avg})/c_1\\&He=En/c_2\\&En'=RANDN(En, He)\\&p_c=\begin{cases}k_1\mathrm{e}^{\frac{-(f'-Ex)^2}{2(En')^2}}, & f'\geqslant f_{avg}\\ k_3, & f'<f_{avg}\end{cases}\end{aligned} \tag{11}$$

(2)云自适应变异率 p_m：

$$\begin{aligned}&Ex=f\\&En=(f_{\max}-f)/c_3\\&He=En/c_4\\&En'=RANDN(En, He)\\&p_m=\begin{cases}k_2\mathrm{e}^{\frac{-(f-Ex)^2}{2(En')^2}}, & f\geqslant f_{avg}\\ k_4, & f<f_{avg}\end{cases}\end{aligned} \tag{12}$$

云自适应产生的交叉率 p_c 或变异率 p_m 如图 3 所示。

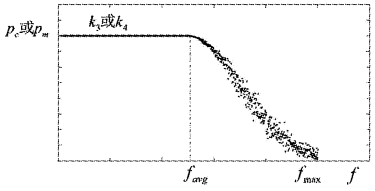


图 3　云自适应产生的交叉率、变异率

3.2.3　云遗传算法的基本步骤

由于云遗传算法是基于遗传算法的思路解决问题的，因此同样包含参数编码、产生初始种群、确定适应度函数、确定遗传操作和确定控制参数等 5 个部分的基本操作。

(1)初始化种群

通常，初始种群的选取与算法的收敛速度有着一定联系。一般的遗传算法是在搜索范围内随机产生个体，形成初始种

群，但这样并不利于提高算法的收敛速度。本文在种群的选择上，通过文献[17]中的最大权重决策树算法进行初始化，生成一个贝叶斯网络，并对该网络的边集在一定概率下进行反向、减边、加边操作，从中选择合理的网络作为初始种群。

(2)计算适应度值

适应度函数是体现种群中个体优劣的衡量标准，适应度越高代表个体越优秀，可以选用合适的评分函数作为适应度函数，用评分的数值大小表示个体在搜索中的好坏。本文适应度函数选用 BIC 评分函数[14]。

BIC 评分是样本最边缘似然函数的一种近似，其明确直观，使用方便，公式为：

$$Score_{BIC}(S|D)=\ln\prod_{i=1}^{n}\prod_{j=1}^{q_i}\prod_{k=1}^{r_i}(\theta_{ijk}^{*})-\frac{1}{2}\ln m\times dim(S) \tag{13}$$

其中，$dim(S)$是贝叶斯网络结构 S 中独立参数的个数，即 $dim(S)=\sum_{i=1}^{n}q_i(r_i-1)$。

(3)选择、复制和遗传

1)把最佳个体复制到下一代；

2)选择比较好的种群，并复制；

3)淘汰最差个体，并随机产生一个外来个体。

(4)交叉操作

(5)变异操作

(6)判断操作

判断是否满足终止条件，若不满足，则转至步骤(2)；反之，则输出网络结构。

在本文的算法中，若总迭代次数超过最大迭代次数或者最大适应度值连续 t 代不变，则算法结束；否则采用新一代个体重复选择、交叉、变异的过程。

4 仿真设计与结果分析

以贝叶斯网络评估空空导弹命中效果为例，学习验证本文描述算法的可行性。

中远程空空导弹能否有效命中目标，受信息域、物理域和认知域中多方面因素的影响，如图 4 所示。

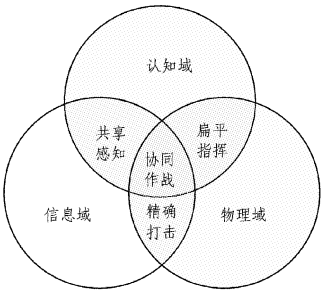


图 4 协同作战条件下作战作用域

物理域是交战双方在空中和地面进行机动、探测及交战的领域，即雷达(包括机载雷达和地面雷达)探测目标、飞机起飞接敌、空空导弹发射、制导、控制等，也是作战平台和连接平台通信网存在的领域；信息域是对目标搜索、跟踪、识别和融合的作战领域，是生成、处理并共享信息的领域；认知域是作战指挥员、飞行员执行战术制定、攻击决策等作战任务的作战领域。经过分析，中远程导弹命中效果评估的全因素贝叶斯模型如图 5 所示。

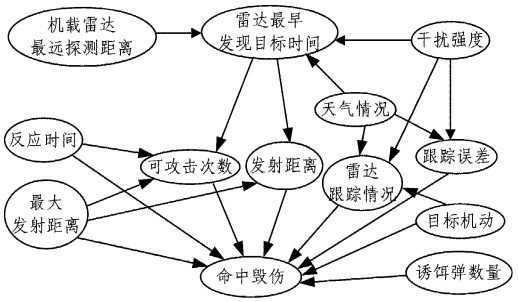


图 5 中远程导弹命中效果评估的全因素贝叶斯模型

本次实验过程中选取其中的 7 个因素，如表 1 所列，并对每个因素的取值进行离散化，得到两个状态。

表 1 网络节点数据设计表

序号	作战想定因素	实验值
1	机载雷达最远探测距离/km	200,300
2	雷达最早发现目标时间/s	20,30
3	干扰强度	0,1
4	反应时间/ms	240,330
5	可攻击次数	0,1
6	发射距离/km	60,100
7	雷达跟踪情况	0,1
8	命中毁伤情况	0,1

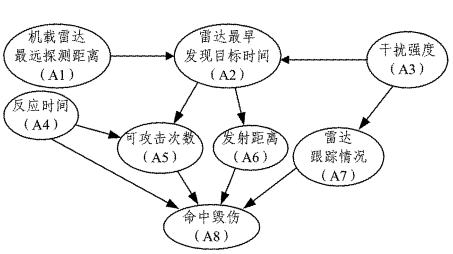
基于文中设定的贝叶斯网络结构，使用 Matlab 中贝叶斯网络工具箱[18]的 sample_bnet 函数，便可按照节点顺序依次进行采样，产生规模为 N 的小数据集；然后基于数据集使用给定节点顺序的 K2 算法以及改进的概率核密度估计的方法分别进行数据拓展；之后利用云遗传算法进行结构学习；最后进行比较。

实验环境为：硬件为 Intel(R) Core(TM) i5-4460T 1.90 GHz；操作系统为 Windows 7；编程软件为 MatlabR2012a。

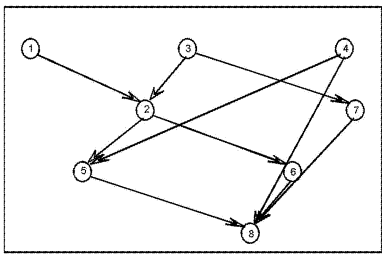
用 *ME* 表示经过学习的贝叶斯结构相比原网络结构缺失的边，*IE* 表示经过学习的贝叶斯结构比原网络结构多余的边，*RE* 表示学习结构与原网络结构相比反转的边，本文在原始样本数为 $N=200,500,700$ 的条件下分别进行实验，为减小偶然误差，实验重复进行 10 次，取平均值作为实验结果。不同样本条件下结构学习的结果比较如表 2 所列。K2 算法和本文算法结构学习效果比较如图 6 所示。

表 2 不同样本条件下结构学习的结果比较

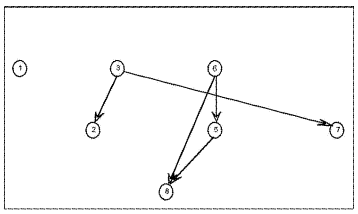
	N=200		N=500		N=700	
	K2 算法	本文算法	K2 算法	本文算法	K2 算法	本文算法
IE	0.3	0.1	0.2	0	0	0
ME	4.2	3.1	3.2	1.4	1.9	0.3
RE	0	0.3	0	0.1	0	0
汉明距离	4.5	3.5	3.4	1.5	1.9	0.3



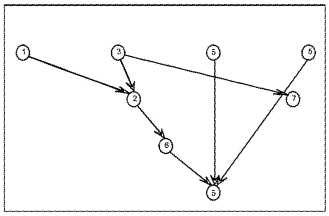
(a)设定的网络结构



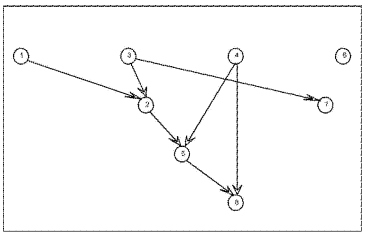
(b)设定的网络结构



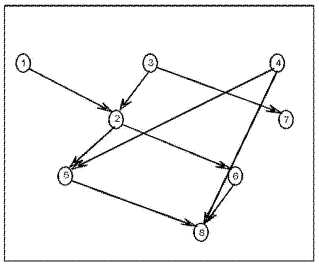
(c)学习后的网络结构(N=200)



(d)改进学习后的网络结构(N=200)



(e)学习后的网络结构(N=500)



(f)改进学习后的网络结构(N=500)

图 6　K2 算法和本文算法结构学习效果比较

其中，汉明距离＝丢失的边(IE)＋反转的边(RE)＋多余的边(ME)。

由实验结果可知：总体来讲，在小样本数据条件下学习贝叶斯网络结构时，本文算法较给定节点顺序时的 K2 算法效果更好，但由于 K2 算法中指定了节点顺序，因此不会产生反向的边，而本文算法在这一方面的学习效果不如 K2 算法。

结束语　本文提出了一种小数据集条件下的贝叶斯网络结构学习的方法，该方法能够对样本数据进行合理的拓展，得到比较大的样本，而后利用云遗传算法对网络结构进行学习。通过仿真验证可知，本文提出的算法具有较强的结构学习能力，改善了小样本条件下 K2 算法学习贝叶斯网络的结构的不足，下一步应在此基础上对参数的学习效果进行分析研究，以更加完善地对贝叶斯网络进行学习。

参 考 文 献

[1] 史志富，张安. 贝叶斯网络理论及其在军事系统中的应用[M]. 北京：国防工业出版社，2012：28-52.

[2] 邸若海，高晓光. 基于限制型粒子群优化的贝叶斯网络结构学习[J]. 系统工程与电子技术，2011，33(11)：423-427.

[3] FRIEDMAN N. The Bayesian structural EM algorithm[C]// San Francisco. CA，USA，1998：129-138.

[4] BORCHANI H，AMOR N B，KHALFALLAH F. Learning and evaluating Bayesian Network equivalence classes from incomplete data[J]. International Journal of Pattern Recognition and Artificial Intelligence，2008，22(2)：253-278.

[5] 徐达明，唐安民，等. 概率密度核估计的 Bootstrap 逼近[J]. 云南民族大学学报(自然科学版)，2007，16(4)：295-298.

[6] 刘伟，龙琼，陈芳，等. Bootstrap 方法的几点思考[J]. 飞行器测控学报，2007，26(5)：78-81.

[7] 韩绍金，李建勋. 基于密度核估计的贝叶斯网络结构学习算法[J]. 计算机工程与应用，2014，50(15)：107-112.

[8] 王金然，郭亚君，吕金凤. 一种改进的密度核估计算法[J]. 大学数学，2008，24(6)：67-71.

[9] RAOP. NonparametricFunctionEstimation[M]. AcademiePress，Inc. 1983.

[10] 朱亚培. 密度核估计的改进及其相关问题的讨论[D]. 兰州：兰州交通大学，2015.

[11] BASHTANNYK D M，ROB，HYNDMAN J. Bandwidth selection for kernel conditional desity estimation [J]. Computation Tatistics & Data Nalysis，2001(36)：63-78.

[12] 王星. 非参数统计[M]. 北京：中国人民大学出版社，2005：213-218.

[13] 郭童. 基于改进鱼蜂群算法的贝叶斯网络结构学习[D]. 杭州：浙江大学，2014

[14] 肖秦琨，高嵩. 贝叶斯网络在智能信息处理中的应用[M]. 此京：国防工业出版社，2012：5-20.

[15] SHETTY S，SONG M. Structure learning of Bayesian network using a semantic genetic algorithm-based approach[C]// Proceedings of the 3rd International Conference on Information Technology：Research and Education. Hsinchu，2005：454-458.

[16] 李德旺，陈兴，喻达磊，等. 多维密度核估计的 Bootstrap 逼近[J]. 西南大学学报(自然科学版)，2007，29(11)：34-37.

[17] CHOW C，LIU C. Approximation Discrete probability Distributions with Dependence Trees [J]. IEEE Transactions on Information Theory，1968，14(3)：462-467.

[18] DAVIDSON-PILON C. 贝叶斯方法概率编程与贝叶斯推断[M]. 辛愿，钟黎，等译. 北京：人民邮电出版社，2017.