

基于拓扑序列和量子遗传算法的贝叶斯网结构学习

赵学武^{1*}, 刘广亮¹, 程新党¹, 冀俊忠²

(1. 南阳师范学院 软件学院, 河南 南阳 473061; 2. 北京工业大学 计算机学院, 北京 100124)

(* 通信作者电子邮箱 zhaoxuewuyonghu@163.com)

摘要: 贝叶斯网是处理不确定性问题知识表示和推理的最重要的理论模型之一, 其结构学习是目前研究的一个热点。提出了一种基于拓扑序列和量子遗传算法的贝叶斯网结构学习算法, 新算法首先利用量子信息的丰富性和量子计算的并行性, 设计出基于量子染色体的拓扑序列生成策略提高了搜索效率, 并为 K2 算法学得高质量的贝叶斯网结构提供了保障; 然后采用带上下界的自适应量子变异策略, 增强了种群的多样性, 提高了算法的搜索能力。实验结果表明, 与已有的一些算法相比, 新算法不仅能获得较高质量的解, 而且还有着较快的收敛速度。

关键词: 贝叶斯网; 结构学习; 量子遗传算法; K2 算法; 拓扑序列; 量子计算

中图分类号: TP181 **文献标志码:** A

Bayesian network structure learning algorithm based on topological order and quantum genetic algorithm

ZHAO Xuewu^{1*}, LIU Guangliang¹, CHENG Xindang¹, JI Junzhong²

(1. School of Software, Nanyang Normal University, Nanyang Henan 473061, China;

2. College of Computer Science and Technology, Beijing University of Technology, Beijing 100124, China)

Abstract: Bayesian network is one of the most important theoretical models for the representation and reasoning of uncertainty. At present, its structure learning has become a focus of study. In this paper, a Bayesian network structure learning algorithm was developed, which was based on topological order and quantum genetic algorithm. With the richness of the quantum information and the parallelism of quantum computation, this paper designed generator strategy of topological order based on a quantum chromosome to improve not only the efficiency of search, but also the quality of Bayesian network structure. And then by using self-adaptive quantum mutation strategy with upper-lower limit, the diversity of the population was increased, so that the search performance of the new algorithm was improved. Compared to some existing algorithms, the experimental results show that the new algorithm not only searches higher quality Bayesian structure, but also has a quicker convergence rate.

Key words: Bayesian network; structure learning; quantum genetic algorithm; K2 algorithm; topological order; quantum computing

0 引言

不确定性问题知识表示和推理是人工智能领域中的一个研究热点, 贝叶斯网 (Bayesian Network, BN) 是处理该问题的一个非常重要的理论模型。近年来, 随着搜索技术的发展和数据挖掘的兴起, 贝叶斯网结构的学习引起了国内外学者的广泛兴趣。到目前为止, 人们已经提出了一些学习贝叶斯网结构的方法, 其中基于蚁群算法^[1]、遗传算法^[2-4]和粒子群算法^[5]的贝叶斯网结构学习是比较新的一些实用而有效的方法。例如, 基于蚁群算法的贝叶斯网结构学习算法 I-ACO-B^[1]首先用 0 阶条件独立性测试发现一些潜在的条件——独立知识并用之压缩搜索空间, 然后利用改进的启发函数使蚁群算法的搜索能力得到提高。文献[5]提出了一种新的基于粒子群的学习贝叶斯网结构的算法——C-PSO-B, 该算法利用定义的规则链模型度量拓扑序列优劣, 有利于发现较高质量的拓扑序列; 然后通过给粒子位置可选择更新的粒子群优

化算法加上动态权重系数, 提高了算法的搜索性能。这两种算法都取得了比较好的成果, 但是它们在学习贝叶斯网结构时仍存在求解精度不高、收敛速度慢等不足。

量子计算是 20 世纪 90 年代被提出的。由于它是以微观世界为物理研究基础, 所以量子计算具有了经典计算无可比拟的优越性。在量子理论中, 量子信息的丰富性、量子计算的并行性和量子态的叠加性等特性, 使量子计算具有了明显的优势。将经典的搜索算法与量子计算相结合是改善上述不足的一种研究思路。到目前为止, 已出现了一些与量子相结合的搜索算法^[6-11], 如量子进化算法^[6-7]、量子遗传算法^[8-10]和量子粒子群算法^[11]等。本文提出了一种基于拓扑序列和量子遗传算法的贝叶斯网结构学习算法 (Bayesian Network Structure Learning Algorithm based on Topological Order and Quantum Genetic Algorithm, T&QGA-B)。为了提高搜索的效率, 新算法首先利用量子计算的优势, 设计出基于量子染色体的拓扑序列生成策略; 然后通过采用带上下界的自适应量子

收稿日期: 2012-12-10; 修回日期: 2013-02-25。 基金项目: 河南省基础与前沿技术研究计划项目 (122300410302, 122300410426, 122300410224, 122300410384, 132300410210); 南阳师范学院青年项目 (QN2010010)。

作者简介: 赵学武 (1983-), 男, 河南南阳人, 助教, 硕士, 主要研究方向: 群集智能、数据挖掘; 刘广亮 (1978-), 男, 安徽阜阳人, 讲师, 硕士, 主要研究方向: 智能信息处理、网络安全; 程新党 (1976-), 男, 河南南阳人, 副教授, 硕士, 主要研究方向: 互联网技术; 冀俊忠 (1969-), 男, 山西晋中人, 教授, 博士生导师, 主要研究方向: 机器学习、Web 智能。

变异策略,增强了种群的多样性。实验结果表明,与新近提出的 I-ACO-B 算法^[1]和 C-PSO-B 算法^[5]相比,新算法在求解精度和收敛速度两方面都有明显提高。

1 贝叶斯网和量子遗传算法

1.1 贝叶斯网

一个贝叶斯网模型可以用一个二元组表示 $B = \langle G, P \rangle$, 其中, G 是一个有向无环图(Directed Acyclic Graph, DAG), 表示了贝叶斯图形结构要素, G 又可以表示为一个二元组 $\langle V, E \rangle$, 这里 V 是有向图节点(和问题域中的随机变量对应)的集合; E 是有向边的集合, 表达了随机变量间直接的依赖关系。因此 G 是对问题域中随机变量间独立—依赖关系的结构性的定性描述。 P 表示概率参数(也称为条件概率参数表)要素, 是贝叶斯网中各个节点在已知其父节点集时各种取值组合下的条件概率的集合, 它可以表示为:

$$P = \{p(X_i | \Pi(X_i)), X_i \in X\}$$

其中 $\Pi(X_i)$ 表示 X_i 的父节点集。因此, P 是在已知贝叶斯网中随机变量间独立—依赖关系的情况下对这种独立—依赖程度的定量描述。因此, 一个完整的贝叶斯网 $\langle G, P \rangle$ 就唯一地确定了在随机变量集合 $X = \{X_1, X_2, \dots, X_n\}$ 上的一个联合分布:

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n p(X_i | \Pi(X_i)) \quad (1)$$

1.2 贝叶斯网结构学习

简单地讲, 贝叶斯网结构的学习过程是在给定一个样本数据集 D 的条件下, 寻求一个能够最佳匹配样本数据集 D 的 DAG G 的过程。该图形结构隐含地表达了节点变量间的部分条件独立性。随着计算技术的发展和理论研究的深入, 贝叶斯网结构学习的研究也经历了一个过程, 其学习方法可以概括为以下两种: 1) 基于随机变量间约束满足的学习方法。它将学习过程看作是约束不断满足的过程, 即在给定的数据集 D 下, 利用信息论的知识测试节点变量间的条件独立性关系, 努力寻找与 D 所蕴含的独立—依赖关系尽可能一致的贝叶斯网结构。该方法的实现简单, 但是随着阶数的增大, 测试数急剧膨胀, 并且测试也越来越不可靠, 尤其是在贝叶斯网的规模较大时, 故得到的贝叶斯网结构的精度难以保证。2) 基于评分搜索的学习方法。在给定数据集下, 选用一个评分标准来判定所生成的贝叶斯网结构与样本数据集的匹配程度, 搜索评分最高的网络结构。该方法由于使用了科学的评分准则, 所以求解的精度可以得到保证, 但是使用此方法搜索节点的父节点集时产生了许多无效搜索, 也浪费了不少时间。由于上述两种方法都有各自的优缺点, 所以将这两种方法有机结合起来成为一种新的研究思路。到目前为止, 已出现了一些基于这种思路的学习贝叶斯网结构的方法, 如基于蚁群算法的学习方法^[1,12]和基于粒子群算法的学习方法^[5]等。本文提出的新算法 T&QGA-B 也属于这种混合型的学习方法。

1.3 量子计算

量子比特是存储量子信息的基本单位, 它具有 $|0\rangle$ 和 $|1\rangle$ 两种基本状态。量子比特与经典比特的区别之处就在于它的状态是不仅仅是这两个基本状态, 而还可以是这两个基本状态的叠加态, 可以表示为这两个基本状态的线性组合, 即:

$$|\varphi\rangle = \alpha|0\rangle + \beta|1\rangle \quad (2)$$

也可以表示为:

$$|\varphi\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \quad (3)$$

其中: α 和 β 都是复数, 称为量子态的概率幅。当我们观测时, 量子比特或者以概率 $|\alpha|^2$ 坍塌到 0, 或者以概率 $|\beta|^2$ 坍塌到 1, 并且满足如下归一性条件:

$$|\alpha|^2 + |\beta|^2 = 1 \quad (4)$$

量子比特与经典比特的区别在于经典比特只有 0 和 1 两种状态, 即 2^1 状态; 而量子比特除了 $|0\rangle$ 和 $|1\rangle$ 两种状态外, 还可以落在 $|0\rangle$ 和 $|1\rangle$ 的任意线性组合态上。类似地, 一个 n 量子比特系统可以表示为:

$$|\varphi\rangle = \sum_{i=1}^{2^n} C_i |\varphi_i\rangle \quad (5)$$

其中: C_i 表示该 n 量子比特系统基本态 $|\varphi_i\rangle$ 的概率幅, $|C_i|^2$ 表示观测时坍塌到 $|\varphi_i\rangle$ 的概率, 也是对应的观测到结果为 i 的概率, 并且也满足如式(6)所示的归一化条件:

$$\sum_{i=1}^{2^n} |C_i|^2 = 1 \quad (6)$$

从式(6)可以看到 n 量子比特系统的状态处于 2^n 个基态的叠加态之中, 也就是说它可以同时表示 2^n 个数。因此, n 位的量子寄存器表示的信息要比 n 位的经典信息系统要丰富得多, 它更具有解决比较复杂问题优势。量子系统的另一个优势是量子计算的并行性, 这个优势可以降低算法的搜索时间。一个 n 位的量子比特系统中存储了 2^n 个经典态, 这样量子计算机对叠加态的演化也就是对其叠加成分的演化, 即是对 2^n 个经典态进行演化, 这就是量子计算的并行性^[10]。

量子旋转门是量子计算中一个非常重要的量子门, 常被用于量子系统更新。其一般形式如式(7)所示:

$$U = \begin{bmatrix} \cos \Delta\theta & -\sin \Delta\theta \\ \sin \Delta\theta & \cos \Delta\theta \end{bmatrix} \quad (7)$$

以单量子比特为例, 更新过程如下:

$$\begin{bmatrix} \alpha' \\ \beta' \end{bmatrix} = \begin{bmatrix} \cos \Delta\theta & -\sin \Delta\theta \\ \sin \Delta\theta & \cos \Delta\theta \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \quad (8)$$

其中: $[\alpha \ \beta]^T$ 表示单量子比特, 也可以是量子染色体的某个量子比特位; $\Delta\theta$ 为旋转角。

2 本文算法

2.1 算法思想

如前所述, 基于约束满足的贝叶斯网结构学习方法是通过对条件独立性测试得到潜在的条件——独立性知识进而得到相应的贝叶斯网结构, 其缺点是精度难以保证。基于评分搜索的贝叶斯网结构学习方法是用科学的评分准则度量得到的贝叶斯网结构, 故精度有保证, 但是效率比较低。因此, 可以考虑将这两种方法有机结合, 实现优势互补。K2 算法是比较经典的基于评分搜索的贝叶斯网结构学习算法, 利用它通常能学到高质量的贝叶斯网结构。然而, 它对输入节点的拓扑序列有着比较强的敏感性。因此研究如何得到高质量的拓扑序列对于利用 K2 算法学习贝叶斯网结构是十分必要的。鉴于此, 本文运用量子信息的丰富性和量子计算的并行性优势, 以量子遗传算法作为搜索算法, 通过基于量子染色体的拓扑序列生成策略得到高质量拓扑序列; 并采用带上下界的自适应量子变异策略, 增强了种群的多样性。这样既提高了算法的搜索性能, 又为 K2 算法提供了较高质量的序列, 弥补了它对拓扑序列敏感的缺陷, 进而保证了求解的质量; 最后, 在使用 K2 算法时先用计算量较小而又可靠的 0 阶条件独立性测

试压缩搜索空间,提高了 K2 算法的效率。

2.2 拓扑序列及其生成策略

2.2.1 拓扑序列

拓扑排序是由某个集合上的一个偏序得到该集合上的一个全序。拓扑排序的结果是一个线性序列,被称为拓扑序列。一个拓扑序列决定了节点间的次序,是对节点间部分关系的反映。例如:有四门课 C1、C2、C3、C4,它们的先修与后修关系如图 1 所示。图 1 是一张有向无环图,它的拓扑序列为 C1C2C3C4 或 C1C3C2C4。这两个拓扑序列就蕴含了一定的选修的先后顺序:C1 是在其他课程之前最先被修的课,而 C4 是在其余课程之后最后才能被修的课;C2 和 C3 是在中间才被修的,它们两个之间并没有必然的谁先谁后的关系。它揭示了 C1、C2、C3、C4 这四门课在选修的先后顺序上的部分关系。

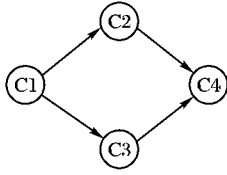


图 1 课程关系意图

2.2.2 基于量子染色体的拓扑序列生成策略

量子遗传算法就是把量子计算融入到遗传算法中形成的算法,具体来说就是用量子系统表示量子染色体,用量子计算实现遗传算法中的遗传操作。用量子遗传算法求解某一实际问题,其中很重要的一点是代表问题解的量子染色体的设计。本文提出的新算法是用量子遗传算法搜索拓扑序列,所以我们要设计出能够得到拓扑序列的量子染色体。对于由 n 个节点组成的拓扑序列而言,量子染色体的具体形式如下所示:

$$qc = \begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \\ \beta_1 & \beta_2 & \dots & \beta_n \end{bmatrix} \quad (9)$$

其中: α_i, β_i 表示第 i 位量子比特位的概率幅。这样的量子染色体形式简单,易于存储。有了量子染色体后,如何从量子染色体得到拓扑序列也是新算法中比较重要的一环。本文设计了基于量子染色体的拓扑序列生成策略,其具体表述如下:

对于某一个量子染色体而言,对其观测可以得到两个信息:一个是由 0 或 1 组成的数组 $bArray = \{a_1, a_2, \dots, a_n\}$, 一个是由 $|\alpha_i|^2$ 或 $|\beta_i|^2$ 组成的实数数组 $fArray = \{f_1, f_2, \dots, f_n\}$ 。这两个数组中元素的获取遵循以下规则:

$$a_i = \begin{cases} 0, & |\alpha_i|^2 > p; \\ 1, & \text{其他} \end{cases}; i = 1, 2, \dots, n \quad (10)$$

$$f_i = \begin{cases} |\alpha_i|^2, & a_i = 0 \\ |\beta_i|^2, & \text{其他} \end{cases}; i = 1, 2, \dots, n \quad (11)$$

其中 p 表示随机数。

就实数数组 $fArray$ 而言,再增加一个同维的从 0 开始的连续自然数序列的整数数组 $tArray = \{0, 1, \dots, n-1\}$, 然后对这两个数组同步调排序,这样排序后的 $tArray$ 就作为一个拓扑序列,可以作为 K2 算法的输入。数组 $bArray$ 的作用有两个:一是参与查旋转角调整策略表,获得旋转角的大小;二是根据它来产生实数数组。例如,当 $n = 5$ 时,某一个量子染色体可以表示为:

$$qc = \begin{bmatrix} \cos \frac{\pi}{6} & \cos \frac{\pi}{12} & \cos \frac{3\pi}{4} & \cos \frac{4\pi}{3} & \cos \frac{7\pi}{4} \\ \sin \frac{\pi}{6} & \sin \frac{\pi}{12} & \sin \frac{3\pi}{4} & \sin \frac{4\pi}{3} & \sin \frac{7\pi}{4} \end{bmatrix}$$

假设经过随机观测得到的 $bArray = \{1, 0, 0, 1, 0\}$, 则可以得到:

$$fArray = \left\{ \left(\sin \frac{\pi}{6} \right)^2, \left(\cos \frac{\pi}{12} \right)^2, \left(\cos \frac{3\pi}{4} \right)^2, \left(\sin \frac{4\pi}{3} \right)^2, \left(\cos \frac{7\pi}{4} \right)^2 \right\}$$

此时的自然数组 $tArray = \{0, 1, 2, 3, 4\}$, 然后将 $fArray$ 和 $tArray$ 同步排序,即可得:

$$fArray = \left\{ \left(\sin \frac{\pi}{6} \right)^2, \left(\cos \frac{3\pi}{4} \right)^2, \cos \left(\frac{7\pi}{4} \right)^2, \left(\sin \frac{4\pi}{3} \right)^2, \left(\cos \frac{\pi}{12} \right)^2 \right\}$$

$$tArray = \{0, 2, 4, 3, 1\}$$

从这里可以看到,这样的量子染色体实际上蕴含着许多个拓扑序列,当我们随机性地观测以后,仅仅得到了其中的一个,因此,量子染色体的信息是丰富的,经过较少代数的观测容易得到较优的拓扑序列。经过排序得到的 $tArray$ 就是根据上述生成策略得到的拓扑序列,它可以作为 K2 算法的输入参数。我们设计的基于量子染色体的拓扑序列生成策略是一种新的策略,这种策略有机地利用了量子概率幅和量子观测的知识,经过排序操作得到相应的拓扑序列。拓扑序列的得到是由两个因素决定的:一个是随机数,一个是实数数组中元素的相对大小。在产生随机数方法相同的情况下,量子旋转门作用以后的量子染色体中各个量子位向较优的量子染色体相对应的量子位靠近,这样就可以产生排列上相差不大的拓扑序列。因此,这种策略既使算法搜索具有一定的随机性,又确保了算法向着较优的方向演化。它充分利用了概率值,把概率值的相对大小用到拓扑序列的构造上来,并有机地结合了量子旋转策略的意图,即把量子位上的靠拢映射为排列顺序上的靠拢。因此,这种策略既可以随机搜索,又确保了算法的收敛性。

与文献[5]中得到拓扑序列的策略相比,本文中的这种策略首先利用了量子信息的丰富性和量子计算的并行性的优势,从而增强了拓扑序列的多样性,也提高了算法的搜索速度;其次,这种策略巧妙地实现了量子染色体到拓扑序列的映射,具有简单、易实现的特点,同时,量子信息本质上是一种概率性的表示方法,这也增加了搜索的宽度。因此,这种策略可以在较少的代数内得到相对较高质量的拓扑序列,为利用 K2 算法学得较高分数的贝叶斯网结构奠定了基础。

2.3 量子旋转门旋转角的更新机制

量子染色体是用量子信息表示的,因此不能采用传统遗传算法中实现遗传算子的方式,而是应当利用量子计算的知识来实现,即采用量子门作用于量子染色体。具体说来就是用量子门作用于量子的基态,使其相互干扰,引起相位发生改变,这也是最主要的遗传操作。因此,量子门的优劣对量子遗传算法的性能将产生直接的影响。本文采用量子旋转门作为主要的操作,对量子染色体中量子位的操作如式(8)所示。其中的量子旋转角按式(12)更新。

$$\Delta\theta_i = S(\alpha_i, \beta_i) \times \theta_i \quad (12)$$

其中: $S(\alpha_i, \beta_i)$ 表示旋转角的旋转方向, θ_i 表示旋转角的大小, $\Delta\theta$ 的值可以根据表 1 和式(12)确定。

量子染色体是量子遗传算法实现搜索进而得到问题解的载体,而量子旋转门是实现量子染色体进化的最主要的方法之一,为搜索提供了较强的动力。因此,本文设计了如表 1 所示的量子旋转门旋转角的更新机制。该机制的基本思想如

下: 将第 t 代的第 i 个个体测量值的评分 $f(x)$ 与该个体的目标评分 $f(b)$ 进行比较, 如果测量值的评分小于目标评分, 则调整个体 x 中相应的概率幅, 使得概率幅朝着有利于 b 出现的方向进化; 反之, 则调整个体 x 中相应的量子比特位的概率幅, 使个体向着更有利于 x 出现的方向进化, 同时把 x 作为下一代的目标 b 。这种概率幅的调整如图 2 所示, 其中, $(\alpha_{ij}, \beta_{ij})$ 表示第 i 个量子染色体的第 j 个量子位, $\Delta\theta_i$ 表示计算出来的量子旋转角, θ_{ij} 是第 i 个量子染色体的第 j 个量子位中的角, 即 $\cos \theta_{ij} = \alpha_{ij}$, $\sin \theta_{ij} = \beta_{ij}$ 。某一代的第 i 个量子当染色体而言, 假设第 i 个量子染色体的评分小于该量子染色体的目标评分, 就应该使第 i 个量子染色体向目标位置靠拢。使用式 (8) 即可实现图 2 表示的调整效果。因此, 这种机制会使迭代的结果朝着越来越好的方向演化, 从而保证了算法的收敛性。

表 1 量子旋转角旋转机制

x_i	b_i	$f(x) < f(b)$	θ_i	$S(\alpha_i, \beta_i)$			
				$\alpha_i \beta_i > 0$	$\alpha_i \beta_i < 0$	$\alpha_i = 0$	$\beta_i = 0$
0	0	FALSE	0	0	0	0	0
0	0	TRUE	0	0	0	0	0
0	1	FALSE	0.01π	-1	1	± 1	0
0	1	TRUE	0.01π	1	-1	0	± 1
1	0	FALSE	0.01π	1	-1	0	± 1
1	0	TRUE	0.01π	-1	1	± 1	0
1	1	FALSE	0	0	0	0	0
1	1	TRUE	0	0	0	0	0

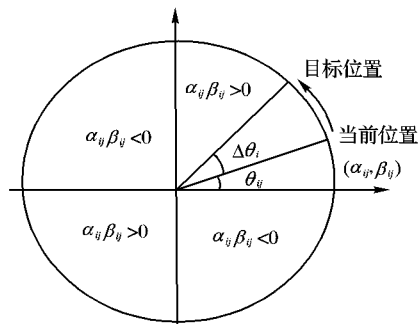


图 2 旋转示意图

2.4 带上下界的自适应量子变异策略

根据上述生成策略, 可以从一个量子染色体中得到一个拓扑序列, 再利用 K2 算法得到一个贝叶斯网结构。不同的量子染色体得到的贝叶斯网结构是不一样的, 那么它的评分也是不一样的。我们把得到的贝叶斯网结构的评分作为相应量子染色体优劣的一个度量, 那么评分越高, 说明越接近客观存在的贝叶斯网结构, 需要做的变异越少; 评分越低, 说明与客观存在的贝叶斯网结构差别较大, 应该做较多的变异。因此, 对所有的量子染色体做统一的变异是不合适的, 应该根据其优劣做不同程度的自适应变异。这样做既符合事物的演化发展规律, 也更有利于在较少的代数内找到较优的解。基于这样一种考虑, 我们根据评分设计出了带上下界的自适应量子变异策略。该策略表示如下:

$$cn_i = \max \{ \min \{ \sqrt{v_g - v_i}, 10 \}, 1 \} \quad (13)$$

其中: v_g 和 v_i 分别表示目标量子染色体评分和当前量子染色体的评分, $\max \{ \}$ 函数和 $\min \{ \}$ 分别表示取两者中较大和较小的一个, cn_i 表示某一代第 i 个量子染色体需要变异多少个量子位。这种带上下界的自适应变异策略, 既考虑了变异的有效性, 又控制了变异的强度, 防止变异过大带来太大的破坏。这种策略用在量子遗传算法中, 根据变异概率给种群做

自适应变异操作, 可以增强种群的多样性, 较为合理地提供新的量子染色体, 从而优化算法的搜索性能。需要说明的是式 (13) 中的“10”是要根据要学习的贝叶斯网的规模并通过实验来确定的。

2.5 算法描述及其分析

T&QGA-B 算法的核心思想是考虑到量子信息相对于经典信息的天然优势和 K2 算法的对节点序列的敏感性和贪婪性, 先用量子遗传算法搜得较优的拓扑序列, 然后利用 K2 算法得到贝叶斯网结构, 最后又对解做了局部优化。本文的创新之处主要体现在: 1) 设计了一种基于量子染色体的拓扑序列的生成策略; 2) 提出了一种自适应的量子变异策略。T&QGA-B 算法的伪代码如下所示:

Algorithm: T&QGA-B

Begin

1) Initialization;

Initialize: $qPopSize, nodeNum, pm, maxGen, lagGen, qChroms$

2) Condition independent test phase;

For every pair of nodes (X_i, X_j) do;

Do order-0 CI tests;

3) Initialization fitness;

For every $qChrom$ do;

Observe(idx);

K2Research(idx);

$G^+ = \arg \max f(G, D)$;

4) Search Phase;

For $i = 1$ to $maxGen$ do;

For $j = i$ to $qPopSize$ do;

Observe(idx);

$G = K2Research(idx)$;

$G = Optimization(G)$;

If $(f(G, D) > f(G^+, D))$ then $G^+ = G$;

If $(lagGen > k)$ then

For $i = 0$ to $qPopSize$

If $(rndm > pm)$ then mutation(i);

5) Return G^+ ;

End

其中: observe() 是用来观测、产生拓扑序列的函数, K2Research() 是用 K2 算法输出贝叶斯网的结构, Optimization(G) 是用来完成对解 G 的局部优化。

T&QGA-B 算法是一种基于量子信息和量子计算的随机搜索算法, 它的主要开销在两个方面: 一个是量子遗传算法搜索拓扑序列的开销; 另一个是利用 K2 学习贝叶斯网结构的开销。量子信息具有丰富性的特点, 量子计算具有并行性的优势。这既可以保持由量子染色体得到的拓扑序列的多样性, 又能减少迭代次数, 加快算法的收敛速度, 进而减少了调用 K2 的次数。带上下界的自适应量子变异策略增强了种群的多样性, 提高了算法的搜索能力。在用 K2 算法学习贝叶斯网的结构时, 我们用比较可靠的 0 阶条件独立性测试产生一个由一些禁忌边组成的禁忌表, 这样就压缩了搜索空间, 减少了搜索节点的父节点集的时间。提供给 K2 算法的节点序列有着较高的质量, 这就在一定程度上保证了求解的质量, 降低了 K2 对节点序列的敏感性所带来的不利影响。最后又进行了优化, 减轻了 K2 的贪婪性所造成的不良影响, 可以在一定程度上保证最后得到较优的解。

3 实验结果及其分析

为了测试 T&QGA-B 算法的性能, 使用公认的 Benchmark

数据集 ALARM 网络(数据集来源于 <http://www.cs.huji.ac.il/labs/compbil/Repository/>)做实验。实验的编程和运行环境:操作系统是 Windows XP2, CPU 是 Pentium4, 内存大小为 512 MB, 本文用 Java 编程语言实现 T&QGA-B 算法。实验所采用的参数: $maxGen = 100$, $popSize = 20$, $pm = 0.1$, $lagGen = 10$ 。

3.1 三种算法实验效果的比较

为了测试 T&QGA-B 算法的整体性能,在 ALARM 网络的不同容量上对 I-ACO-B 算法、C-PSO-B 算法和 T&QGA-B 算法做实验。

表 2 给出了对这些算法独立运行 10 次的实验结果,其中, K2 为算法运行得到的贝叶斯结构的评分值,括号内的值为运行 10 代中得到的最佳评分值; Ad、De 和 In 分别表示与标准网络结构相比较时额外增加的弧数、意外去除的弧数和错误指向的弧数; It 为算法达到最优解时所需要最小迭代次数。在这里我们采用统一的形式 $\mu \pm \sigma$ 表示上述三种算法独立运行 10 次所得到的这些指标的平均值 μ 和标准偏差 σ 。

从表 2 中可以看到,本文提出的新算法在收敛速度上和

求解质量上都要优于 I-ACO-B 算法和 C-PSO-B 算法。究其原因,量子遗传算法使用了量子计算的知识,运用了量子信息的丰富性和量子计算的并行性优势,增强了种群的多样性,提高了算法的收敛速度。我们设计的基于量子染色体的拓扑序列生成策略是一种新的策略,这种策略具有一定的随机性,观测时概率的相对大小决定了拓扑序列的生成。对同一个量子染色体多次观测得到的 $fArray$ 是不一样的,因此观测具有局部搜索的能力,这就加大了搜索的密度,增加了搜索的广度,有利于高质量的拓扑序列出现,从而在一定程度上为 K2 学得精度较高的贝叶斯网结构提供了保障。

带上下界的自适应量子变异策略,既体现了不同量子染色体变异强度的不同,又在一定程度上控制了变异强度,防止变异过大带来的不利影响。在量子旋转门的作用下,量子染色体又会不断地向着较优的方向演化。

从表 2 还可以看到,新算法 T&QGA-B 收敛的速度比较快,但是收敛代数的波动性比较大,这是因为对量子染色体的观测是一种比较随机的方法,这种随机性带来了搜索上的优势,同时也带来了收敛代数波动性大的现象。

表 2 三种算法在 ALARM 数据集上的实验结果

Sample Capacity	参数	I-ACO-B ^[1]	C-PSO-B	T&QGA-B
2000	K2	-9717.47 \pm 0.00(-9717.46)	-9716.14 \pm 0.15(-9716.13)	-9715.51 \pm 0.12(-9715.13)
	Ad	3.0 \pm 0.0(3)	2.93 \pm 0.2(3)	3.0 \pm 0.10(3)
	De	1 \pm 0.00(1)	1.0 \pm 0.0(1)	1.0 \pm 0.00(1)
	In	1.1 \pm 0.1(1)	1.1 \pm 0.0(1)	1.1 \pm 0.0(1)
	It	50.00 \pm 7.45(20)	42.25 \pm 7.59(31)	37.64 \pm 9.54(31)
3000	K2	-14401.66 \pm 0.10(-14401.29)	-14389.45 \pm 0.10(-14389.27)	-14384.52 \pm 0.14(-14384.37)
	Ad	1.4 \pm 0.16(2)	1.3 \pm 0.3(2)	1.16 \pm 0.27(2)
	De	1.0 \pm 0.00(1)	1.0 \pm 0.0(1)	1.0 \pm 0.00(1)
	In	1.4 \pm 0.16(2)	1.1 \pm 0.12(1)	1.21 \pm 0.23(1)
	It	50.00 \pm 8.03(20)	34.90 \pm 6.26(21)	27.40 \pm 7.25(19)
4000	K2	-19099.64 \pm 0.65(-19098.41)	-19087.25 \pm 0.32(-19086.37)	-19086.59 \pm 0.54(-19085.93)
	Ad	1.9 \pm 0.18(2)	1.7 \pm 0.22(2)	1.65 \pm 0.37(2)
	De	1.0 \pm 0.00(1)	1.0 \pm 0.00(1)	1.0 \pm 0.00(1)
	In	1.8 \pm 0.33(2)	1.3 \pm 0.11(1)	1.29 \pm 0.27(1)
	It	54.00 \pm 6.70(40)	45.57 \pm 7.04(30)	30.60 \pm 8.46(23)

3.2 三种算法收敛性能的比较

为了验证和比较本文算法的收敛性能,本文对 I-ACO-B、C-PSO-B 算法和 T&QGA-B 算法在 10 个样本容量上分别独立运行 10 次获得各自最优评分值时所需要的平均迭代次数的比较,如图 3 所示。

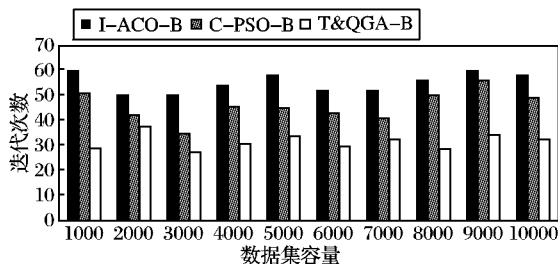


图 3 三种算法收敛代数的比较

从图 3 可以看出,算法收敛的迭代次数和数据量的大小没有明显的特定关系,这体现出量子遗传算法的随机性的一面,尤其是量子观测的随机性。但是本文算法的平均迭代次数在上述 10 个样本容量上都比 I-ACO-B 算法和 C-PSO-B 算法好很多。新算法运用了量子信息的丰富性和量子计算的并行性优势,量子观测的随机性使得每个量子染色体每次被观

测后得到的结果并不一样,这意味着随机性的观测具有局部搜索的能力,会使得根据生成策略得到的拓扑序列的多样性增强,这无疑会提高算法的搜索速度。带上下界的自适应量子变异策略根据量子染色体的质量做自适应性变异,这有利于增强量子染色体的多样性,有利于量子染色体向更好的方向去搜索,也有助于提高算法的收敛速度。

4 结语

本文以量子遗传算法为基础,提出一种基于拓扑序列和量子遗传算法的贝叶斯网结构学习算法。本文算法首先利用量子计算的优势,设计出简单易实现的基于量子染色体的拓扑序列的生成策略,以此得到高质量的拓扑序列,为使用 K2 算法学得较优的贝叶斯网结构提供了保障;然后使用带上下界的自适应量子变异策略,增强了种群的多样性和算法的搜索能力。在通用的 ALARM 数据集上的实验表明新算法是有效的。今后研究工作是对本文算法中旋转角的动态变化和迭代次数波动较大的问题做进一步研究,进而丰富量子群集智能算法。

(下转第 1603 页)

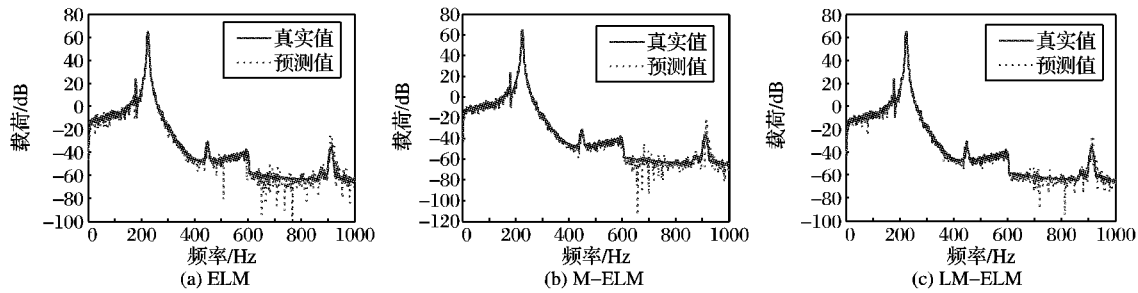


图2 激励点1三种算法对比预测效果

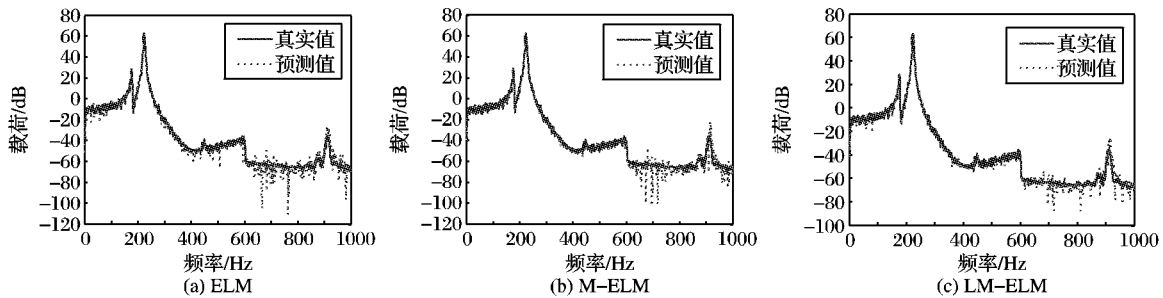


图3 激励点4三种算法对比预测效果

4 结语

ELM中冗余隐神经元的存在,不可避免地会减弱ELM模型泛化能力。进行模型选择,是扩展ELM应用范围、提高实用效果的关键步骤。本文在经典ELM隐神经元层之上,又引入了隐特征空间的概念,通过添加隐层空间到低维隐特征空间的映射函数矩阵的Frobenius范数作为正则项,构建了新的优化目标,最终采用交替优化的方法,自适应识别最优映射函数和ELM模型参数,并最终在UCI回归数据和圆柱壳载荷识别数据集验证了该算法的有效性。

参考文献:

- [1] HUANG G B, ZHU X, SIEW C. Extreme learning machine: theory and applications[J]. *Neurocomputing*, 2006, 70(1/2/3): 489 - 501.
- [2] HUANG G B, DING X J, ZHOU H. Optimization method based extreme learning machine for classification[J]. *Neurocomputing*, 2010, 74(1/2/3): 155 - 163.
- [3] 高光勇, 蒋国平. 采用优化极限学习机的多变量混沌时间序列预测[J]. *物理学报*, 2012, 61(4): 040506.
- [4] 张弦, 王宏力. 具有选择与遗忘机制的极端学习机在时间序列预测中的应用[J]. *物理学报*, 2011, 60(8): 080504.
- [5] HUANG G B, WANG D H, LAN Y. Extreme learning machines: a

survey[J]. *International Journal of Machine Learning and Cybernetics*, 2011, 2(2): 107 - 122.

- [6] FENG G R, HUANG G B, LIN Q P, *et al.* Error minimized extreme learning machine with growth of hidden nodes and incremental learning[J]. *IEEE Transactions on Neural Network*, 2009, 20(8): 1352 - 1357.
- [7] MAO W T, TIAN M, CAO X Z, *et al.* Model selection of extreme learning machine based on multi-objective optimization[J]. *Neural Computing & Applications*, 2013, 22(3): 521 - 529.
- [8] 刘学艺, 李平, 郜传厚. 极限学习机的快速留一交叉验证算法[J]. *上海交通大学学报*, 2011, 45(8): 6 - 11.
- [9] ZHU Q Y, QIN A K, SUGANTHAN P N, *et al.* Evolutionary extreme learning machine[J]. *Pattern Recognition*, 2005, 38(2): 1759 - 1763.
- [10] ZHENG V W, PAN S J, YANG Q, *et al.* Transferring multi-device localization models using latent multi-task learning[C]// *Proceedings of the 23rd AAAI Conference on Artificial Intelligence*. Palo Alto: AAAI Press, 2008: 1427 - 1432.
- [11] BEZDEK J C, HATHAWAY R J. Convergence of alternation optimization[J]. *Neural Parallel Scientific Computing*, 2003, 11(4): 351 - 368.
- [12] 毛文涛. 支持向量回归机模型选择研究及在综合力学环境预示中的应用[D]. 西安: 西安交通大学, 2011.

(上接第1599页)

参考文献:

- [1] 冀俊忠, 张鸿勋, 胡仁兵, 等. 基于独立性测试和蚁群优化的贝叶斯网结构学习算法[J]. *自动化学报*, 2009, 35(3): 281 - 288.
- [2] FAULKNER E. K2GA: Heuristically guided evolution of Bayesian network structures from data[C]// *Proceedings of the 2007 IEEE Symposium on Computational Intelligence and Data Mining*. Washington, DC: IEEE Computer Society, 2007: 18 - 25.
- [3] LARRANAGE P, POZA M, UURRAMENDI Y, *et al.* Structure learning of Bayesian networks by genetic algorithms: A performance analysis of control parameters[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1996, 18(9): 912 - 925.
- [4] 刘大有, 王飞, 卢奕南, 等. 基于遗传算法的Bayesian网络结构学习研究[J]. *计算机研究与发展*, 2001, 38(8): 916 - 922.
- [5] 赵学武, 冀俊忠, 程亮, 等. 基于链模型和粒子群的贝叶斯网结构学习算法[J]. *计算机工程*, 2011, 37(37): 181 - 184.
- [6] HAN K H, KIM J H. Quantum-inspired evolutionary algorithm for a class of combinatorial optimization[J]. *IEEE Transactions on Evolu-*

tionary Computation, 2002, 6(6): 580 - 593.

- [7] PLATELT M D, SCHLIEBS S, KASABOV N. A versatile quantum-inspired evolutionary algorithm[C]// *IEEE Congress on Evolutionary Computation*. Piscataway: IEEE, 2007: 423 - 430.
- [8] TALBI H, DRAA A, BATOCHE M. A new quantum-inspired genetic algorithm for solving the travelling salesman problem[C]// *Proceedings of IEEE International Conference on Industrial Technology*. Piscataway: IEEE, 2004: 1192 - 1197.
- [9] 邢焕来, 潘炜, 邹喜华. 一种解决组合优化问题的改进型量子遗传算法[J]. *电子学报*, 2007, 35(10): 1999 - 2002.
- [10] 熊焰, 陈欢欢, 苗付友, 等. 一种解决组合优化问题的量子遗传算法 QGA[J]. *电子学报*, 2004, 32(11): 1855 - 1858.
- [11] KE M, HONG G, ZHAO Y D, *et al.* Quantum-inspired particle swarm optimization for valve-point economic load dispatch[J]. *IEEE Transactions on Power Systems*, 2010, 25(1): 215 - 222.
- [12] 冀俊忠, 胡仁兵, 张鸿勋, 等. 一种混合的贝叶斯网结构学习算法[J]. *计算机研究与发展*, 2009, 46(9): 1498 - 1507.