

# 基于约束的局部-全局LWF链图结构学习算法

曹付元<sup>1</sup>, 杨淑晶<sup>1</sup>, 王雲霞<sup>1</sup>, 俞奎<sup>2</sup>

(1. 山西大学计算机与信息技术学院, 山西太原 030006; 2. 合肥工业大学计算机与信息学院, 安徽合肥 230601)

**摘要:** LWF链图结构学习旨在发现链图中所有节点的父节点、子节点、邻居节点以及配偶节点。然而, 目前最新的LWF链图结构学习算法是基于Growing-Shrinking(GS)思想得到节点的局部结构(即节点的马尔科夫毯)来学习全局网络结构, 该类算法的条件独立测试是以整个马尔科夫毯为条件集的, 为了保证条件独立测试的可靠性, 算法要求样本数量是马尔科夫毯大小的指数级, 从而使得算法的数据效率较差。针对该问题, 本文提出了一种基于约束的局部-全局LWF链图结构学习算法。该算法通过迭代的学习邻接集和配偶集来降低对数据样本量的要求; 与此同时, 在学习邻接集时采用后向策略保障了条件独立测试的正确性。算法的基本思想如下: 首先学习网络中每个节点的马尔科夫毯, 将节点马尔科夫毯学习拆分为学习邻接集和学习配偶集; 然后利用节点的马尔科夫毯信息恢复网络骨架, 根据链图复合体有向边的特点, 利用条件独立测试确定网络复合体有向边, 从而恢复链图结构。理论分析证明了该算法的正确性, 在仿真数据集和标准数据集上的实验测试验证了算法的有效性。

**关键词:** LWF链图; 马尔科夫毯; 条件独立测试; 数据效率

**基金项目:** 国家自然科学基金(No.61976128)

**中图分类号:** TP18

**文献标识码:** A

**文章编号:** 0372-2112(2023)06-1458-10

**电子学报URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20210134

## Local-Global LWF Chain Graph Structure Learning Algorithm Based on Constraints

CAO Fu-yuan<sup>1</sup>, YANG Shu-jing<sup>1</sup>, WANG Yun-xia<sup>1</sup>, YU Kui<sup>2</sup>

(1. School of Computer and Information Technology, Shanxi University, Taiyuan, Shanxi 030006, China;

2. School of Computer Science and Information Engineering, Hefei University of Technology, Hefei, Anhui 230601, China)

**Abstract:** LWF chain graph structure learning aims to find the parents, children, neighbours and spouses of all nodes in the chain graph. Currently, the state-of-the-art LWF chain graph structure learning algorithms obtain the local structure of nodes to learn the global network structure based on Growing-Shrinking (GS) idea. The conditional independence test of these algorithms takes the whole Markov blanket (MB) as the condition set. In order to ensure the reliability of conditional independence test, the number of samples is required to be exponential level of the size of Markov blanket, which makes the data efficiency of the algorithm poor. To alleviate this problem, we propose a Local-Global LWF chain graph structure learning algorithm based on constraints, which reduces the requirement of sample size of data by iterative learning adjacencies and spouses; while learning adjacencies, it further improves the accuracy of the conditional independence test by backward strategy. The basic idea of the algorithm as follows: firstly, the Markov blanket of each node in the network is learned, and the Markov blanket learning of node is divided into learning the adjacencies and the spouses; secondly, we use the Markov blanket information of nodes to recover the network skeleton and take advantage of conditional independent test to discover its complexes, which restores the chain graph structure, according to the characteristics of directed edges of chain graph complexes. Theoretical analysis demonstrates the correctness of the algorithm. Moreover, experiments on the generated datasets and standard datasets show the effectiveness of the algorithm.

**Key words:** LWF chain graph; Markov blanket; conditional independence test; data efficiency

**Foundation Item(s):** National Natural Science Foundation of China (No.61976128)

## 1 引言

概率图模型<sup>[1,2]</sup>是表示概率条件独立信息的一种方便而直观的形式,其中经典的两种模型是贝叶斯网络(有向图)和马尔科夫网络(无向图).贝叶斯网络中的边都是有向边,马尔科夫网络中的边都是无向边.但是在现实生活的许多领域中,存在一些特定的条件独立结构,无法仅用有向图或者无向图来构建.例如图1所示血脂紊乱与糖尿病这两种疾病之间的关联<sup>[3]</sup>,因为它们在生理学有某种程度的相关,需要利用包含无向边和有向边的混合图来模拟它们之间的关系.因此,链图的引入是非常必要的.

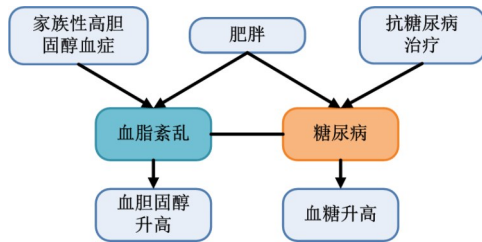


图1 血脂紊乱与糖尿病的关联图

链图是贝叶斯网络和马尔科夫网络两种经典模型的自然推广,既包含有向边又包含无向边.依赖于无向边的解释,链图主要有三种不同的解释:第一种是1989年由Lauritzen等人<sup>[4]</sup>和Frydenberg<sup>[5]</sup>提出的LWF解释;第二种是1996年由Andersson等人<sup>[6]</sup>提出的AMP解释;第三种是1993年由Cox等人<sup>[7,8]</sup>提出的MVR(MultiVariate Regression)多元回归解释.Sonntag等人<sup>[9]</sup>详细叙述了三种解释之间的关系.本文主要研究LWF解释下的链图(LWF链图).LWF链图是有向无环图与无向图的结合,从因果关系的角度来看,LWF链图中的有向边表示直接的因果效应,而LWF链图中的无向边表示由于干扰而产生的因果效应<sup>[10]</sup>.

LWF链图作为统计应用中的建模工具已经得到了越来越多的关注,并被用于人体活动识别和图像分离<sup>[11]</sup>.但是,LWF链图在现实世界中的应用仍然非常有限,一个重要的原因是缺乏现实可靠的LWF链图结构学习算法.

目前,现有的LWF链图结构学习算法大多都是对贝叶斯结构学习算法的推广.Ma等人<sup>[12]</sup>提出的基于分解技术的LWF链图结构学习(Learn Chain graph via Decomposition, LCD)算法,采用了分而治之的策略提高了搜索效率,但是建立分离树的过程非常耗时耗力.王静云等人<sup>[13]</sup>提出了基于GS(Growing-Shrinking)思想的GSCG(GS Chain Graph)算法,该算法通过学习节点的局部结构恢复整个网络结构,但是GSCG算法存在条件独立测试不可靠、数据效率低和学习节点局部边界时

引入错误节点等不足.Javidian等人<sup>[14]</sup>提出了MbLWF算法(基于GS思想),但是该算法要求样本数量是节点马尔科夫毯大小的指数级,来保证条件独立测试的可靠性,从而导致算法的数据效率较低.

为了提高条件独立测试的可靠性、提高算法的数据效率及减少在学习节点马尔科夫毯时添加过多的错误节点,在概率分布满足忠实性假设条件下,本文提出了一种基于约束的局部-全局LWF链图结构学习(Local-Global LWF Chain Graph structure learning algorithm based on constraints, LGCG)算法,该算法将网络结构恢复问题转化为局部学习马尔科夫毯问题,利用节点的马尔科夫毯信息恢复整个网络结构.本文主要贡献如下:

(1)将马尔科夫毯的学习拆分为学习邻接集和配偶集,不仅可以减少在学习马尔科夫毯时因错误节点的加入导致的级联错误,还可以避免以马尔科夫毯为条件集出现在条件独立测试中,从而提高算法条件独立测试的可靠性和算法的数据效率.

(2)在学习邻接集时采用后向策略,使得条件独立测试给定的条件集总是能够基于最小的分离集,从而保障了条件独立测试的正确性.

(3)恢复链图结构时,将骨架恢复与有向边的恢复同时进行,提高了算法的效率.

## 2 相关工作

LWF链图结构学习方法大致分为两类<sup>[15]</sup>:基于约束的方法和基于随机搜索机制的评分搜索方法.由于目前很难找到有效的评分函数,链图结构学习主要是基于约束的方法.基于约束的LWF链图结构学习算法是对两个节点进行依赖性或独立性测试,以决定是否在两个节点之间添加边,如果两个节点之间存在边,则需要利用条件独立测试确定边的方向.下面具体介绍目前基于约束的主要方法.

Verma等人<sup>[16]</sup>提出了PC Like算法,该算法是一个因果模型的典型表示,并且该算法定义了一个判断等价类的有效图形标准.Studený<sup>[17]</sup>基于PC Like算法学习LWF链图,提出了LCG算法.Peña<sup>[18]</sup>和Sonntag等人<sup>[19]</sup>分别将PC Like算法用于学习另外两种解释的链图.Ma等人<sup>[12]</sup>提出的基于分解技术的结构学习算法(LCD算法),是对Xie等人<sup>[20]</sup>提出的学习贝叶斯网络分解算法的推广,该算法采用了分而治之的策略提高了搜索效率.Peña等人<sup>[21]</sup>在扩展Meek规则的基础上提出了学习包含最优链图结构的CKES算法,该算法基于学习贝叶斯网络结构的KES<sup>[22]</sup>算法思想,放松了对概率分布忠实性的假设,由于链图等价类空间不可枚举,因此该算法适用于低维数据问题.王静云等人<sup>[13]</sup>提出了基于GS

思想的LWF链图等价类结构学习算法GSCG算法,该算法是通过学习节点的局部结构恢复整个网络结构. Peòà等人<sup>[23]</sup>提出了PCMB算法,该算法是在无须学习贝叶斯网络的条件下,从数据中学习马尔科夫边界的算法,Wang等人<sup>[24]</sup>受PCMB算法的启发提出了有错误发现率控制的局部链图结构学习算法.Javidian等人<sup>[14]</sup>扩展了马尔科夫毯的概念,提出了MbLWF算法(基于GS思想)来发现LWF链图中的马尔科夫毯.

上述研究工作中算法要求样本数量是节点马尔科夫毯大小的指数级,来保证条件独立测试的可靠性,从而导致算法的数据效率较低.为此,本文采用了拆分策略,该策略的优势是避免以马尔科夫毯为条件集出现在条件独立测试里,从而提高条件独立测试的可靠性,提升算法的数据效率;此外,本文结合了后向策略的优势,使得条件独立测试的结果更准确.

### 3 基本定义和性质

在这一节中,主要介绍相关的定义和性质.表1列出了本文中使用的符号.

表1 符号

符号	表示意义
$G, V, E, P$	链图,节点集,边集,联合概率分布
$X, Y, T$	一个节点
$W, R, S, Z$	$V$ 上的集合
$\text{Pa}(X)$	节点 $X$ 的父集
$\text{Ch}(X)$	节点 $X$ 的子集
$\text{Ne}(X)$	节点 $X$ 的邻居集
$\text{Adj}(X)$	节点 $X$ 的邻接集
$\text{SP}(X)$	节点 $X$ 的配偶集
$\text{Mb}(X)$	节点 $X$ 的马尔科夫毯
$\text{Adjs}$	用于存放 $V$ 中所有节点的邻接集
$\text{Sepset}_{X,Y}$	存放 $X$ 与 $Y$ 的分离集
$\text{Sepset}$	用于存放 $V$ 中两两节点间的分离集
$X \perp Y   S$	给定条件集 $S$ , $X$ 与 $Y$ 条件独立
$X \not\perp Y   S$	给定条件集 $S$ , $X$ 与 $Y$ 条件依赖
$ \cdot $	集合的大小

**定义 1**<sup>[17]</sup> 链图是一个二元组 $(G, P)$ ,其中 $G=(V, E)$ , $V$ 是有限节点集合, $E$ 是由 $V$ 中成对节点组成的集合称为边集, $P$ 是定义在有限集合 $V$ 上的联合概率分布.

链图中,对于任意互异节点 $X, Y \in V$ ,如果 $(X, Y) \in E$ 且 $(Y, X) \in E$ ,则称节点之间的边为无向边,表示为 $X-Y$ 或者 $Y-X$ ;如果 $(X, Y) \in E$ 且 $(Y, X) \notin E$ ,则称节点之间的边为有向边,表示为 $X \rightarrow Y$ .本文将 $V$ 中任意一个节点 $X$ 的父集、子集、邻居集和邻接集定义如下:

$$\text{Pa}(X) = \{Y \in V | Y \rightarrow X\},$$

$$\text{ch}(X) = \{Y \in V | X \rightarrow Y\},$$

$$\text{Ne}(X) = \{Y \in E | X-Y\},$$

$$\text{Adj}(X) = \text{Pa}(X) \cup \text{Ch}(X) \cup \text{Ne}(X)$$

图 $G$ 中有一条从 $X$ 到 $Y$ 长度为 $n(n \geq 2)$ 的路径,该路径是由互异节点序列 $\{X \equiv X_0, X_1, \dots, X_n \equiv Y\} \subseteq V$ 构成的,并且满足 $(X_{i-1}, X_i) \in E, \forall i=1, 2, \dots, n$ ,其中称路径 $X_0, X_1, \dots, X_n(n \geq 0)$ 为下降路径;如果对对所有 $i=1, 2, \dots, n$ ,有 $X_{i-1} \rightarrow X_i$ 或者 $X_{i-1} - X_i$ 在该路径中,称 $X_n$ 为 $X_0$ 的后代节点,称相应的 $X_0$ 为 $X_n$ 的祖先节点. $V$ 的一个子集 $Z$ 的祖先节点是 $Z$ 中所有节点的祖先节点集合,表示为 $\text{An}(Z)$ .

**定义 2**<sup>[17]</sup> 图 $G$ 中称互异节点 $X_1, X_2, \dots, X_n(n \geq 3)$ 产生的子图为复合体,若满足 $X_1 \rightarrow X_2 - X_3 - \dots - X_{n-2} - X_{n-1} \leftarrow X_n$ ,其中节点 $X_1$ 和 $X_n$ 称为复合体的父节点.

如果两个互异节点 $X$ 和 $Y$ 为相同复合体的父节点,则称 $X$ (或 $Y$ )是 $Y$ (或 $X$ )的配偶节点,记 $X$ 的所有配偶节点为集合 $\text{SP}(X)$ .

链图 $G$ 的骨架图<sup>[12]</sup> $G^*=(V, E^*)$ 是一个无向图,其中 $E^*$ 是去掉 $G$ 中所有的有向边的方向后得到的边集;由 $G$ 所决定的道德图<sup>[12]</sup> $G^m=(V, E^m)$ 为无向图,其中 $E^m$ 是 $E^*$ 与 $G$ 中连接每个复合体父节点的无向边的并集;链图 $G$ 的模式图<sup>[12]</sup> $G^{**}=(V, E^{**})$ ,其中 $E^{**}$ 是将 $G$ 中不在 $G$ 的任何复合体中的有向边变成无向边而得到的图.

如果 $W, R, S$ 为 $V$ 中的3个不相交的子集, $W$ 和 $R$ 为非空集合, $W$ 和 $R$ 在给定条件 $S$ 时关于概率分布 $P$ 是相互独立的( $W$ 和 $R$ 被 $S$ 集合 $c$ -分离<sup>[25]</sup>,称 $S$ 为分离集),表示为 $W \perp R | S [P]$ .

**定义 3**<sup>[25]</sup> 给定链图 $(G, P)$ ,概率分布 $P$ 忠实于图 $G=(V, E)$ ,当且仅当 $X \perp Y | S [G] \Leftrightarrow X \perp Y | S [P]$ .

定义3表明, $X$ 和 $Y$ 在给定条件 $S$ 时关于概率分布 $P$ 是相互独立的,等价于 $X$ 和 $Y$ 在给定条件 $S$ 时关于图 $G$ 是相互独立的.

**定义 4**<sup>[14]</sup> 在忠实性条件下,给定链图 $(G, P)$ ,节点 $X$ 的马尔科夫毯 $\text{Mb}(X)$ 是 $X$ 的父集、 $X$ 的子集、 $X$ 的邻居集和 $X$ 的配偶集组成的集合,即 $\text{Mb}(X) = \text{Pa}(X) \cup \text{Ch}(X) \cup \text{Ne}(X) \cup \text{SP}(X)$ .

**定义 5**<sup>[17]</sup> 如果两个链图包含相同的条件独立信息(相同的复合体和相同的骨架图,即有相同的模式图),则称它们是属于同一个马尔科夫等价类.

**性质 1**<sup>[17]</sup> 给定链图 $(G, P), G=(V, E)$ ,概率分布 $P$ 满足全局马尔科夫性,如果图 $(G_{\text{An}(W \cup R \cup S)})^m$ 中 $S$ 分离 $W$ 和 $R$ ,则 $W \perp R | S [P]$ . $(G_{\text{An}(W \cup R \cup S)})^m$ 是 $\text{An}(W \cup R \cup S)$ 在道德图 $G^m$ 中的导出子图.

图2为链图a,其中 $B-D$ 为链图中的一条无向边,

$D \rightarrow F$ 为链图中的一条有向边,从图2可知,节点 $F$ 的父集、子集、邻居集、邻交集、配偶集和马尔科夫毯分别为  $Pa(F)=\{D\}$ ,  $Ch(F)=\{K\}$ ,  $Ne(F)=\{E\}$ ,  $Adj(F)=\{D, E, K\}$ ,  $SP(F)=\{G\}$ ,  $Mb(F)=\{D, E, G, K\}$ ;由马尔科夫等价类的定义可知,链图a和图3所示的链图b属于同一马尔科夫等价类,因为它们具有相同的复合体和相同的骨架图,其中图4为它们的骨架图,图5为它们的道德图,并且两个链图的模式图均为链图b.

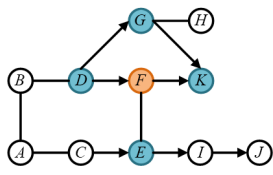


图2 链图a

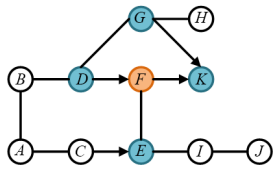


图3 链图b

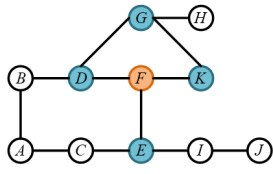


图4 骨架图

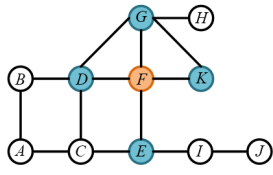


图5 道德图

### 4 学习LWF链图结构的LGCG算法

在这一节中,本文首先分析现存的基于GS思想的LWF链图结构学习算法的不足之处,然后详细介绍本文提出的LGCG算法,最后用一个实例对LGCG算法进行解释.

#### 4.1 基于GS思想的算法

目前最新的GSCG算法和MbLWF算法是基于GS思想的,它们均是利用马尔科夫毯发现算法学习节点邻域信息,虽然它们均提高了算法的搜索效率,但是GSCG算法和MbLWF算法存在一些不足之处:第一,算法的条件独立测试是以整个马尔科夫毯为条件集的,为了保证条件独立测试的可靠性,算法要求样本数量是马尔科夫毯的指数级,从而使得算法的数据效率较

差,所以当样本量较少时,条件独立测试的可靠性较低;第二,算法在学习马尔科夫毯信息的Growing阶段,由于配偶节点较晚进入,可能导致引入错误节点,而一旦有错误节点进入,将持续误导随后迭代里的判断,从而导致更多的错误节点进入,产生级联错误.

例如在图6中,以 $D$ 节点为例,从图中可以得知, $D$ 节点的马尔科夫毯  $Mb(D)=\{B, C, F, G\}$ . 在 Growing 阶段,如果将 $C$ 节点添加到 $D$ 节点的 $Mb(D)$ 中,集合 $Mb(D)$ 中至少包含 $F$ 节点和 $E$ 节点. 而 $E$ 节点是不属于 $Mb(D)$ 的,一旦节点 $E$ 提早进入 $Mb(D)$ 中,将会持续误导后面迭代里的判断,产生级联错误.

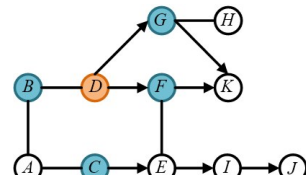


图6 产生级联错误的例子

为了解决以上算法的不足,本文提出了一种基于约束的局部-全局LWF链图结构学习算法,即LGCG算法,下面将详细介绍该算法.

#### 4.2 LGCG算法

本文提出的LGCG算法将链图结构学习分为两个阶段实现:第一阶段学习链图中节点的马尔科夫毯,第二阶段利用节点的马尔科夫毯信息恢复整个链图结构.

##### 4.2.1 学习节点的马尔科夫毯

本文将学习节点的马尔科夫毯拆分为两步:一是学习节点的邻交集,提出了RecognizeAdj算法(如算法1所示);二是学习节点的配偶集并保存到节点的马尔科夫毯中,提出了findMb算法(如算法2所示). 这个拆分学习策略可减少学习马尔科夫毯时因错误节点的加入而导致的级联错误,也可避免以马尔科夫毯为条件集出现在条件独立测试里,这样不仅有助于提高条件独立测试的可靠性,而且有助于提高算法的数据效率,继而让条件独立测试的结果所做的决策(比基于GS思想的算法)更准确.

RecognizeAdj算法用于学习所有节点邻交集. 该算法采用后向策略,即假设除目标节点以外的所有节点均是目标节点的候选邻交集,通过条件独立测试删除候选邻集中的错误节点,得到最终的邻交集. 后向策略使得判断目标节点与错误节点的条件独立性时,条件独立测试给定的条件集总是能够基于最小的分离集,保障了条件独立测试的正确性.

算法1中,步骤6~24学习目标节点 $T$ 的邻接节点,首先初始化  $Adj(T)=\mathcal{V}\setminus\{T\}$ ,  $cutSetSize=0$ , 如果给定条

**算法 1** RecognizeAdj 学习所有节点的邻接集输入:  $D$ :Data,  $\varepsilon$ :Threshold.

输出: Adj:Adjacency set of all nodes.

Sepset:Separation set of all nodes.

```

1. NonAdj =  $\emptyset$ ;
2. Adj =  $\emptyset$ ;
3. FOR ( $T \in V$ )DO
4.   adjt =  $V \setminus \{T\}$ ;
5.   cutSetSize = 0;
6.   WHILE ( $|\text{adjt}| > \text{cutSetSize}$ )DO
7.     FOR ( $X \in \text{adjt}$ )DO
8.       FOR ( $S \in \text{adjt} \setminus \{X\}, |S| = \text{cutSetSize}$ )DO
9.         IF ( $\text{pvalue}(T, X|S) \leq \varepsilon$ )THEN
10.          NonAdj = NonAdj  $\cup \{X\}$ ;
11.          Sepset $_{T,X} = S$ ;
12.        END IF
13.      END FOR
14.    END FOR
15.    IF ( $|\text{NonAdj}| > 0$ )THEN
16.      adjt = adjt  $\setminus$  NonAdj;
17.      cutSetSize = cutSetSize + 1;
18.      NonAdj =  $\emptyset$ ;
19.    ELSE
20.      cutSetSize = cutSetSize + 1;
21.    END IF
22.    Adj( $T$ ) = adjt;
23.  END WHILE
24. END FOR
25. RETURN Adj, Sepset;

```

件  $S, S \subseteq \text{Adj}(T) \setminus \{X\}, |S| = \text{cutSetSize}$ , 节点  $X$  与  $T$  独立, 则将节点  $X$  加入一个非邻接集合 NonAdj 中, 等待 adjt 中节点全部遍历完毕, 从 adjt 中删除 NonAdj 中的节点; 然后 cutSetSize 加 1, 重复以上步骤, 直到  $|\text{adjt}| \geq \text{cutSetSize}$  终止循环, 最终输出目标节点  $T$  的邻接集 adjt, 并将 adjt 存放在 Adj 中.

findMb 算法利用节点的邻接集用于学习节点的马尔科夫毯. 该算法首先对节点的邻接集进行修剪, 删除错误节点, 目的是避免后面迭代里产生的级联错误; 算法在学习目标节点的配偶集时, 其条件集不是马尔科夫毯, 而是上一阶段得到的分离集, 这样可以提高算法的数据效率, 保证算法的效率不依赖于样本的大小.

算法 2 中, 步骤 2~7 对节点的邻接集进行修剪, 如果两个节点邻接, 则两个节点分别存在于彼此的邻接集中; 如果  $X \in \text{Adj}(T)$ , 但是  $T \notin \text{Adj}(X)$ , 从  $\text{Adj}(T)$  删除  $X$ , 得到精度高的邻接集; 然后在步骤 9~18 通过遍历节点  $X \in \text{Adj}(T)$  和节点  $Y \in V \setminus \{\text{Adj}(T) \cup \{T\}\}$ , 如果满足

**算法 2** findMb 学习节点的马尔科夫毯输入:  $D$ :Data,  $T$ :Target, Adj, Sepset,  $\varepsilon$ :Threshold.输出: Mb( $T$ ): Mb of  $T$ .

```

1. cmb = Adj( $T$ );
2. FOR ( $X \in \text{Adj}(T)$ )DO
3.   IF ( $T \notin \text{Adj}(X)$ )THEN
4.     cmb = cmb  $\setminus \{X\}$ ;
5.     Sepset $_{T,X} = \text{Sepset}_{X,T}$ ;
6.   END IF
7. END FOR
8. Adj( $T$ ) = cmb;
9. FOR ( $X \in \text{Adj}(T)$ )DO
10.  candids =  $V \setminus \text{cmb} \setminus \{T\}$ ;
11.  FOR ( $Y \in \text{candids}$ )DO
12.    pval1 = pvalue( $T, Y | \text{Sepset}_{T,Y}$ );
13.    pval2 = pvalue( $T, Y | \text{Sepset}_{T,Y} \cup \{X\}$ );
14.    IF ( $\text{pval1} < \varepsilon$  &  $\text{pval2} > \varepsilon$ )THEN
15.      cmb = cmb  $\cup \{Y\}$ ;
16.    END IF
17.  END FOR
18. END FOR
19. continue = TRUE;
20. IF ( $|\text{cmb}| == 0$ )THEN
21.  continue = FALSE;
22. END IF
23. WHILE (continue)DO
24.   $p_X = \text{pvalue}(T, X | \text{cmb} \setminus \{X\}_{X \in \text{cmb}})$ ;
25.  p.val.max =  $\max_{X \in \text{cmb}} p_X$ ;
26.  candidates =  $\{X \in \text{cmb} | p_X = \text{p.val.max}\}$ ;
27.  IF ( $\text{p.val.max} < \varepsilon$ ) THEN
28.    cmb = cmb  $\setminus$  (candidates[1]);
29.  ELSE
30.    continue = FALSE;
31.  END IF
32. END WHILE
33. Mb( $T$ ) = cmb;
34. RETURN Mb( $T$ );

```

$T$  与  $Y$  在  $\text{Sepset}_{T,Y}$  下条件独立, 在  $\text{Sepset}_{T,Y}$  与  $X$  的并集下条件依赖, 则  $Y$  为  $X$  的一个配偶. 根据全局马尔科夫性质, 这两个条件共同保证了  $Y \in \text{SP}(T)$ ; 最后在步骤 19~33 修剪  $T$  的马尔科夫毯, 以得到精确的马尔科夫毯.

给出如下定理 1~3, 其中定理 3 给出了算法 1 和算法 2 的正确性证明.

**定理 1**<sup>[17]</sup> 如果概率分布  $P$  忠实于  $G$ , 则  $G$  中的每一对节点  $X$  和  $Y$  是邻接的, 当且仅当  $\forall Z \subseteq V, X \not\perp Y | Z[P]$ .

**定理 2**<sup>[14]</sup> 如果概率分布  $P$  忠实于  $G$ , 则  $G$  中的任意节点  $X$  满足:  $X \perp \bigvee \{X, \text{Mb}(X)\} | \text{Mb}(X)$ .

**定理 3** 链图  $(G, P)$ , 概率分布  $P$  忠实于图  $G=(V, E)$ , 假定独立性测试是正确的, 且数据集是独立抽样的样本集, 则算法 1 和算法 2 能够得到节点  $X \in V$  的马尔科夫毯.

**证明** (1) 在假定条件独立测试正确的前提下, 如果在图  $G$  中  $X$  是  $T$  的邻接节点, 由定理 1 可知不存在一个分离集, 使得节点  $X$  与  $T$  条件独立, 因此节点  $X$  不会被删除;  $X$  不是  $T$  的邻接节点, 由定理 1 可知存在一个分离集使得  $X$  与  $T$  条件独立, 因此节点  $X$  会被删除.

(2) 如果条件独立测试出现不可靠的情况, 算法 2 中步骤 2~7 对节点的邻接集进行修剪, 使得邻接集更加精确. 如果  $T$  和  $X$  是邻接的, 则一定有  $X \in \text{Adj}(T)$  且  $T \in \text{Adj}(X)$ , 则此阶段显然是正确的, 并且此阶段称为对称约束.

(3) 在假定条件独立测试正确的前提下, 如果  $X$  是  $T$  的配偶节点, 由定理 3 及  $\text{Mb}(T)$  的定义可知,  $T$  与  $Y$  在分离集下条件独立, 在分离集与  $X$  的并集下条件依赖, 因此节点  $Y$  被添加到  $\text{Mb}(T)$  中.

综上所述,  $\text{Mb}(T)$  中包含了节点的马尔科夫毯, 即邻居节点、父节点、子节点及配偶节点.

证毕

#### 4.2.2 链图的结构化恢复

利用得到的节点的马尔科夫毯信息, 恢复链图的骨架, 通过条件独立测试恢复复合体的有向边. 由于同一马尔科夫等价类中的任何两个链图结构都是不可区分的, 所以在结构化恢复阶段是恢复链图的模式图.

算法 3 根据链图中节点的邻接节点及配偶节点的性质, 从  $\text{Mb}(T)$  中分离出  $\text{Adj}(T)$  和  $\text{SP}(T)$ , 属于  $\text{Adj}(T)$  集合的节点, 直接用无向边连接两个节点, 属于  $\text{SP}(T)$  集合的节点, 通过条件独立测试判断两个节点之间边的方向, 最终得到链图结构. 与之前的算法相比, 本文算法的不同之处是每次往集合  $\text{Adj}(T)$  或  $\text{SP}(T)$  添加节点, 就触发去判断两个节点之间是有向边还是无向边, 这个策略是将骨架恢复与复合体有向边恢复同时进行, 进一步减少了算法的运行时间.

**定理 4**<sup>[13]</sup> 链图  $(G, P)$ , 概率分布  $P$  忠实于图  $G=(V, E)$ , 假定独立性测试是正确的, 且数据集是独立抽样的样本集, 则算法 3 能够恢复链图骨架和所有复合体.

#### 4.3 LGCG 算法举例

在这一节, 本文利用一个实例, 来介绍 LGCG 算法的执行.

如图 7 所示, 以链图 c 中的  $D$  节点为例, 首先初始化  $\text{Adj}(D)=\{A, B, C, E, F, G, H, I, J, K\}$  (图 8), 条件集大小

#### 算法 3 结构化恢复算法

输入:  $\text{Mb}$ : Mb of all nodes.

输出: global LWF Chain Graph structure.

1.  $\text{Adj}(s) = \emptyset;$
2.  $\text{SP} = \emptyset;$
3. FOR( $X \in \text{Mb}[T]$ )DO
4. IF( $\text{pvalue}(T, X|B) > \alpha, \exists B \subseteq \text{Mb}[T]$ )THEN
5.  $\text{Adj}(T) = \text{Adj}(T) \cup \{X\};$
6. add edge  $T-X;$
7. ELSE
8.  $\text{SP}(T) = \text{SP}(T) \cup \{X\};$
9. IF( $\text{pvalue}(T, X|B \cup \{Y\}) > \epsilon, \exists Y \in \text{Adj}(T), \forall B \subseteq \text{Adj}(T) \setminus \{Y\}$ ) THEN
10. orient  $T \rightarrow Y;$
11. END IF
12. END IF

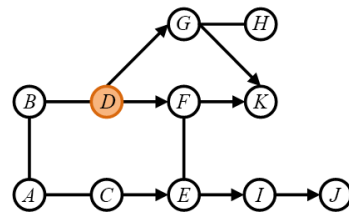


图 7 链图 c

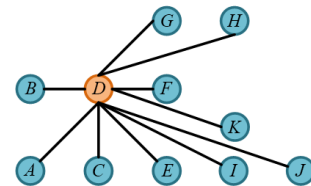


图 8 初始化

$\text{cutSetSize} = 0$ , 当邻接集的大小大于当前条件集的大小时, 开始进行迭代, 在第一次迭代中, 当前条件集的大小  $\text{cutSetSize} = 0$ , 依次遍历  $\text{Adj}(D)$  中的节点, 在给定空集, 判断节点  $D$  与  $C$  的条件独立性, 两个节点条件独立, 将节点  $C$  从  $\text{Adj}(D)$  中删除, 此时  $\text{Adj}(D) = \{A, B, E, F, G, H, I, J, K\}$  (图 9); 然后条件集的大小加一  $\text{cutSetSize} = 1$ , 开始第二次迭代, 继续判断节点  $D$  与  $\text{Adj}(D)$  中每个节点的条件独立性, 给定条件为  $K$ , 节点  $D$  分别与节点  $J, I, H$  独立, 从  $\text{Adj}(D)$  中删除它们, 此时  $\text{Adj}(D) = \{A, B, E, F, G, K\}$  (图 10); 条件集的大小继续加一  $\text{cutSetSize} = 2$ , 开始第三次迭代, 判断节点  $D$  与  $\text{Adj}(D)$  中每个节点的条件独立性, 给定条件为  $\{K, B\}$ , 节点  $D$  与节点  $A$  条件独立, 给定条件为  $\{G, F\}$ , 节点  $D$  与节点  $K$  条件独立, 从  $\text{Adj}(D)$  中删除节点  $A$  和  $K$ , 此时  $\text{Adj}(D) = \{B, E, F, G\}$ ; 条件集的大小加 1, 此时  $\text{cutSetSize} = 4$ , 现在节点  $D$  的邻接集的大小

$|\text{Adj}(D)|=4$ , 二者的大小相等, 结束查找当前节点  $D$  的邻接集, 最终输出节点  $D$  的邻接集为  $\text{Adj}(D)=\{B, E, F, G\}$  (图 11). 同理得节点  $E$  的邻接集  $\text{Adj}(E)=\{C, F, I\}$ .

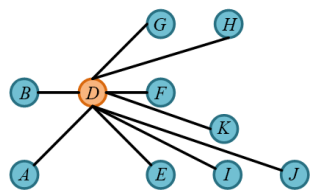


图9 第一次迭代结果

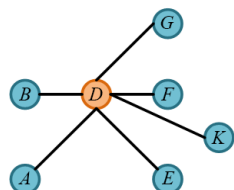


图10 第二次迭代的结果

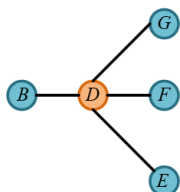


图11 第三次迭代的结果

已知如果两个节点  $D$  和  $E$  是邻接的, 那么一定存在  $D \in \text{Adj}(E), E \in \text{Adj}(D)$ . 由于  $D \notin \text{Adj}(E), E \in \text{Adj}(D)$ , 所以将节点  $E$  从  $\text{Adj}(D)$  中删除, 得到  $\text{Adj}(D)=\{B, F, G\}$  (图 12). 根据得到的邻接集去学习节点  $D$  的配偶集, 并将  $D$  的配偶集添加到  $\text{Mb}(D)$ . 首先初始化  $\text{Mb}(D)=\text{Adj}(D)$ , 节点  $F \in \text{Adj}(D)$ , 节点  $C \in V \setminus \{\text{Adj}(D), \{D\}\}$ , 节点  $D$  与节点  $C$  在二者的分离集下条件独立, 在分离集与节点  $F$  的并集下条件依赖, 则  $C$  为节点  $D$  的一个配偶, 将节点  $C$  添加到  $\text{Mb}(D)$ , 最终节点  $D$  的  $\text{Mb}(D)=\{B, C, F, G\}$  (图 13).

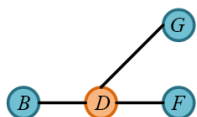


图12  $D$ 的邻接集

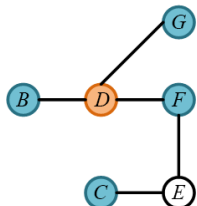


图13  $D$ 的马尔科夫毯

根据节点  $D$  的马尔科夫毯进行链图的结构化恢复. 将  $\text{Mb}(D)$  拆分为  $\text{Adj}(D)$  和  $\text{SP}(D)$ , 分别为  $\text{Adj}(D)=\{B, F, G\}$  和  $\text{SP}(D)=\{C\}$ , 属于邻接集的节点直接用无向边连接 (图 14), 属于配偶集的节点, 在邻接集的参与下, 利用条件独立测试, 判断出边的方向 (图 15), 由此得到节点  $D$  的局部结构, 遍历学习所有节点局部结构, 得到最终的全局链图结构. 由于属于同一马尔科夫等价类的两个链图无法区分, 最终学到链图的模式图.

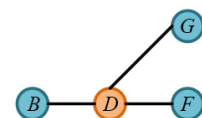


图14  $D$ 的骨架恢复图

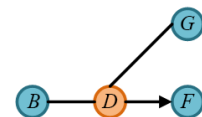


图15  $D$ 的结构恢复图

### 5 实验测试与分析

本节在仿真数据集和标准数据集上将本文提出的 LGCG 算法与 GSCG 算法和 MbLWF 算法从不同的评价指标上进行了对比分析.

本文使用了 5 类评价指标.

- (1) TPR: 恢复结构骨架中正确确定边占所有边的比例.
- (2) FPR: 恢复结构骨架中错误确定边的个数与空缺边个数的比值.
- (3) SHD: 恢复结构与标准结构相比, 需要操作的次数, 即
  - (a) 添加或移除一条边;
  - (b) 增加、去除或者反转一条边的方向.
- (4) ACC: 恢复结构骨架中正确确定边占所有边和空缺边总和的比例.
- (5) TIME: 算法的运行时间.

原则上, 较大的 TPR 和 ACC, 以及较小的 FPR, SHD 和 TIME, 表示性能良好. 所有实验均通过 R 实现.

#### 5.1 仿真数据集的实验测试与分析

本文采用的仿真数据集是依据文献 [12] 提供的方法生成的, 该方法首先生成节点个数为  $P$ 、平均邻接节点个数为  $N$  的链图, 然后依据生成的链图抽样样本容量为  $n$  的满足高斯分布的独立同分布样本. 其中实验参数  $P \in \{10, 20, 40, 60, 80, 100\}$ ,  $N \in \{2, 3, 5, 10\}$ ,  $n \in \{100, 300, 1\ 000, 3\ 000, 10\ 000, 30\ 000\}$ , 样本容量  $n$  取对数后为  $\log_{10} n \in \{2, 2.5, 3, 3.5, 4, 4.5\}$ . 实验测试过程使用显著性水平  $\varepsilon = \{0.01, 0.05\}$ .

在节点个数为100、平均相邻节点个数为10的不同样本容量下,算法运行30次的平均实验结果如图16所示.图16从左到右依次为TPR,FPR,SHD,ACC,TIME,实验对比图横坐标为样本容量,纵坐标为指标值,红色、蓝色和黑色分别表示LGCC算法、MbLWF算法和GSCG算法,实线和虚线分别表示显著性水平 $\varepsilon=0.01$ 和 $\varepsilon=0.05$ ,其他实验结果见<https://github.com/ysjzuihang/LGCC>.

从图16中可以看出,TPR指标值和ACC指标值随着样本容量的增加不断增大且逐渐趋近于1,FPR指标值和SHD指标值随着样本容量的增加而降低(较缓慢)且逐渐趋近于0,TIME指标值会随着样本容量的增加不断增大.

实验结果表明:(1)在大多数情况下,LGCC算法的TPR指标值略优于GSCG算法和MbLWF算法,并且LGCC算法比其他两个算法的FPR指标值低,不同的显著性水平对LGCC算法FPR指标值的影响较小;(2)在SHD指标上,LGCC算法得到了一定的提升,当样本容量增加时,LGCC算法的SHD指标值呈较明显的递减趋势,且较优于其他两个算法;(3)当 $N=2$ 时,LGCC算法的ACC指标值优于GSCG算法,当 $N=3,5,10$ 时,LGCC算法的ACC指标值优于GSCG算法和MbLWF算法,并且在ACC指标上LGCC算法受显著性水平的影响较小;(4)在大多数情况下,LGCC算法的TIME指标值提升较明显.综上所述,实验结果表明LGCC算法表现出较好的性能.

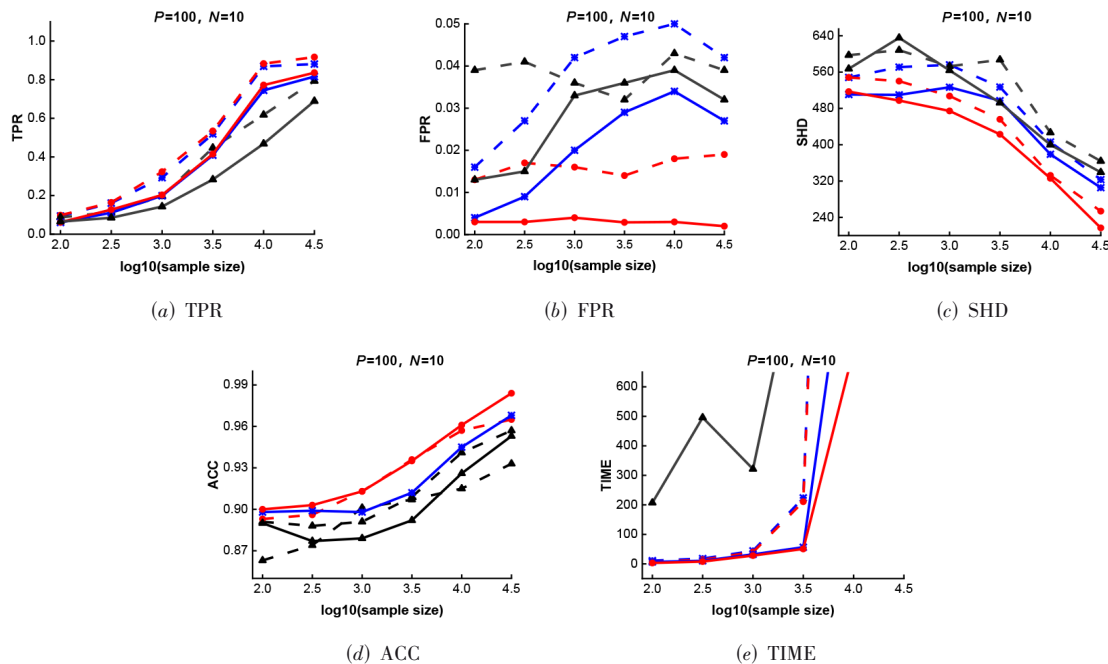


图16  $P=100, N=10$ 的TPR,FPR,SHD,ACC,TIME曲线

## 5.2 标准数据集的实验测试与分析

本文测试了来源于现实生活中的两个标准数据集 Alarm 和 Mildew ([https://pages.mtu.edu/~lebrown/supplements/mmhc\\_paper/mmhc\\_index.html](https://pages.mtu.edu/~lebrown/supplements/mmhc_paper/mmhc_index.html)).其中 Alarm 网是一个用于监测病人的报警信息网络,它总共包含37个节点、46条边和509个参数,该网络的平均马尔科夫毯的大小是3.51;Mildew网是一个冬小麦的真菌疾病粉状霉菌影响网络,它总共包含35个节点、46条边和540150个参数,该网络的平均马尔科夫毯的大小是4.57.本文在两个标准数据集上对LGCC算法、GSCG算法和MbLWF算法进行实验测试( $\varepsilon = \{0.01, 0.05\}$ ).算法运行30次的平均实验结果如表2和表3所示.实验结果表明:LGCC算法在ACC指标上得到一定提升,并且

随着样本容量的增加,ACC指标值呈递增趋势,且较优于其他两个算法;在两个标准数据集上,不同的样本容量不同的显著性水平上,LGCC算法的TIME指标值均低于其他两个算法.综上所述,实验结果表明LGCC算法表现出较好的性能.

## 5.3 显著性测试比较算法之间的差异

为了进一步比较LGCC算法与GSCG算法和MbLWF算法的效率,在零假设(即算法间性能相同)且5%的显著性水平上,本文进行Friedman检验,在仿真数据集和标准数据集下该检验结果为零假设均被拒绝,说明算法间有显著性差异.在仿真数据集上该检验计算得到的LGCC算法、GSCG算法和MbLWF算法的平均序值分别为3,1.75,1.25,在标准数据集上得到的LGCC算法、

表 2 LWF 链图结构学习算法的实验( $\varepsilon=0.01$ )

Network	Algorithm	Size=500					Size=5000				
		TPR	FPR	SHD	ACC	TIME	TPR	FPR	SHD	ACC	TIME
Alarm	GSCG	0.602	0.022	42.95	0.952	1.158	0.729	0.022	30.37	0.962	4.011
	MbLWF	0.831	0.013	27	0.976	0.862	0.946	0.022	26.9	0.977	4.367
	LGCG	0.832	0.001	20.315	0.986	0.661	0.955	0.0007	9.45	0.994	3.526
Mildew	GSCG	0.518	0.045	55.8	0.921	4.922	0.593	0.039	47.365	0.933	9.825
	MbLWF	0.692	0.02	26.965	0.958	1.532	0.883	0.032	24.885	0.962	8.688
	LGCG	0.704	0.003	23.264	0.974	1.351	0.899	0.002	11.983	0.99	6.262

表 3 LWF 链图结构学习算法的实验( $\varepsilon=0.05$ )

Network	Algorithm	Size=500					Size=5000				
		TPR	FPR	SHD	ACC	TIME	TPR	FPR	SHD	ACC	TIME
Alarm	GSCG	0.469	0.3	67.585	0.913	5.727	0.449	0.058	71.1	0.908	12.884
	MbLWF	0.85	0.018	28.25	0.975	1.071	0.957	0.024	28.815	0.975	5.302
	LGCG	0.86	0.003	18.548	0.987	0.96	0.959	0.003	10.335	0.994	4.736
Mildew	GSCG	0.487	0.089	81.4	0.879	25.929	0.441	0.087	82.299	0.877	36.926
	MbLWF	0.77	0.028	27.97	0.957	2.015	0.909	0.04	27.7	0.957	10.56
	LGCG	0.78	0.006	21.55	0.977	1.579	0.919	0.005	12.2	0.989	9.559

GSCG 算法和 MbLWF 算法的平均序值分别为 3, 1, 2, 其中平均序值越高, 算法效率越高. 显然不论在仿真数据集还是在标准数据集上, LGCG 算法的效率都是较高的.

为了进一步区分算法间的性能差异, 本文继续进行了 Nemenyi 后续检验(若两个算法的平均序值之差超出了临界值域 CD, 则表示两个算法性能不相同, 且平均序值越大算法效率越高), 在仿真数据集和标准数据集上得到的临界值域 CD 均为 1.655, 结果表明, 在仿真数据集上 LGCG 算法比 MbLWF 算法的显著性差异更明显, 如图 17 和图 18 所示(显著性测试代码参照文献[26]).

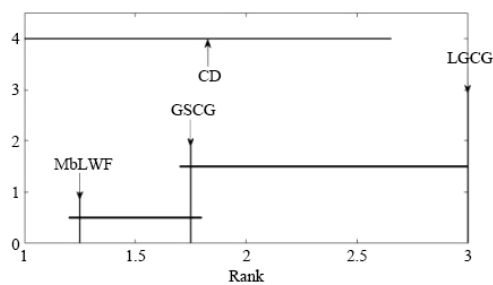


图 17 在仿真数据集上 Nemenyi 检验的关键差异图

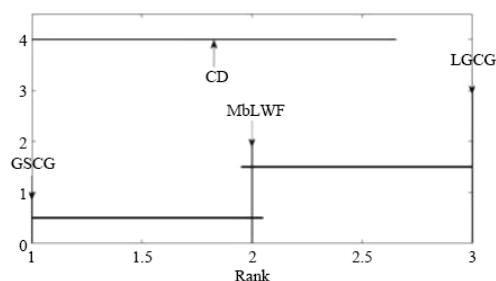


图 18 在标准数据集上 Nemenyi 检验的关键差异图

## 6 结论

LWF 链图模型作为一种工具, 日益得到重视, 从数据中学习链图结构并进行推理和计算已成为该研究领域的重点和难点问题. 本文以基于约束的方法为基础, 提出了学习 LWF 链图结构的 LGCG 算法, 并从理论上证明了算法的有效性, 为链图结构学习提供了新的方案. 本文提出的算法的优势在于算法在学习马尔科夫毯时采用了拆分策略, 将马尔科夫毯拆分为学习邻接集和学习配偶集, 这个拆分策略可以减少学习马尔科夫毯时因错误节点的加入而导致的级联错误, 也可以避免以马尔科夫毯为条件集出现在条件独立测试里, 这样不仅有助于提高条件独立测试的可靠性, 也有助于提高算法的数据效率; 在查找节点的邻接集时采用的是后向策略, 使得条件独立测试给定的条件集总是能够基于最小的分离集, 进一步保障了条件独立测试的正确性. 但是当网络中节点的邻接节点较多时, 高阶条件独立测试计算复杂, 故模型学习时间复杂度会很大. 因此, 下一步的重点工作是针对大规模复杂网络结构的学习问题, 提出相应的改进策略.

## 参考文献

- [1] FEIZI S, MAKHDOUMI A, DUFFY K, et al. Network maximal correlation[J]. IEEE Transactions on Network Science and Engineering, 2017, 4(4): 229-247.
  - [2] 张宏毅, 王立威, 陈瑜希. 概率图模型研究进展综述[J]. 软件学报, 2013, 24(11): 2476-2497.
- ZHANG H Y, WANG L W, CHEN Y X. Research progress of probabilistic graphical models: A survey[J]. Journal

- of Software, 2013, 24(11): 2476-2497. (in Chinese)
- [3] LAPPENSCHAAR M, HOMMERSOM A, LUCAS P J F. Qualitative chain graphs and their application[J]. International Journal of Approximate Reasoning, 2014, 55(4): 957-976.
- [4] LAURITZEN S L, WERMUTH N. Graphical models for associations between variables, some of which are qualitative and some quantitative[J]. The Annals of Statistics, 1989, 17(1): 31-57.
- [5] FRYDENBERG M. The chain graph Markov property[J]. Scandinavian Journal of Statistics, 1990, 17(4): 333-353.
- [6] ANDERSSON S A, MADIGAN D, PERLMAN M D. An alternative Markov property for chain graphs[C]//Proceedings of the Twelfth International Conference on Uncertainty in Artificial Intelligence. Portland: ACM, 1996: 40-48.
- [7] COX D R, WERMUTH N. Linear dependencies represented by chain graphs[J]. Statistical Science, 1993, 8(3): 204-218.
- [8] COX D R, WERMUTH N. Multivariate Dependencies: Models, Analysis and Interpretation[M]. New York: Chapman and Hall/CRC, 2014.
- [9] SONNTAG D, PEÑA J M. Chain graph interpretations and their relations revisited[J]. International Journal of Approximate Reasoning, 2015, 58: 39-56.
- [10] BHATTACHARYA R, MALINSKY D, SHPITSER I. Causal inference under interference and network uncertainty [J]. Uncertainty in Artificial Intelligence, 2019, 2019: 372.
- [11] ZHANG L, ZENG Z, JI Q. Probabilistic image modeling with an extended chain graph for human activity recognition and image segmentation[J]. IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society, 2011, 20(9): 2401-2413.
- [12] MA Z M, XIE X C, GENG Z. Structural learning of chain graphs via decomposition[J]. Journal of Machine Learning Research, 2008, 9: 2847-2880.
- [13] 王静云, 刘三阳, 朱明敏. 基于条件独立测试的链图结构学习算法[J]. 电子学报, 2017, 45(10): 2443-2448.  
WANG J Y, LIU S Y, ZHU M M. Structure learning of chain graphs using the conditional independence tests[J]. Acta Electronica Sinica, 2017, 45(10): 2443-2448. (in Chinese)
- [14] JAVIDIAN M ALI, VALTORTA M, JAMSHIDI P. Learning LWF chain graphs: A Markov blanket discovery approach[J]. Uncertainty in Artificial Intelligence, 2020, 124: 1069-1078.
- [15] DAPHNE KOLLER. 概率图模型:原理与技术[M]. 王飞跃, 韩素青, 译. 北京: 清华大学出版社, 2015.
- [16] VERMA T, PEARL J. Equivalence and synthesis of causal models[C]//Proceedings of the Sixth Annual Conference on Uncertainty in Artificial Intelligence. New York: ACM, 1990: 255-270.
- [17] STUDENÝ M. A recovery algorithm for chain graphs[J]. International Journal of Approximate Reasoning, 1997, 17(2/3): 265-293.
- [18] PEÑA J M. Learning AMP chain graphs under faithfulness[C]//Proceedings of the 6th European Workshop on Probabilistic Graphical Models. Granada: PGM, 2012: 251-258.
- [19] SONNTAG D, PEÑA J M. Learning multivariate regression chain graph under faithfulness[C]//Proceedings of the 6th European Workshop on Probabilistic Graphical Models. Granada: PGM, 2012: 299-306.
- [20] XIE X C, GENG Z, ZHAO Q. Decomposition of structural learning about directed acyclic graphs[J]. Artificial Intelligence, 2006, 170(4/5): 422-439.
- [21] PEÑA J M, SONNTAG D, NIELSEN J D. An inclusion optimal algorithm for chain graph structure learning[C]//Proceedings of the 17th International Conference on Artificial Intelligence and Statistics. Reykjavik: PMLR, 2014: 778-786.
- [22] NIELSEN J D, KOCKA T, PEÑA J M. On local optima in learning bayesian networks[C]//Proceedings of the 19th Conference on Uncertainty in Artificial Intelligence. Acapulco: Morgan Kaufmann, 2003: 435-442.
- [23] PEÑA J M, NILSSON R, BJÖRKEGREN J, et al. Towards scalable and data efficient learning of Markov boundaries[J]. International Journal of Approximate Reasoning, 2007, 45(2): 211-232.
- [24] WANG J Y, LIU S Y, ZHU M M. Local structure learning of chain graphs with the false discovery rate control [J]. Artificial Intelligence Review, 2019, 52(1): 293-321.
- [25] PEÑA J M. Faithfulness in chain graphs: The Gaussian case[C]//Proceedings of the 14th International Conference on Artificial Intelligence and Statistics. Fort Lauderdale: JMLR, 2011, 15: 588-599.
- [26] LING Z L, YU K, ZHANG Y W, et al. Causal learner: A toolbox for causal structure and Markov blanket learning [J]. Pattern Recognition Letters, 2022, 163(C): 92-95.

#### 作者简介



曹付元 男, 1974年5月出生于山西省阳高县. 现为山西大学计算机与信息技术学院教授. 主要研究方向为数据挖掘、机器学习与因果推断. 在国内外发表学术论文70余篇.

E-mail: cfy@sxu.edu.cn



杨淑晶 女, 1996年1月出生于山西省临汾市. 现为山西大学计算机与信息技术学院博士研究生. 主要研究方向为数据挖掘、机器学习与因果推断.

E-mail: ysjzuibang@126.com