一种条件集动态加权的因果结构学习算法

曹冬蕾[1],曹付元[1,2*],王雲霞[1],高小方[1,2]
(1. 山西大学计算机与信息技术学院,山西太原 030006;
2. 山西大学计算智能与中文信息处理教育部重点实验室,山西太原 030006)

摘 要: 基于约束的因果结构学习算法具有不依赖特定函数模型假设的优势且计算效率较高,但其对撞结构定向阶段高度依赖特定条件集的条件独立性检验(Conditional Independence Test,CIT)结果. 尽管近来有研究者提出Shapley-PC算法利用Shapley值整合多个条件集检验以降低CIT误差的影响,但仍未充分考虑不同条件集对定向决策的具体影响,忽略部分关键条件集的重要性,降低定向准确性. 为此,本文提出一种条件集动态加权的因果结构学习算法(Dynamically Weighted Causal Structure Learning,DW-CSL). 该方法核心机制为针对相同规模的条件集通过归一化p-value与Shapley值构建动态权重,细粒度地量化不同条件集对定向决策的贡献差异,从而显著抑制CIT误差在对撞结构中的定向传播. 具体而言,该方法首先基于PC-Stable框架构建因果骨架;其次在对撞结构定向阶段,基于Shapley值提出条件集动态加权的定向决策规则,通过归一化p-value量化条件集贡献差异,使不同条件集的CIT结果具有可比性,再将归一化值作为Shapley边际贡献的权重,实现对未遮蔽三元组的精准定向;最后通过Meek规则定向剩余无向边. 实验结果表明,在合成与基准数据上,该方法较对比方法在对撞结构识别准确性上平均提高4.75%,在边定向准确性上平均提高5.5%,有效提高了因果结构学习的稳定性和准确性.

关键词: 因果结构学习;条件独立性检验;条件集动态加权;归一化p-value;Shapley值

基金项目: 国家自然科学基金(No.U24A20323,No.62376145);山西省科技创新人才团队专项资助项目(No.202204051002016);太行山西省实验室技术攻关专项资助项目(No.THYF-JSZX-24010700)

中图分类号: TP18 **文献标识码:** A **文章编号:** 0372-2112(2025)09-3274-13

电子学报URL:http://www.ejournal.org.cn **DOI:**10.12263/DZXB.20250637

A Causal Structure Learning Algorithm with Dynamic Weighted Condition Set

CAO Dong-lei[1], CAO Fu-yuan[1,2*], WANG Yun-xia[1], GAO Xiao-fang[1,2]
(1. *School of Computer and Information Technology, Shanxi University, Taiyuan, Shanxi* 030006, *China*;
2. *Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Shanxi University, Taiyuan, Shanxi* 030006, *China*)

Abstract: Constraint-based causal structure learning algorithms have the advantage of not relying on specific functional model assumptions and generally offer high computational efficiency. However, their V-structure orientation stage heavily depends on the results of conditional independence tests (CIT) on specific conditioning sets. Although the recently proposed Shapley-PC algorithm integrates multiple condition sets through Shapley value evaluation to mitigate CIT errors, it still fails to adequately account for the varying influence of different condition sets on orientation decisions, thereby overlooking the importance of certain key sets and reducing orientation accuracy. To address this issue, we propose a dynamically weighted causal structure learning (DW-CSL) algorithm. The core idea of the method is to combine normalized p-values with Shapley values to assign dynamic weights to condition sets of the same size, thereby finely quantifying their contribution differences to orientation decisions and effectively suppressing the propagation of CIT errors in V-structure orientation. Specifically, the algorithm first constructs the causal skeleton based on the PC-Stable framework; then, during V-structure orientation, it introduces a dynamic weighted orientation rule that incorporates normalized p-values into Shapley value calculations, making CIT results from different condition sets comparable and enabling precise orientation of unshielded tri-

收稿日期:2025-07-21;录用日期:2025-09-16;责任编辑:覃怀银
*通讯作者:曹付元

ples; finally, the remaining undirected edges are oriented using Meek's rules. Experimental results on both synthetic and benchmark datasets demonstrate that, compared with baseline methods, DW-CSL improves V-structure recognition accuracy by an average of 4.75% and edge orientation accuracy by an average of 5.5%, thereby enhancing the stability and overall accuracy of causal structure learning.

Key words: causal structure learning; conditional independence test; dynamic weighted condition set; normalized *p*-value; Shapley value

Foundation Item(s): National Natural Science Foundation of China (No.U24A20323, No.62376145); Shanxi Province Science and Technology Innovation Talent Team Special Project (No.202204051002016); Key Technologies Program of Taihang Laboratory in Shanxi Province (No.THYF-JSZX-24010700)

1 引言

因果结构学习(Causal Structure Learning, CSL)[1~3]是数据驱动因果推理领域中的核心任务之一,其目标是提取数据中变量之间的因果关系,并以有向无环图(Directed Acyclic Graph, DAG)[4]进行表示,广泛应用于生态气候学[5,6]、生物医学[7~9]、经济学[10,11]、社会科学[12]及人工智能[13,14]等多个领域[15]. 因果结构学习的主流方法之一是基于约束的方法,核心思想是利用条件独立性检验(Conditional Independence Test, CIT)[16,17]结合d-分离(d-separation)理论[4]来推断变量间的因果结构. 该类方法不依赖特定函数模型假设,且计算效率较高[18]. 代表性算法包括Peter-Clark(PC)算法[19]及其改进版本[如PC-Stable[20]、Conservative PC(Conservative Peter-Clar)[21]等]. 这类算法通常包括2个阶段:在无向图学习阶段,算法从完全图出发,使用CIT等统计方法识别变量间的独立关系并据此移除边,构建出变量间的无向图结构;在方向学习阶段,算法依据*V*-结构(*V*-Structure)等局部结构特征来确定部分边的方向[22]. 具体而言,当存在未遮蔽三元组(Unshielded Triplet, UT) $X_i-X_j-X_k$ 满足 $X_i\perp X_k|S$ 并且 $X_j\notin S$,则推断该UT为 $X_i\to X_j\leftarrow X_k$,且 X_j 为碰撞点,条件集 S 表示在条件独立性检验中用于检验变量之间是否独立的变量集. 然而对撞结构的判定过程高度依赖于CIT结果的准确性,而现有方法的正确性依赖于所有CIT判断均准确的理想假设. 但在现实中,这一理想假设往往难以满足,致使CIT结果不稳定,导致对撞结构误判[23],并在后续定向阶段造成累积性偏差,最终削弱结构学习的可靠性和有效性[24].

为了提高基于约束算法的可靠性,Zhalama等人[25]提出CIT执行得到的不确定结果可以通过一种图形推理的方法纠正. 郝志峰等人[23]提出一种增强的条件独立性检验方法,通过抑制待测变量的外部噪声干扰提升检验精度,但依旧局限于单一条件集的结果,未能整合多条件集间的统计信息. Russo等人[26]提出了基于Shapley值的因果结构学习方法(Shapley-PC),引入了合作博弈论中用于量化个体在合作中贡献的Shapley值理论[27~29],将每个条件集视为"参与者",将对撞结构定向决策规则视为"合作博弈",以量化各条件集在CIT决策中的边际贡献,缓解定向阶段对单一条件集选择的敏感性,但在计算边际贡献时对相同规模的条件集作等权处理,忽略了条件集组成差异反映出的依赖性强弱信息. 上述方法尽管在纠错能力、检验精度和局部判定稳定性方面取得进展,但未能从条件集整体信息出发刻画依赖性差异,未充分考虑不同条件集对定向判断的具体影响,忽略部分关键条件集的重要性,降低定向准确性. 例如图1展示了考虑变量 X_1 与 X_2 的依赖性时,2个规模相同但组成不同的条件集 $\{X_3,X_4\}$、$\{X_5,X_6\}$ 在CIT的结果上分别为条件独立和条件依赖(即 $X_1\perp X_2|\{X_3,X_4\}$, $X_1\not\perp X_2|\{X_5,X_6\}$). 若对所有条件集赋予相同权重,则无法反映其对定向决策的重要性差异,致使因噪声等因素产生误判的条件集与判定正确的条件集被等同处理,进而导致错误信息掩盖真实依赖关系,最终引发对撞结构定向的不稳定性和准确性下降.

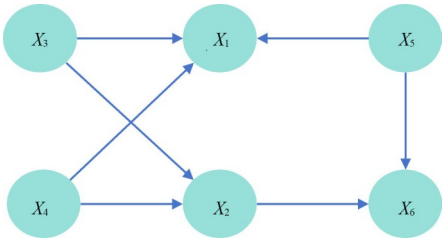

图1　不同条件集对变量间独立性的不同影响

为了全面衡量不同条件集在定向决策中的贡献并提升算法在有限样本下的稳定性与准确性,本文提出一种条件集动态加权的因果结构学习方法. 该方法通过归一化*p*-value使不同条件集的条件独立性检验结果具有可比性,进而将归一化后的值作为Shapley边际贡献的权重,以量化不同条件集对依赖性判定的贡献差异. 该方法不仅保留了Shapley值衡量平均边际贡献的公平性原则,整合了多条件集的依赖性信息,还能在概率意义上凸显条件集所体现的依赖性差异,实现相同规模条件集间依赖性差异的细粒度衡量. 方法基于PC-Stable框架实现,在保持原有结构发现机制不变的基础

上融入新的定向决策规则. 该方法能够在多条件集的统计信息中自适应抑制异常结果的干扰,降低检验误差对定向结果的影响,显著增强方法的鲁棒性及对撞结构定向的稳定性与准确性. 实验基于多个合成与基准数据集展开,结果显示本文方法在因果结构恢复的准确性方面较基线算法具有显著优势.

2 相关工作

近年来,因果结构学习方法在从数据中恢复变量间因果关系方面取得了长足进展. 众多研究者提出的算法主要可归纳为3大类:基于约束的方法、基于评分的方法以及基于函数因果模型的方法,其中基于约束的方法因其原理直观、计算效率相对较高,在众多应用场景中受到广泛关注[30~33]. 作为该领域最具代表性的算法,PC算法[19]通过逐步增加条件集合的大小,对变量对进行条件独立性检验,构建因果图的骨架,并依据分离集信息定向未屏蔽三元组(UT),最终生成部分有向无环图(Completed Partially Directed Acyclic Graph, CPDAG)[4]. 然而,PC算法的实际应用效果受到其内在机制的限制,特别是在处理有限样本和高维数据的复杂场景时,PC算法对条件独立性检验结果的高度依赖,使定向阶段容易受到噪声和误差的干扰,导致因果图恢复的不稳定性和准确性下降[34].

针对PC算法的稳定性缺陷,研究者们提出了多种改进方案. Murphy等人[20]提出的PC-Stable算法解决了骨架构建过程中的顺序敏感性,其核心在于通过固定条件集规模,对所有变量对进行独立性检验后统一删除边,克服了原PC算法中因变量检验顺序不同引起的不稳定性. 尽管PC-Stable在骨架构建方面取得了一定改进,但其在定向阶段仍然直接继承自PC算法,依赖于条件独立性检验的结果,错误检验的影响难以完全避免. 为进一步提高定向阶段的鲁棒性,后续研究聚焦于改进对撞结构的判定规则. Ramsey等人[21]提出了Conservative PC(CPC)算法,该方法将忠实性假设拆分为邻接忠实性和定向忠实性,并要求在定向未屏蔽三元组时,候选碰撞结点必须在所有分离集中均未出现才能被认定为真正的碰撞器(*V*-结构). 这种极其严格的条件虽然在一定程度上降低了因单个CIT错误带来的错误定向风险,但也使定向结果显得过于保守,可能导致真实存在的对撞结构无法被识别. 随后,为了在保守性和灵活性之间寻求平衡,Murphy等人[20]进一步提出了Majority PC(MPC)算法. MPC算法在定向过程中采用了多数决策略,即候选碰撞结点只需要在少于1/2的分离集中出现,即可认为存在足够证据支持其作为对撞结构. 这种相对宽松的判定标准提升了识别真实对撞结构的灵活性. 除了上述规则改进的方法外,也有研究者探索利用CIT的连续统计量信息. Ramsey提出的PC-Max算法进一步利用了独立性检验中p值的连续信息. 该算法在对每个未屏蔽三元组进行定向时,选取p值最大的条件集合作为依据,仅在该条件集合中候选变量不出现时才定向为对撞结构. PC-Max算法在有限样本条件下表现出较高的鲁棒性,但其本质仍未脱离对单一CIT结果的依赖,未能有效整合其他条件集提供的潜在补充信息.

为解决PC系列算法在定向过程中对条件独立性检验结果依赖性过高的问题,近年来Russo等人[26]在Shapley-PC算法中首次引入合作博弈论中的Shapley值,该方法将参与对撞结构定向决策的不同条件集视为“参与者”,以合作博弈的形式评估其对定向决策的边际贡献. 该方法有效缓解对单一条件集选择的敏感性,是提升定向鲁棒性的重要探索. 然而Shapley-PC在计算边际贡献时对所有条件集进行等权处理,未能充分考虑关键条件集的重要性.

3 基础知识

本研究关注基于约束的因果结构学习问题,致力于解决定向阶段对CIT结果高度依赖对结构学习带来的不确定性问题. 本文的理论结果和可靠性算法是基于因果图模型建立的. 为此,为了使后续方法描述更加严谨,在本节中对本文涉及的基本符号、图论概念、条件独立性检验以及相关统计量进行形式化定义.

在因果图模型[35]中,使用有向无环图(Directed Acyclic Graph, DAG)描述变量间的因果关系. 设因果图 $\boldsymbol{G}=(V,\boldsymbol{E})$[36] 由节点集 $V=\{X_1,X_2,\cdots,X_a\}$ 和边集 $\boldsymbol{E}\subseteq V\times V$ 组成,节点表示随机变量,而边反映了变量之间的关系. 若 $(X_i,X_j)\in\boldsymbol{E}$,则表示 X_i 对 X_j 具有直接影响. 对于任一节点 X_i,其邻接集合定义为 $\mathrm{Adj}(\boldsymbol{G},X_i)=\{X_j|(X_i,X_j)\in\boldsymbol{E}\text{或}(X_j,X_i)\in\boldsymbol{E}\}$,而父节点集合则为 $\mathrm{Pa}(\boldsymbol{G},X_i)=\{X_j|X_j\to X_i\}$. 此外,若存在一条从 X_j 到 X_i 的有向路径,则称 X_i 为 X_j 的后裔节点. 对于3个节点 X_i,X_j,X_k 来说,若 X_i 与 X_k 在图中不直接相邻,而均与 X_j 相邻,则称 (X_i,X_j,X_k) 为未遮蔽三元组(UT),记作 $X_i-X_j-X_k$.

在基于约束的因果结构学习方法中,条件独立性检验是确定变量间直接因果边的核心步骤. 对于变量 X_i 与 X_j 以及条件集 $S\subseteq V\backslash\{X_i,X_j\}$,若在给定条件集 S 下,X_i 与 X_j 条件独立,则记作 $X_i\perp X_j|S$. 为便于形式化描述,定义指示函数:

$$I\left(X_i,X_j|S\right)=\begin{cases}1, & \text{若}X_i\perp X_j|S\text{成立}\\ 0, & \text{否则}\end{cases} \tag{1}$$

对于因果图 $\boldsymbol{G}$ 中的任意2个相邻节点 X_i 与 X_j,其分离集 $\mathrm{SepSet}(X_i,X_j)$ 定义为满足以下条件的最小条件集

$S\subseteq V\backslash\{X_i,X_j\}$:若$X_i$与$X_j$在条件集$S$下独立,且对任意条件集$S$的真子集$S'$,$X_i$与$X_j$在$S'$下均不独立,则$S$为$X_i$与$X_j$的最小分离集.

在实际应用中,通常通过统计方法计算一个p值$p(X_i,X_j|S)$,用于衡量在原假设H_0:$X_i\perp X_j|S$下观察到数据的概率. 当$p(X_i,X_j|S)\geqslant\alpha$(其中$\alpha$为预设显著性水平)时,认为$X_i$与$X_j$在条件集$S$下独立;反之,则视为条件依赖.

为了保证基于约束的因果结构学习方法的有效性,本文基于以下基本假设构建理论基础.

(1)因果充分性假设[4]:观测变量集合V包含了所有直接影响V内变量的因果因素,即不存在隐藏的共同原因.

(2)因果马尔科夫假设[4]:对于具有因果充分性的变量集V而言,在已知变量父亲节点的条件下,所有变量与他们的非后裔节点互相条件独立.

(3)因果忠实性假设[4]:观测数据中的所有条件独立性关系均由因果图$\boldsymbol{G}$所隐含,反之亦然. 也就是说,若在$\boldsymbol{G}$中某条件集S将X_i与X_j进行d-分离,则在真实数据分布中$X_i\perp X_j|S$.

在因果结构定向问题中引入Shapley值的思想,可以客观衡量各条件集对对撞结构定向决策的贡献. Shapley值源于合作博弈论,用以量化各参与者对整体收益的边际贡献. 设有参与者集合$N=\{1,2,\cdots,n\}$以及价值函数v: $2^N\to\mathbb{R}$,则对于任一参与者$i\in N$,其Shapley值定义为

$$\phi_v(i)=\sum_{T\subseteq N\backslash\{i\}}\frac{|T|!(n-|T|-1)!}{n!}\left[v(T\cup\{i\})-v(T)\right] \quad (2)$$

其中,$T\subseteq N\backslash\{i\}$表示不包含参与者的任意子集.

在本文的研究方法(见第4节)中将“参与者”对应于条件集中的某一变量,价值函数$v(T)$与条件独立性检验所获得的p值相关联,以反映在给定条件集下变量间依赖性的强度.

4　研究方法

本节针对因果图定向阶段在有限样本条件下因单次条件独立性检验结果易受噪声干扰、导致定向不准确的问题,考虑了如何充分利用条件集的依赖程度信息提高因果结构学习方法的准确性和鲁棒性,提出一种基于条件集动态加权的因果结构学习方法(Dynamically Weighted Causal Structure Learning, DW-CSL). 该方法的基本思想是:在因果无向图学习阶段,基于PC-Stable框架构建因果骨架;在因果方向学习阶段,基于Shapley值提出条件集动态加权的定向决策规则,并结合归一化权重提升定向判断的精度与鲁棒性.

4.1　问题定义

尽管研究者们提出的方法有效缓解了对单一条件集选择的依赖,然而直接将Shapley值引入定向阶段存在一个明显问题:虽然不同条件集的大小可能相同,但内部组成的变量不同,这些不同的条件集在判断变量之间的依赖性方面发挥着截然不同的作用,现有的方法为所有大小相同的条件集赋予了相同的权重,忽略了各条件集对最终定向决策的重要性差异.

针对图1中展示的现象,本节进一步对2个规模相同但组成不同的条件集在CIT中的表现差异进行具体分析,以揭示其对对撞结构定向可能造成的影响. 图中包含节点X_1、X_2、X_3、X_4、X_5和X_6,其中箭头代表直接因果关系. 假设在对变量X_1与X_2进行条件独立性检验时,分别采用不同的条件集进行测试:当条件集选取为$\{X_3,X_4\}$时,检验结果显示X_1与X_2条件独立;而当条件集选取为$\{X_5,X_6\}$时,检验结果却表明X_1与X_2存在依赖关系. 当直接将Shapley值引入定向阶段时,对于所有具有相同大小的条件集会赋予同等权重,使得来自$\{X_3,X_4\}$和$\{X_5,X_6\}$的检验结果在定向决策中具有同等影响. 然而,实际上不同组成元素的条件集对判断对撞结构的影响存在差异. 尤其是在有限样本条件下,CIT结果可能因噪声或样本不足而出现偏差,这种错误检验结果如果与其他条件集赋予相同的权重,就可能对最终对撞结构的定向产生影响. 例如,在上述情形中,若条件集$\{X_3,X_4\}$错误地判定X_1与X_2条件独立,而条件集$\{X_5,X_6\}$正确地反映了二者的依赖关系,在Shapley定向策略中,由于两者被视为同等重要,错误结果可能“掩盖”正确结果,导致因果图定向的不稳定、不准确.

因此,本文所研究的问题为:如何在条件独立性检验结果存在不确定性的条件下,通过全面地考虑不同条件集对定向决策的影响,提高基于约束方法中对撞结构定向的准确性,减少CIT误差对定向决策的干扰,更准确地恢复观测数据中的真实因果关系.

4.2　因果无向图学习

本方法的因果无向图学习过程即因果骨架构建过程基于PC-Stable框架实现,具体包含以下4步. 首先,初始化完全无向图,构建包含所有变量的全连接无向图$\boldsymbol{G}$,初始图中任意节点对均通过无向边连接. 其次,执行条件独立性检验,从阶数$l=0$(即空条件集)开始,逐级递增至预设最大阶数max_degree,在每阶l中遍历所有相邻节点对(X_i,X_j),并枚举大小为l的候选条件集$S\left\{S|S\subseteq V\backslash\{X_i,X_j\},|S|=l\right\}$. 接着进行边删除,若存在条件集$S$使得$X_i$与$X_j$满足条件独立,则删除无向边$X_i-X_j$并记录分离集SepSet$(X_i,X_j)=S$. 最后,在每阶完成所有节点对检验后统一删除独立边,避免传统PC算法因检验顺序差异导致的结果波动.

4.3　因果方向学习

在因果方向学习阶段,本文提出一种条件集动态加权的定向策略:在对某个UT进行判定时,将所有大

小相同的条件集的p-value归一化，然后利用归一化结果作为权重因子，衡量候选节点在不同条件集合中对变量间独立性变化的平均边际贡献，将该加权贡献值称为归一化的Shapley独立性值（Weighted Shapley Independence Value，WSIV）.

设对于一个已知的骨架$\boldsymbol{C}$，有未遮蔽三元组$X_i-X_j-X_k$，令Z为对X_i与X_k构成条件集的所有集合，即

$$Z=\left\{S|S\subseteq\left(\mathrm{Adj}\left(\boldsymbol{C},X_i\right)\cup\mathrm{Adj}\left(\boldsymbol{C},X_k\right)\right)\backslash\{X_j\}\right\} \tag{3}$$

并将这些条件集按照其变量个数进行分组，每个分组内所有条件集均为相同大小. 记组内任一条件集为S^q_{num}（其中下标num表示该组合中条件集包含的变量数，上标q为该组合中各条件集的编号），$p(X_i,X_k|S^q_{\mathrm{num}})$表示在大小为num的第$q$个条件集下对$X_i$与$X_k$进行条件独立性检验得到的$p$-value，$n$表示对于所考虑的UT中所有参与构成邻接集合的变量数量. 归一化过程即计算同组内所有条件集的p-value之和，即$\sum_{a=1}^{m}p(X_i,X_k|S^a_{\mathrm{num}})$归一化后整体的权重为$w^q_{\mathrm{num}}$，即

$$w^q_{\mathrm{num}}=\frac{p(X_i,X_k|S^q_{\mathrm{num}})}{\sum_{a=1}^{m}p\left(X_i,X_k|S^a_{\mathrm{num}}\right)}\times\frac{\mathrm{num}!(n-\mathrm{num}-1)!}{n!} \tag{4}$$

则X_j所考虑的UT中的归一化Shapley独立性值（WSIV）为

$$\phi\left(X_j,\{X_i,X_k\}\right)=\sum_{S\in Z}w^q_{\mathrm{num}}\cdot\left[I\left(X_i,X_k|S^q_{\mathrm{num}}\cup\{X_j\}\right)-I\left(X_i,X_k|S^q_{\mathrm{num}}\right)\right] \tag{5}$$

其中，$I(\cdot)$为条件独立性检验的指示函数，其取值依据预设的显著性水平α判定[见式(1)]. 该式通过使用大小相同但组成不同的条件集的p-value进行归一化处理，令大小相同的同一组条件集在比较贡献时具有可比性，从而更细粒度地衡量候选节点X_j对X_i与X_k独立性状态所引入的变化. 如果计算得出的WSIV值为负，则表明候选节点的加入使原本独立的X_i与X_k之间依赖性增强，则将该未遮蔽三元组UT定向为对撞结构，即确定X_i与X_k均指向X_j.

在条件集动态加权定向决策规则中，核心在于如何精确衡量候选节点在各条件集组合中的边际贡献. Shapley值公式用于合作博弈中量化个体对整体收益的平均边际增益，能够在所有可能的条件集排列中公平分配贡献. 在本方法中，将“玩家”映射为组成条件集的变量，而“价值函数”$v(T)$则取为条件独立性指示函数得到的结果，表示在给定条件集S下，X_i与X_k的独立性程度. 由此，候选节点X_j对条件集S^q_{num}的边际贡献为

$$\Delta_{S^q_{\mathrm{num}}}(X_j)=I\left(X_i,X_k|S^q_{\mathrm{num}}\cup\{X_j\}\right)-I(X_i,X_k|S^q_{\mathrm{num}}) \tag{6}$$

鉴于相同大小的条件集可能因其具体组成而在实际检验中表现出不同显著性水平，本文在Shapley值计算中引入p-value归一化机制，即使用归一化权重$p(X_i,X_k|S^q_{\mathrm{num}})\Big/\sum_{a=1}^{m}p(X_i,X_k|S^a_{\mathrm{num}})$来进行加权. 由于Shapley值的加权因子$|S|!(n-|S|-1)!/n!$保证了所有排列的公平性，故最终归一化后的WSIV值不仅反映了单个条件集中p-value的大小，也综合了所有可能组合的平均效应. 这种定向策略不仅综合了多个条件集的信息，降低了CIT误差对定向判断的影响，而且将p-value的大小信息引入定量评估中，更细粒度地考虑了不同条件集对定向判断的贡献，使决策过程更具灵敏性和鲁棒性.

算法1展示了DW-CSL算法的伪代码，该算法分为3个阶段：骨架构建阶段从一个完全无向图开始，逐步进行条件独立性检验，删除不符合条件的边并记录分离集，从而构建稳定的骨架图（第2~7行）；对撞结构定向阶段通过动态权重计算WSIV，判定碰撞节点（第9~20行）；剩余边定向阶段应用Meek规则确保无环性（第21行）. 理论上可以证明，在满足因果忠实性、样本量充足且CIT判断无误的理想条件下，当候选节点X_j的确为碰撞节点时，对应的WSIV值为负；反之，若X_j非碰撞节

算法1 DW-CSL

输入：数据集$\boldsymbol{D}$，显著性水平α，最大阶数max_degree

输出：部分有向无环图CPDAG

1. 骨架构建(基于PC-Stable)
2. 初始化完全无向图
3. FOR l= 0 TO max_degree DO
4. 对每对相邻节点(X_i,X_j)，遍历大小为l的条件集S
5. IF $\exists S$ 使得$X_i\perp X_j|S$ THEN
6. 删除边X_i-X_j，记录S为分离集
7. END FOR
8. 对撞结构定向
9. FOR 每个未遮蔽三元组$X_i-X_j-X_k$ IN $\boldsymbol{C}$ DO
10. $Z=\left\{S|S\subseteq\left(\mathrm{Adj}(\boldsymbol{C},X_i)\cup\mathrm{Adj}(\boldsymbol{C},X_k)\right)\backslash\{X_j\}\right\}$
11. FOR num = 0 TO $|Z|$ DO
12. FOR 每个条件集S^q_{num} IN $|S^q_{\mathrm{num}}|=\mathrm{num}$ DO
13. 计算$p\left(X_i,X_k|S^q_{\mathrm{num}}\right)$
14. 计算归一化权重w^q_{num}
15. END FOR
16. END FOR
17. 计算$\phi\left(X_j,\{X_i,X_k\}\right)$
18. IF $\phi\left(X_j,\{X_i,X_k\}\right)<0$ THEN
19. 定向为$X_i\rightarrow X_j\leftarrow X_k$
20. END FOR
21. 应用Meek规则定向剩余边

点,则WSIV值为正. 该性质为因果图中对撞结构的判定提供了明确的理论依据.

此外,条件集动态加权的定向策略在实际中还具有显著的鲁棒性优势:即使存在部分条件集的误判,因其对应p-value较低而被赋予较小权重,降低了误判结果对整体定向的主导影响,显著提升了因果结构恢复的稳定性与准确性.

4.4　理论分析

在因果忠实性假设成立的前提下,DW-CSL方法具有良好的理论收敛性与鲁棒性. 当样本量趋于无穷时,条件独立性检验的p-value收敛至真实概率,归一化权重准确反映条件集对依赖性判断的贡献程度,WSIV趋于负值当且仅当候选节点为真实碰撞节点,从而正确识别对撞结构,确保输出CPDAG收敛至真实因果图的马尔可夫等价类,该过程保留了PC框架所具有的渐近一致性属性. 在有限样本条件下,条件集动态加权的定向策略有效抑制了因噪声或样本不足导致的异常p值影响,使整个V构定向步骤在理论上满足PC算法要求的渐近一致性[37]、完备性与正确性[32].

引理1　在因果充分性假设、因果马尔科夫假设和因果忠实性假设成立的基础上,DW-CSL的权重机制设置是合理且有效的.

证明　在条件独立性检验中,设检验统计量服从零假设$H_0: X_i \perp X_j | S$下的渐进分布,得到的p值可视为对H_0成立与否的证据强度刻画. 当H_0成立时,p值在$[0,1]$之间服从均匀分布;当H_0不成立时,p值的分布依赖于样本量及真实统计量值,p值越小意味着依赖性证据越强,因此不同的p值能够反映条件集对变量依赖程度的差异. 证毕.

由引理1可知,p-value归一化能够使同规模条件集的统计检验结果具有可比性,Shapley值保证了对所有条件集排列的公平性与完备性. 二者结合不仅保留了Shapley值刻画平均边际贡献的公平性原则,还能在概率意义上凸显条件集所体现的依赖性差异.

在有限样本条件下,由于噪声扰动或样本不足,单个条件集的检验结果容易偏离真实独立关系,进而影响对撞结构定向的准确性. DW-CSL所引入的归一化权重机制,有效抑制异常p-value对定向结果的放大影响. 通过整合同一规模下多个条件集的信息,DW-CSL能显著缓解CIT误差对最终结构恢复的干扰,从而在样本规模受限情境下依然具备较强的鲁棒性与判别能力.

综上所述,本文提出的DW-CSL方法在因果结构学习中的优势体现在多个方面:在理论上,保留了传统PC算法在充分样本条件下的正确性与一致性;在方法机制上,通过提出新颖的定向决策机制整合多个条件集的统计证据,显著提升了定向决策对局部误差的容忍度;在实践中,为因果图定向提供了一种更具精度与稳健性的判断机制,尤其适用于样本受限、图结构复杂等现实应用场景.

4.5　时间复杂度分析

本算法主要包括骨架构建、对撞结构定向以及Meek规则定向三个阶段,其中时间复杂度主要由前两个阶段决定. 在骨架构建阶段,算法基于PC-Stable框架对每对变量执行条件独立性检验,检验次数与图中变量数d和条件集最大规模K有关,其时间复杂度为$O(d^{K+2}\cdot t)$,其中t表示单次条件独立性检验的计算开销. 在对撞结构定向阶段,针对每个未遮蔽三元组,算法遍历所有大小不超过$2K$的候选条件集,进行归一化Shapley独立性值计算,复杂度近似为$O(d^3\cdot 2^{2K}\cdot t)$. 当最大邻接度$K$为常数时,两部分的总时间复杂度可简化为$O(d^3\cdot K)$,在因果图结构稀疏的实际应用中具有良好的可扩展性. Meek规则定向阶段复杂度为$O(d^2)$,相对较低.

5　实验与结果分析

5.1　实验设计

5.1.1　数据集与参数设置

本实验旨在系统评估提出的DW-CSL方法在合成数据与基准数据中的性能,验证其在有限样本、噪声干扰及复杂图结构下的鲁棒性与准确性.

在合成数据实验中,首先生成2类常见的图结构:Erdős-Rényi(ER)随机图与Scale-Free(SF)无标度图,节点数覆盖$|V|\in\{10,50\}$,边密度设定为$\rho\in\{2,4\}$. 给定真实的DAG,合成数据的生成机制服从线性结构方程模型(Structural Equation Model,SEM):

$$X_j=\sum_{X_i\in \mathrm{Pa}(X_j)}\omega_{ij}\cdot X_i+\varepsilon_j \tag{7}$$

其中,权重ω_{ij}服从均匀分布$[\omega_{ij}\sim U(-1,1)]$,用于量化父节点对子节点的直接影响强度. ε_j表示随机生成的噪声,实验中引入高斯、指数及均匀噪声模拟真实场景的干扰.

在基准数据实验中,为全面评估算法的性能,本研究选取了bnlearn数据库中的一系列经典贝叶斯网络数据集. 这些数据集在相关领域获得广泛验证,涵盖了离散变量、连续变量及混合型变量场景,充分体现了现实数据的多样性特征. 尤为关键的是,其底层网络结构并非随机生成,而是源于真实世界领域的专业知识或由专家构建的因果模型(如医疗诊断、工业流程控制、基因调控等),确保了所蕴含因果关系的合理性与实际意义. 这为客观验证因果结构学习算法在恢复精度与鲁棒性方面的性能提供了坚实的基础. 本文选择了部分数据集进行实验,具体数据集细节如表1所示.

在异常CIT敏感性实验中,为系统评估DW-CSL对

表1 贝叶斯网络数据集结构信息

数据集	类型	节点数	边数	平均度
Alarm	离散	37	46	1.24
Insurance	离散	27	52	1.93
Ecoli70	连续	46	70	1.52
Mehra	混合	24	71	3.64

异常条件独立性检验结果的抑制能力，本文基于bnlearn标准贝叶斯网络数据集开展高斯噪声注入实验．在骨架构建阶段，对每一次CIT返回的p-value注入高斯噪声$N(0,\sigma^2)$，噪声比例rate$=\{1\%,2\%,5\%\}$，其中σ设为0.002，噪声注入通过环境变量统一控制，确保所有方法在同一噪声场景下输入一致．

参数设置方面，样本量按比例因子生成，比例因子为$s\in\{100,500,1\,000\}$，每组实验重复10次以计算性能指标的均值与标准差，确保结果的统计稳健性．

5.1.2 基线方法与评价指标

为了全面评估DW-CSL方法在因果结构恢复中的表现，本文选取多种有代表性的基于约束的因果发现算法作为基线进行比较，涵盖经典PC算法及其主要改进版本，包括：解决顺序敏感性的PC-Stable算法，采用保守定向策略的Conservative PC算法，采用多数投票策略的Majority PC算法，基于最大p值策略的PC-Max算法，以及使用Shapley值的Shapley-PC算法．

为了全面衡量因果结构学习过程中结构恢复和边方向定向的整体性能，本文引入了三类具有代表性的评价指标，分别从不同维度评估算法的效果．ArrowHead F_1（AH-F_1）用于评估因果图中所有有向边（即箭头）的预测准确性，该指标综合了方向预测的精确率与召回率，能够反映算法在边定向方面的整体能力．V-structure F_1（V-F_1）用于衡量对UT中对撞结构的识别效果，是对局部结构定向能力的敏感评估指标，能够直观反映算法在Structural Hamming Distance（SHD）用于量化估计图与真实图在结构上的差异．以上三个评价指标覆盖了全局定向精度、局部因果结构识别能力及结构整体差异性，兼顾了宏观与微观评估维度，从多个角度全面、细致地反映算法在因果结构学习任务中的表现优劣．

5.2 实验结果

5.2.1 合成数据集实验与分析

在合成数据集上的实验结果表明，本文提出的DW-CSL算法在边定向精度（AH-F_1）、对撞结构识别（V-F_1）及结构汉明距离（SHD）3项核心指标上均优于现有方法，展现出优异的因果结构恢复能力与鲁棒性，如表2及图2~图4所示．

表2 10节点和50节点的ER和SF图的AH-F_1和V-F_1

节点	方法	ER2		ER4		SF2		SF4	
		AH-F_1	V-F_1	AH-F_1	V-F_1	AH-F_1	V-F_1	AH-F_1	V-F_1
10	PC	0.36±0.28	0.46±0.36	0.15±0.16	0.22±0.32	0.67±0.23	0.81±0.29	0.32±0.27	0.49±0.41
	CPC	0.42±0.26	0.59±0.33	0.15±0.17	0.23±0.33	0.74±0.13	0.88±0.18	0.34±0.23	0.54±0.36
	MPC	0.42±0.26	0.58±0.34	0.15±0.17	0.22±0.33	0.74±0.13	0.88±0.18	0.36±0.23	0.59±0.36
	PC-Max	0.50±0.22	0.67±0.3	0.07±0.12	0.09±0.24	0.73±0.2	0.90±0.22	0.39±0.26	0.59±0.4
	SPC	0.63±0.16	0.86±0.16	0.25±0.16	0.36±0.42	0.82±0.08	0.99±0.02	0.54±0.16	0.84±0.23
	DW-CSL	**0.68±0.10**	**0.88±0.05**	**0.31±0.10**	**0.45±0.35**	**0.85±0.07**	**0.99±0.01**	**0.58±0.13**	**0.88±0.18**
50	PC	0.25±0.3	0.31±0.37	0.04±0.09	0.08±0.17	0.63±0.37	0.67±0.4	0.25±0.35	0.28±0.4
	CPC	0.4±0.36	0.5±0.44	0.06±0.1	0.14±0.24	0.53±0.44	0.58±0.48	0.4±0.39	0.46±0.45
	MPC	0.35±0.34	0.42±0.41	0.04±0.09	0.08±0.18	0.51±0.44	0.57±0.49	0.36±0.39	0.42±0.45
	PC-Max	0.63±0.27	0.81±0.34	0.05±0.11	0.12±0.27	0.56±0.44	0.61±0.48	0.59±0.37	0.69±0.43
	SPC	0.75±0.06	0.98±0.03	0.19±0.15	0.48±0.36	0.9±0.04	**1.0±0.0**	0.83±0.07	0.99±0.03
	DW-CSL	**0.78±0.05**	**0.99±0.0**	**0.22±0.10**	**0.52±0.30**	**0.92±0.05**	0.99±0.01	**0.85±0.08**	**0.99±0.01**

注：加粗数据表示最优结果．

在稀疏网络（ER2/SF2）场景下，DW-CSL表现出突出的定向稳定性：对于10节点ER2网络，其AH-F_1达到0.68±0.10，较Shapley-PC提升5%，且V-F_1为0.88±0.05；在50节点SF2网络中，AH-F_1提升至0.92±0.05，较PC-Max提升36%，验证了算法在大规模稀疏图中的可靠性．针对高密度网络（ER4/SF4）的噪声干扰问题，DW-CSL通过归一化p值加权策略有效抑制误差传播，例如在50节点ER4场景下，其V-F_1为0.52±0.30，较Shapley-PC提升4%，标准差减少6%．从结构恢复精度来看，DW-CSL在复杂拓扑中表现尤为突出：50节点ER4网络的SHD均值为168.3，较Shapley-PC降低2.0%，且误差范围缩小0.6%；而在50节点SF4网络中，其SHD为24.5±10.7较基线方法PC降低10.1%，误差范围缩小2.3%．

5.2.2 基准数据集实验与分析

在基准数据集（Alarm、Insurance、Ecoli70、Mehra）上的实验结果表明，DW-CSL算法在因果结构恢复的精确

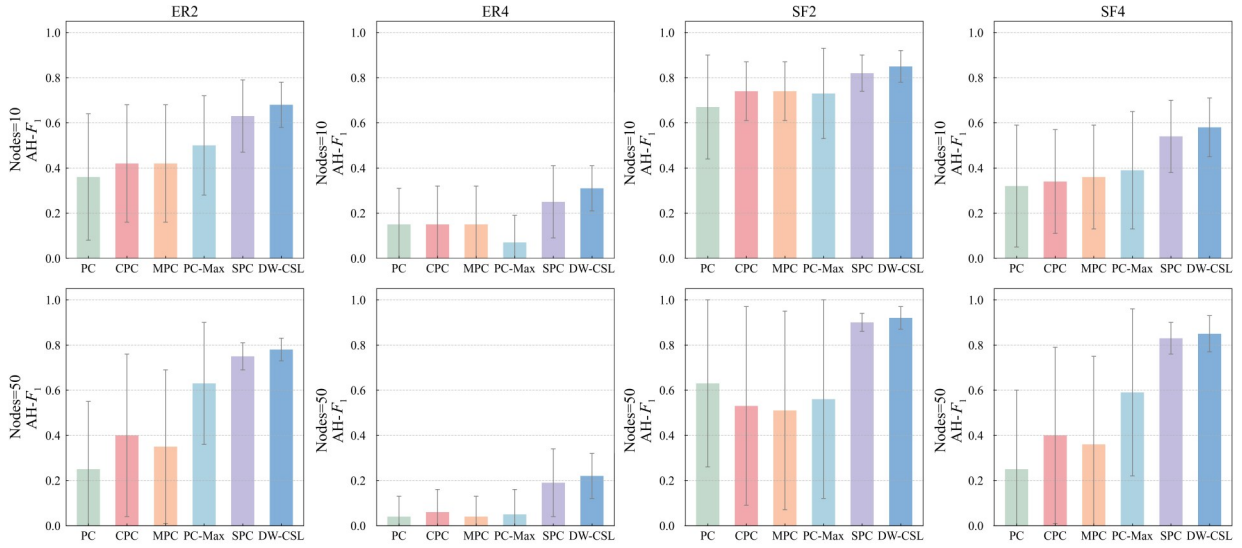

图2　10节点和50节点的ER和SF图的AH-F_1柱状图

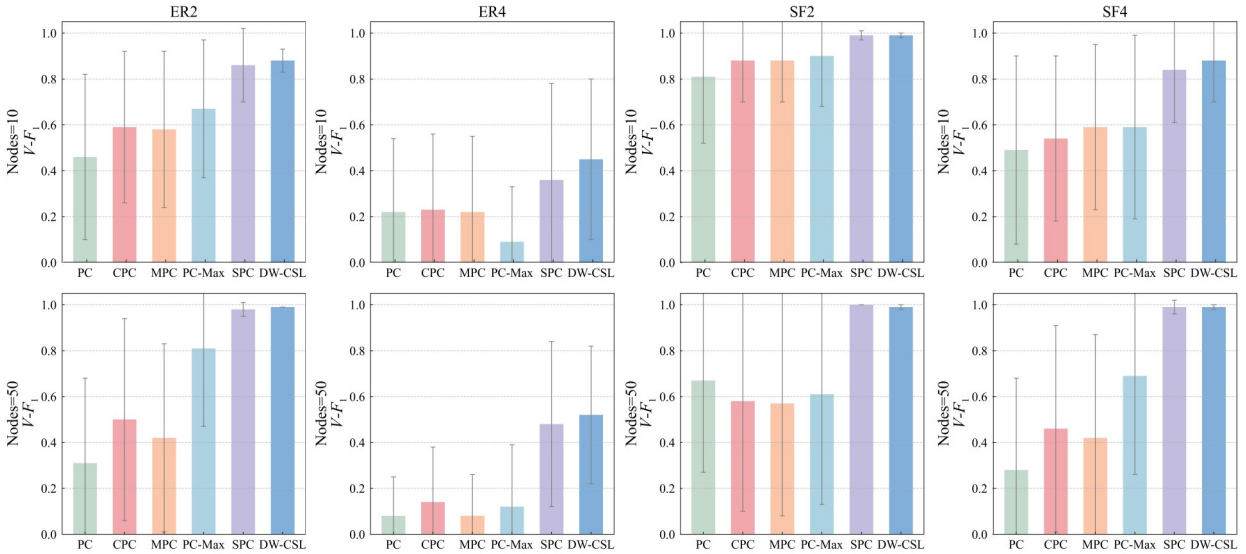

图3　10节点和50节点的ER和SF图的V-F_1柱状图

性与稳定性方面均显著优于现有方法. 从AH-F_1(箭头方向预测精度)和V-F_1(对撞结构识别精度)两项核心指标来看,DW-CSL算法在高复杂度网络与混合变量类型场景下展现出更强的适应性,如表3和图5所示. 在Alarm数据集上,其AH-F_1较PC算法提升47%;在Ecoli70数据集中,V-F_1为0.89±0.04与CPC(0.01±0.0)实现数量级突破. 此外,DW-CSL在跨数据集泛化性上表现突出:尽管Mehra网络具有最高的边密度(|E|=71),其AH-F_1仍达到0.51±0.18,显著高于PC-Max(0.15±0.2)和MPC(0.14±0.2),较Shapley-PC提升10%. 此外,DW-CSL方法在所有所选数据集下的SHD(Structural Hamming Distance)指标均优于其余基线方法,表现出更强的结构恢复能力,如图6所示.

实验结果表明,DW-CSL提出的动态权重分配机制发挥了关键作用. 该机制通过自适应地调整不同条件独立性检验(CIT)结果在全局结构学习中的置信权重,有效抑制了单次检验误差对整体因果图构建的负面影响. 这种对局部不确定性的鲁棒性处理,使模型在包含复杂依赖关系的真实场景数据上,展现出更可靠、更稳定的因果推理能力. 其提升的结构学习精度(SHD)与定向质量(V-F_1,AH-F_1)不仅体现在综合指标上,更在多个挑战性场景中得到验证. 因此,DW-CSL为需要高可信度因果关系的应用领域(如基于Alarm数据集的医疗监测系统故障原因分析、基于Mehra模型的工业过程控制系统关键变量间因果建模)提供了更具鲁棒性的因果发现技术支撑,有助于提升这些领域基于数据驱动的决策质量.

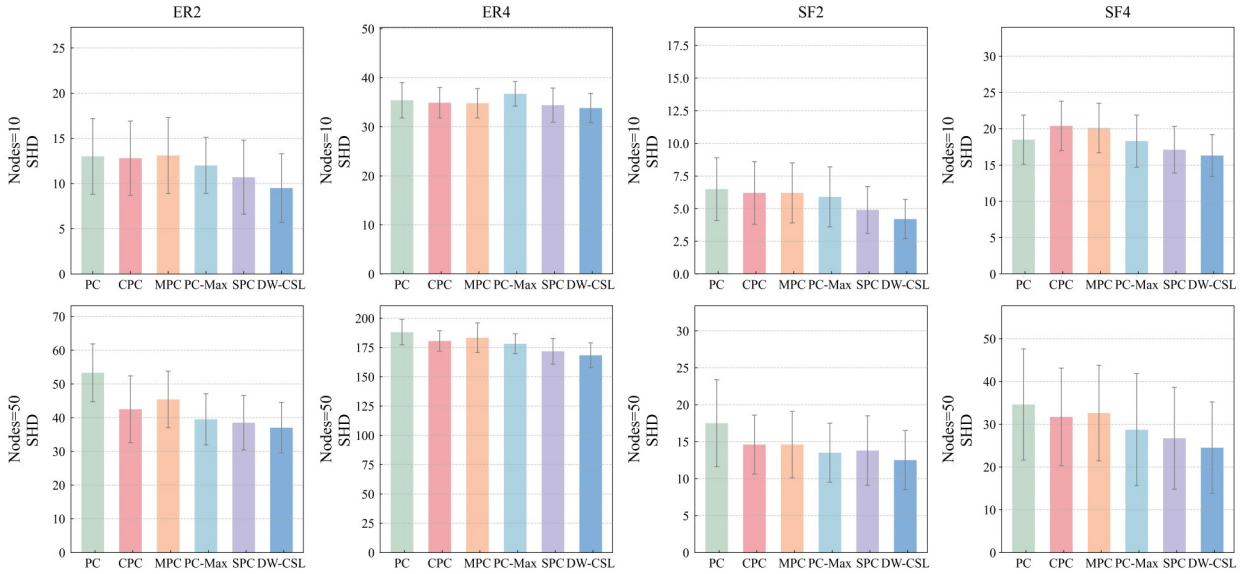

图4 10节点和50节点的ER和SF图的SHD柱状图

表3 贝叶斯网络生成数据集的AH-F_1和V-F_1

方法	Alarm		Insurance		Ecoli70		Mehra	
	AH-F_1	V-F_1	AH-F_1	V-F_1	AH-F_1	V-F_1	AH-F_1	V-F_1
PC	0.13±0.02	0.18±0.3	0.02±0.1	0.03±0.1	0.32±0.2	0.29±0.2	0.01±0.0	0.01±0.0
CPC	0.37±0.03	0.53±0.4	0.04±0.1	0.05±0.1	0.01±0.0	0.01±0.0	0.27±0.2	0.45±0.4
MPC	0.16±0.03	0.24±0.4	0.02±0.1	0.02±0.1	0.55±0.1	0.60±0.1	0.14±0.2	0.26±0.4
PC-Max	0.16±0.03	0.2±0.4	0.04±0.1	0.07±0.2	0.60±0.2	0.72±0.3	0.15±0.2	0.28±0.5
SPC	0.57±0.00	0.85±0.1	0.21±0.1	0.42±0.2	0.73±0.1	0.89±0.1	0.41±0.2	0.76±0.3
DW-CSL	**0.60±0.15**	**0.89±0.08**	**0.32±0.03**	**0.58±0.11**	**0.75±0.1**	**0.89±0.2**	**0.5±0.14**	**0.79±0.1**

注：加粗数据表示最优结果.

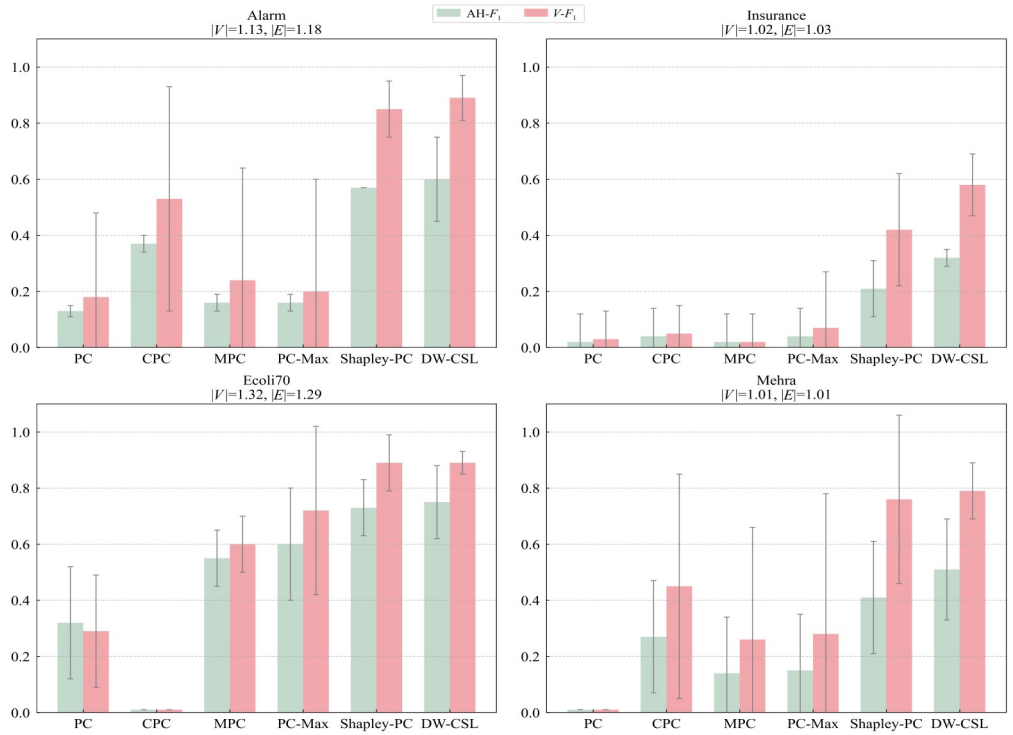

图5 贝叶斯网络生成数据集的AH-F_1和V-F_1柱状图

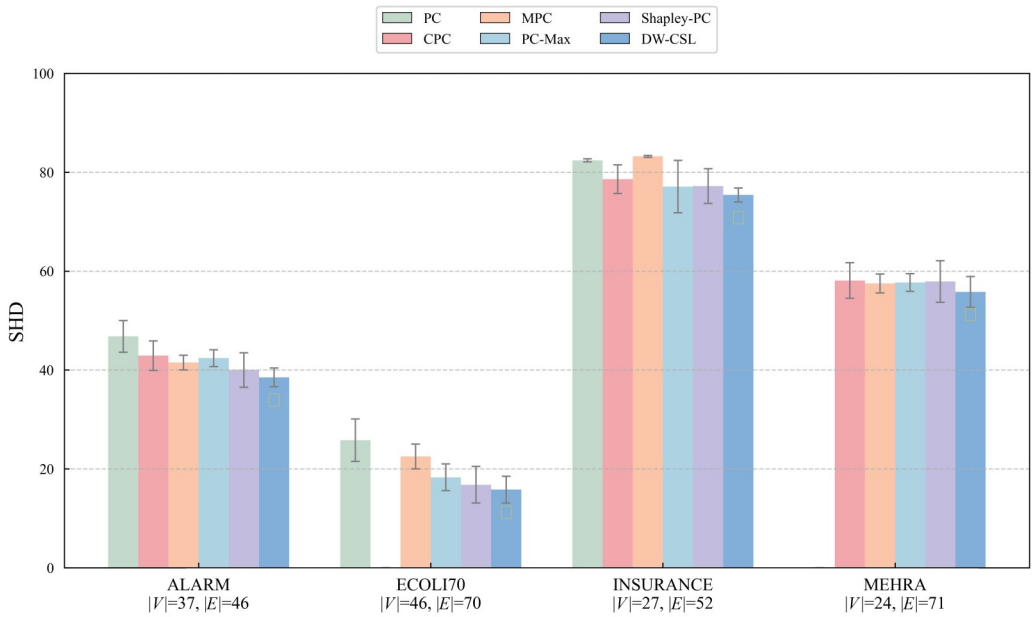

图6　贝叶斯网络生成数据集的SHD柱状图

5.2.3　异常CIT敏感性实验分析

在异常CIT敏感性实验上表明,DW-CSL算法的动态加权机制可以有效抑制异常CIT结果,如图7和图8所示. 随着噪声强度增大,所有方法的性能曲线均呈整体下降的趋势,但下降幅度存在显著差异. DW-CSL的两条曲线始终位于最上方且斜率最平缓. 相比之下,PC、CPC、MPC等基线方法曲线陡峭,部分指标在5%噪声时已接近甚至低于0.01. Shapley-PC与PC-Max介于两者之间,但在高噪声区间仍出现明显衰减. 综合折线图可见,DW-CSL的鲁棒性较高,显著优于对比算法,进一步验证了动态加权机制对异常条件独立性检验结果的有效抑制.

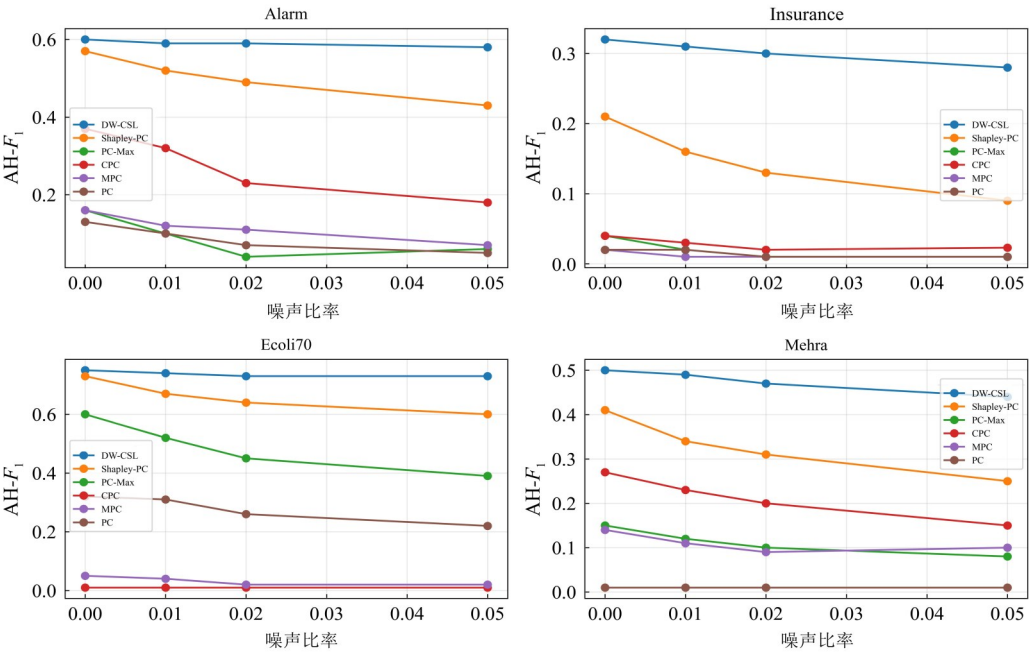

图7　异常CIT敏感性实验的AH-F_1折线图

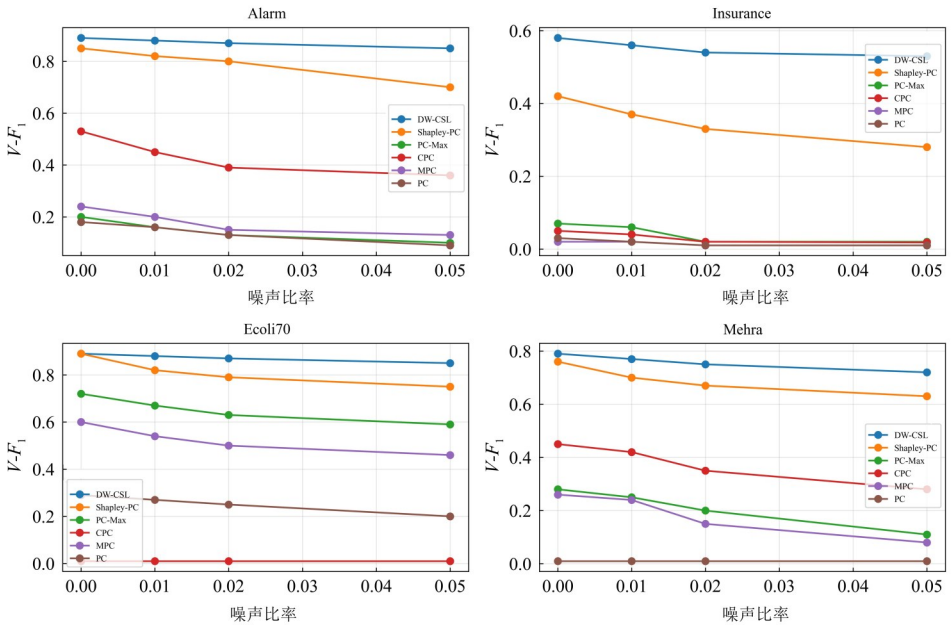

图8 异常CIT敏感性实验的V-F_1折线图

6 结论

本文针对传统基于约束的因果结构学习方法在有限样本条件下对条件独立性检验结果依赖性强、易引发对撞结构误判的问题，提出了一种条件集动态加权的因果结构学习算法(DW-CSL). 该方法基于Shapley值提出条件集动态加权的定向决策规则，并结合归一化权重提升定向判断的精度与鲁棒性. 在理论上本方法具备一致性与正确性. 通过在合成与基准数据集上的对比实验，DW-CSL在结构恢复精度(SHD)与定向质量(V-F_1、AH-F_1)方面均优于现有主流算法，尤其在高维、低样本或噪声扰动较强的场景下仍能保持较高的性能，具有良好的稳定性与扩展性. 该方法的局限性主要在于计算效率，当处理邻接节点较多的复杂结构时，其所依赖的条件独立性测试的计算复杂度显著增加，导致整体学习的时间开销较大. 此外，本文方法建立在因果充分性与线性因果关系模型假设之上，在非线性因果关系模型下，本方法使用的条件独立性检验刻画变量间依赖强度的准确性下降，致使归一化p-value的统计意义减弱，从而影响Shapley值权重的精确性；本方法因果关系模型的建立过程不适用于存在隐变量的情况，但本方法的思路可以推广到存在隐变量的场景中. 因此，未来工作将聚焦于开发更高效的策略，同时进一步探索该方法在非线性模型和存在隐变量场景中的适应性与推广性.

参考文献

[1] SPIRTES P, ZHANG K. Causal discovery and inference: Concepts and recent methodological advances[J]. Applied Informatics, 2016, 3: 3.

[2] PEARL J, MACKENZIE D. The book of why: The new science of cause and effect[J]. Journal of the American Statistical Association , 2020, 115(529): 482-485.

[3] CAO F Y, WANG Y X, YU K, et al. Causal discovery from unknown interventional datasets over overlapping variable sets[J]. IEEE Transactions on Knowledge and Data Engineering, 2024, 36(12): 7725-7742.

[4] NEUBERG L G. Causality: Models, reasoning, and inference[J]. Econometric Theory, 2003, 19(4): 675-685.

[5] CAI R C, ZHANG Z J, HAO Z F, et al. Understanding social causalities behind human action sequences[J]. IEEE Transactions on Neural Networks and Learning Systems, 2017, 28(8): 1801-1813.

[6] RUNGE J, NOWACK P, KRETSCHMER M, et al. Detecting and quantifying causal associations in large nonlinear time series datasets[J]. Science Advances, 2019, 5(11): eaau4996.

[7] CAI R C, ZHANG Z J, HAO Z F. Causal gene identification using combinatorial V-structure search[J]. Neural Networks, 2013, 43: 63-71.

[8] YANG J, AN N, ALTEROVITZ G. A partial correlation statistic structure learning algorithm under linear structural equation models[J]. IEEE Transactions on Knowledge and Data Engineering, 2016, 28(10): 2552-2565.

[9] SHEN X P, MA S S, VEMURI P, et al. Challenges and opportunities with causal discovery algorithms: Application to Alzheime's pathophysiology[J]. Scientific Reports, 2020, 10: 2975.

[10] CHEN W Q, HAO Z F, CAI R C, et al. Multiple-cause discovery combined with structure learning for high-dimensional discrete data and application to stock prediction[J]. Soft Computing, 2016, 20(11): 4575-4588.

[11] LI Y Z, TORRALBA A, ANANDKUMAR A, et al. Causal discovery in physical systems from videos[EB/OL]. (2020-11-29)[2025-07-21]. https://arxiv.org/abs/2007.00631.

[12] YANG J, LI N, AN N, et al. An efficient causal structure learning algorithm for linear arbitrarily distributed continuous data[J]. The Journal of Supercomputing, 2020, 76(5): 3355-3363.

[13] BELTHANGADY C, GIAMPANIS S, JANKOVIC I, et al. Causal deep learning reveals the comparative effectiveness of antihyperglycemic treatments in poorly controlled diabetes[J]. Nature Communications, 2022, 13: 6921.

[14] SCHÖLKOPF B, LOCATELLO F, BAUER S, et al. Toward causal representation learning[J]. Proceedings of the IEEE, 2021, 109(5): 612-634.

[15] WANG Y, CAO F Y, YU K, et al. Federated causal structure learning with non-identical variable sets[C]//Proceedings of the 42nd International Conference on Machine Learning. Cambridge: PMLR, 2025: 302-324.

[16] HUEGLE J, HAGEDORN C, SCHLOSSER R. A KNN-based non-parametric conditional independence test for mixed data and application in causal discovery[C]//Machine Learning and Knowledge Discovery in Databases: Research Track. Cham: Springer, 2023: 541-558.

[17] ZHANG K, PETERS J, JANZING D, et al. Kernel-based conditional independence test and application in causal discovery[C]//Proceedings of the 27th Conference on Uncertainty in Artificial Intelligence. New York: ACM, 2011: 804-813.

[18] YANG S J, CAO F Y, YU K, et al. Learning causal chain graph structure via alternate learning and double pruning[J]. IEEE Transactions on Big Data, 2024, 10(4): 442-456.

[19] SPIRTES P, GLYMOUR C N, SCHEINES R, et al. Causation, Prediction, and Search [M]. 2nd Ed. Cambridge, Mass: MIT Press, 2000.

[20] MURPHY K, SCHÖLKOPF B, COLOMBO D, et al. Order-independent constraint-based causal structure learning[J]. Journal of Machine Learning Research, 2014, 15(1): 3741-3782.

[21] RAMSEY J, SPIRTES P, ZHANG J J. Adjacency-faithfulness and conservative causal inference[C]//Proceedings of the 22th Conference on Uncertainty in Artificial Intelligence. New York: ACM, 2006: 401-408.

[22] 蔡瑞初, 陈薇, 张坤, 等. 基于非时序观察数据的因果关系发现综述[J]. 计算机学报, 2017, 40(6): 1470-1490.
CAI R C, CHEN W, ZHANG K, et al. A survey on non-temporal series observational data based causal discovery[J]. Chinese Journal of Computers, 2017, 40(6): 1470-1490. (in Chinese)

[23] 郝志峰, 汪菲霞, 陈正鸣, 等. 基于增强条件独立性检验的鲁棒因果发现算法[J]. 软件学报, 2025, 36(9): 4134-4152.
HAO Z F, WANG F X, CHEN Z M, et al. Robust causal discovery algorithm based on enhanced conditional independence tests[J]. Journal of Software, 2025, 36(9): 4134-4152. (in Chinese)

[24] MARGARITIS D, THRUN S. Bayesian network induction via local neighborhoods[C]//Proceedings of the 13th International Conference on Neural Information Processing Systems. New York: ACM, 1999: 505-511.

[25] ZHALAMA Z, ZHANG J J, EBERHARDT F, et al. ASP-based discovery of semi-Markovian causal models under weaker assumptions[C]//Proceedings of the 28th International Joint Conference on Artificial Intelligence. California: IJCA, 2019: 1488-1494.

[26] RUSSO F, TONI F. Shapley-PC: Constraint-based causal structure learning with Shapley values[C]//Proceedings of the 4th Conference on Causal Learning and Reasoning. Lausanne, Switzerland: PMLR, 2025: 292-339.

[27] AUMANN R J, HART S. Handbook of Game Theory with Economic Applications[M]. Amsterdam, Tokyo: Elsevier, 1992.

[28] SUNDARARAJAN M, NAJMI A. The many Shapley values for model explanation[C]//Proceedings of the 37th International Conference on Machine Learning. New York: ACM, 2020: 9269-9278.

[29] ROZEMBERCZKI B, WATSON L, BAYER P, et al. The shapley value in machine learning[C]//Proceedings of the 31th International Joint Conference on Artificial Intelligence. Vienna: IJCAI, 2022: 5572-5579.

[30] 曹付元, 杨淑晶, 王雲霞, 等. 基于约束的局部-全局LWF链图结构学习算法[J]. 电子学报, 2023, 51(6): 1458-1467.
CAO F Y, YANG S J, WANG Y X, et al. Local-global LWF chain graph structure learning algorithm based on constraints[J]. Acta Electronica Sinica, 2023, 51(6): 1458-1467. (in Chinese)

[31] VOWELS M J, CAMGOZ N C, BOWDEN R. D'ya like DAGs? A survey on structure learning and causal discovery[J]. ACM Computing Surveys, 2023, 55(4): 1-36.

[32] ZANGA A, OZKIRIMLI E, STELLA F. A survey on

causal discovery: Theory and practice[J]. International Journal of Approximate Reasoning, 2022, 151: 101-129.
[33] GLYMOUR C, ZHANG K, SPIRTES P. Review of causal discovery methods based on graphical models[J]. Frontiers in Genetics, 2019, 10: 524.
[34] ALIFERIS C F, STATNIKOV A, TSAMARDINOS I, et al. Local causal and Markov blanket induction for causal discovery and feature selection for classification part I: Algorithms and empirical evaluation[J]. Journal of Machine Learning Research, 2010, 11: 171-234.
[35] PETERS J, JANZING D, SCHÖLKOPF B. Elements of Causal Inference: Foundations and Learning Algorithms[M]. Cambridge: MIT Press, 2017.
[36] PEARL J. The foundations of causal inference[J]. Sociological Methodology, 2010, 40(1): 75-149.
[37] GONG C, ZHANG C Z, YAO D, et al. Causal discovery from temporal data: An overview and new perspectives[J]. ACM Computing Surveys, 2024, 57(4): 1-38.

作者简介

曹冬蕾 女,2001年1月出生于山西省大同市. 山西大学计算机与信息技术学院硕士研究生. 主要研究方向为机器学习与因果推断.
E-mail: caodonglei@sxu.edu.cn

王雲霞 女,1996年6月出生于山西省朔州市. 山西大学计算机与信息技术学院讲师. 主要研究方向为机器学习与因果推断.
E-mail: wyx@sxu.edu.cn

曹付元 男,1974年5月出生于山西省大同市. 山西大学计算机与信息技术学院教授、博士生导师. 主要研究方向为机器学习与因果推断.
E-mail: cfy@sxu.edu.cn

高小方 女,1978年2月出生于山西省安泽县. 山西大学计算机与信息技术学院副教授、硕士生导师. 主要研究方向为数据挖掘与机器学习.
E-mail: gxfhtp@sxu.edu.cn