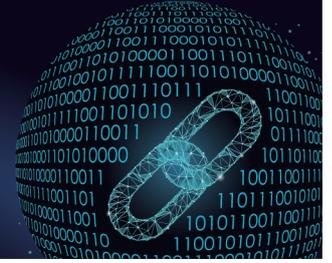




# 控制与决策

CONTROL AND DECISION



## 一种基于互信息度量的时序数据因果发现方法

李德志, 鲁云军, 吴健平, 李强

引用本文:

李德志, 鲁云军, 吴健平, 李强. 一种基于互信息度量的时序数据因果发现方法[J]. 控制与决策, 2024, 39(9): 3151–3159.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2023.0248>

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 一种基于柯氏复杂度的因果网络定向方法

A causal network orientation method based on Kolmogorov complexity

控制与决策. 2021, 36(9): 2241–2248 <https://doi.org/10.13195/j.kzyjc.2020.0005>

#### 基于MI-SVR模型的航空旅客出行指数预测方法研究

Air passenger index prediction method based on MI-SVR mode

控制与决策. 2021, 36(7): 1619–1626 <https://doi.org/10.13195/j.kzyjc.2019.1446>

#### 基于互信息操作变量曲线参数化的间歇过程批内修正优化

Intra-batch correction optimization of batch process with manipulated variable trajectory parameterization based on mutual information

控制与决策. 2021, 36(1): 234–240 <https://doi.org/10.13195/j.kzyjc.2019.0825>

#### 时滞累积TDAGM( $\beta, N, t$ )模型及其在粮食生产中的应用

Time-delayed accumulative TDAGM( $\beta, N, t$ )

控制与决策. 2021, 36(8): 2002–2012 <https://doi.org/10.13195/j.kzyjc.2019.1630>

#### 基于波段影像统计信息量加权K-means聚类的高光谱影像分类

Algorithm based on band statistical information weighted K-means for hyperspectral image classification

控制与决策. 2021, 36(5): 1119–1126 <https://doi.org/10.13195/j.kzyjc.2019.1516>

# 一种基于互信息度量的时序数据因果发现方法

李德志, 鲁云军<sup>†</sup>, 吴健平, 李 强

(国防科技大学 信息通信学院, 武汉 430019)

**摘要:** 在时序数据因果关系发现研究中,传统算法针对时间窗口内时序数据之间的因果关系进行分析,存在因果关系识别准确率受限、算法复杂度较高等问题.为解决该问题,首先对概要因果图、因果概要互信息和条件因果概要互信息进行定义,在此基础上推导出基于因果互信息的时序变量定向规则,而后区分是否存在混杂因子,结合PC(Peter and Clark)和FCI(fast causal inference)算法分别提出改进的PCSMI(Peter and Clark summary mutual information)和FCISMI(fast causal inference summary mutual information)算法.实验结果表明改进后算法能够在低复杂度条件下有效提升时序数据因果发现的准确率.

**关键词:** 因果发现; 时序数据; 概要因果图; 因果概要互信息; 条件因果概要互信息; PC; FCI

中图分类号: TP181

文献标志码: A

DOI: 10.13195/j.kzyjc.2023.0248

**引用格式:** 李德志,鲁云军,吴健平,等.一种基于互信息度量的时序数据因果发现方法[J].控制与决策,2024,39(9): 3151-3159.

## A causal discovery method for time series data based on mutual information measurement

LI De-zhi, LU Yun-jun<sup>†</sup>, WU Jian-ping, LI Qiang

(College of Information and Communication, National University of Defense Technology, Wuhan 430019, China)

**Abstract:** In the research on causal discovery of time series data, the traditional algorithm analyse the causal relationship between time series data in the time window, which has problems such as limited causality recognition accuracy and high algorithm complexity. In order to solve this problem, this paper first defines the summary causal diagram, causal summary mutual information and conditional causal summary mutual information, derives the temporal series variable orientation rule based on causal mutual information, and then distinguishes whether there are confounding factors, and proposes improved Peter and Clark summary mutual information (PCSMI) and fast causal inference summary mutual information (FCISMI) algorithms combined with Peter and Clark (PC) and fast causal inference (FCI) algorithms, respectively. Experimental results show that the improved algorithm can effectively improve the accuracy of causal discovery of time series data under low complexity conditions.

**Keywords:** causal discovery; time series data; summary causal diagram; causal summary mutual information; conditional causal summary mutual information; PC; FCI

## 0 引言

时序数据是传感器或实验在一段时间内产生的观测结果的集合,在研究中通常也指代时序变量的集合.从时序数据中进行因果发现能够揭示数据背后的机制,推理干预后果,对于医疗、工业和社会学等领域研究具有重要意义<sup>[1]</sup>.

因果发现常用的算法<sup>[2-3]</sup>主要分为以下几类:格兰杰系列算法<sup>[4]</sup>、基于评分类算法、基于函数类算法、基于约束类算法<sup>[5-7]</sup>和基于深度学习类算法<sup>[8-11]</sup>.格兰杰系列算法具有代表性的是MVGC (multivariate

granger causality)算法,基本思路是原因变量有助于解释结果变量未来变化,主要适用于无混杂因子的场景<sup>[12-14]</sup>.基于评分类具有代表性的是DYNOTEARS (dynamic non combinatorial optimization via trace exponential and augmented lagRangian for structure learning)算法<sup>[15]</sup>,基本思路是通过最大化图评分找到最符合实际数据的因果关系<sup>[16]</sup>,主要适用于变量具有高维特性的场景.基于函数类具有代表性的是TiMINo (time series models with independent noise)算法,基本思路是通过识别噪声和潜在原因之间的独

收稿日期: 2023-03-06; 录用日期: 2023-09-07.

基金项目: “十四五”装备预研项目(315057206).

<sup>†</sup>通讯作者. E-mail: lu\_yunjun@hotmail.com.

立性来发现因果关系<sup>[17-18]</sup>,主要适用于存在混杂因子的场景. 基于约束类是因果发现中最流行的算法,基本思路是通过检验变量之间的独立性或条件独立性识别其因果关系<sup>[19]</sup>,最具代表性的是PC(Peter and Clark)和FCI(fast causal inference)算法<sup>[20-21]</sup>. 近年来也有许多学者对其加以改进完善,例如文献[22]提出的PCMCI(Peter and Clark momentary conditional independence)算法将因果发现区分为两个阶段,第1阶段按照经典PC算法思路迭代独立性测试删除无关变量,第2阶段利用第1阶段结果执行MCI测试;文献[23]提出的tsFCI(time series fast causal inference)算法利用先验知识对经典FCI算法进行改进,通过对历史时间节点集的约束,重新定位不同时间点各变量之间的因果方向. 基于深度学习类具有代表性的是CGTST(causality-gated time series transformer)算法<sup>[24]</sup>,基本思路是通过因果门和时序Transformer组合来预测因果关系,主要适用于大规模或超大规模数据因果发现场景. 上述这些因果发现算法能够发现时序数据中的因果关系,但由于对变量自因果特性带来的影响考虑不足,使其因果关系识别准确率受限. 而每个时间窗口都需要进行独立性检验以及对概率密度函数的估计,使得算法复杂度较高.

在以上研究成果的基础上,针对上述算法不足,本文按照基于约束类算法的基本思路,通过构建概要因果图将变量自因果特性纳入互信息计算范畴并尽可能减少检验次数,在此基础上对因果概要互信息和条件因果概要互信息进行估计和检验,区分是否存在混杂因子,而后对经典PC和FCI算法进行改进,提出PCSMI(Peter and Clark summary mutual information)和FCISMI(fast causal inference summary mutual information)算法,力求以较低的算法复杂度实现较高的因果发现准确率.

### 1 理论基础

本文研究主要基于以下基本假设:一是所有时间序列数据在时间上对齐;二是所有时间序列采样时刻为整数;三是所有时间序列满足一阶马尔可夫自因果性. 在满足以上假设前提下,本文提出的方法主要涉及以下基本概念.

#### 1.1 概要因果图

概要因果图是指抽象出不同时间序列变量之间的因果关系,忽略其不同采样时刻的具体因果关系而形成的因果图. 概要因果图可以理解为窗口因果图的概要表示形式,窗口因果图主要是从完整时间序列图中截取有限长时间序列进行分析,它反映的是各变

量时间序列之间在各时刻全部因果关系. 概要因果图反映的是变量之间的总体因果关系(包括变量因时间流逝产生的自因果关系),其基本定义如下.

**定义1** 设 $X$ 是一个多变量离散时间序列, $G = (V, E)$ 为概要因果图. 该图中的顶点集合 $V$ 由集合 $X^1, \dots, X^d$ 组成. 图的边 $E$ 定义如下:当且仅当在时间 $t$ 时,时延 $0 \leq i(p \neq q)$ 或时延 $0 < i(p = q)$ 条件下,若变量 $X_{t-i}^p$ 是变量 $X_t^q$ 的原因,则变量 $X^p$ 和变量 $X^q$ 通过箭头连接,变量自身因果关系通过自环箭头表示.

概要因果图是有向无环图的一种形式,如图1所示. 本文主要通过构建概要因果图来识别变量之间的因果关系,以便于缩减计算规模.

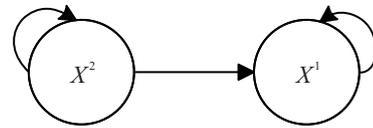


图1 概要因果图

#### 1.2 因果概要互信息

因果概要互信息是以两个时间序列变量历史数据为条件,在一定时间窗口内,所有因果关系产生标准互信息的最大值. 要确定因果概要互信息,需要先划定最优时间窗口以确保完全包含变量之间因果关系.

划定时间窗口后,将因果概要互信息表示中涉及的变量定义如下: $\gamma_{\max}$ 表示两个时间序列变量 $X^p$ 与 $X^q$ 之间最大时延,最大的窗口大小可表示为 $\lambda_{\max} = \gamma_{\max} + 1$ . 时间序列变量 $X^p$ 关于 $X^q$ 的时间窗口集合表示为 $X^{(p;\lambda_{pq})}$ ,时间序列变量 $X^p$ 关于 $X^q$ 的时间窗口集合表示为 $X^{(q;\lambda_{qp})}$ . 在此基础上给出因果概要互信息定义如下.

**定义2** 对于时间序列变量 $X^p$ 和 $X^q$ ,本文将 $X^p$ 与 $X^q$ 之间因果概要互信息(CSMI)定义为

$$CSMI(X^p; X^q) = \max_{(\lambda_{pq}, \lambda_{qp}, \gamma_{pq}) \in C^{(p,q)}} I(X_t^{(p;\lambda_{pq})}; X_{t+\gamma_{pq}}^{(q;\lambda_{qp})} | X_{t-1}^{(p;1)}, X_{t+\gamma_{pq}-1}^{(q;1)}) \triangleq I(X_t^{(p;\bar{\lambda}_{pq})}; X_{t+\bar{\gamma}_{qp}}^{(q;\bar{\lambda}_{qp})} | X_{t-1}^{(p;1)}, X_{t+\bar{\gamma}_{qp}-1}^{(q;1)}), \quad (1)$$

其中 $I$ 表示互信息. 设 $C^{(p,q)}$ 表示时间窗口大小和相应时延的集合,由于 $C^{(p,q)}$ 中不同值均可能取得最大值,本文设 $\bar{\gamma}_{pq}$ 为最大值, $\bar{\lambda}_{pq}$ 和 $\bar{\lambda}_{qp}$ 为最小值. $\bar{\gamma}_{pq}$ , $\bar{\lambda}_{pq}$ 和 $\bar{\lambda}_{qp}$ 与其对应的最优时延和最优窗口大小相一致. 将 $\bar{\gamma}_{pq}$ 设为最大值是因为可以忽略不重要时刻

的因果关系,  $\bar{\lambda}_{pq}$  和  $\bar{\lambda}_{qp}$  设为最小值是因为增加窗口大小只会增加互信息,取最小值为最优时间窗口,因此有

$$\begin{aligned}
 & I(X_t^{(p;\lambda_{pq})}; X_{t+\gamma_{pq}}^{(q;\lambda_{qp})} | X_{t-1}^{(p;1)}, X_{t+\gamma_{pq}-1}^{(q;1)}) = \\
 & I((X_t^{(p;\lambda_{pq}-1)}, X_{t+\lambda_{qp}-1}^{(p;1)}); X_{t+\gamma_{pq}}^{(q;\lambda_{qp})} | X_{t-1}^{(p;1)}, X_{t+\gamma_{pq}-1}^{(q;1)}) = \\
 & I(X_t^{(p;\lambda_{pq}-1)}; X_{t+\gamma_{pq}}^{(q;\lambda_{qp})} | X_{t-1}^{(p;1)}, X_{t+\gamma_{pq}-1}^{(q;1)}) + \\
 & I(X_{t+\lambda_{qp}-1}^{(p;1)}; X_{t+\gamma_{pq}}^{(q;\lambda_{qp})} | X_{t-1}^{(p;1)}, X_{t+\gamma_{pq}-1}^{(q;1)}, X_t^{(p;\lambda_{pq}-1)}) \geq \\
 & I(X_t^{(p;\lambda_{pq}-1)}; X_{t+\gamma_{pq}}^{(q;\lambda_{qp})} | X_{t-1}^{(p;1)}, X_{t+\gamma_{pq}-1}^{(q;1)}). \quad (2)
 \end{aligned}$$

本文目标是构建概要因果图来进行因果发现,因此不需要考虑两个时间序列变量之间所有的因果关系,通过CSMI能够有效识别两个时序变量之间的因果关系,确保不会遗漏.此外,本文借鉴文献[25]转移熵的思路,将时间窗口大小和时延基本单位设置为1个采样时间间隔,并通过变量历史数据为条件的方式消除自因果关系.

CSMI的另一个重要性质定义如下.

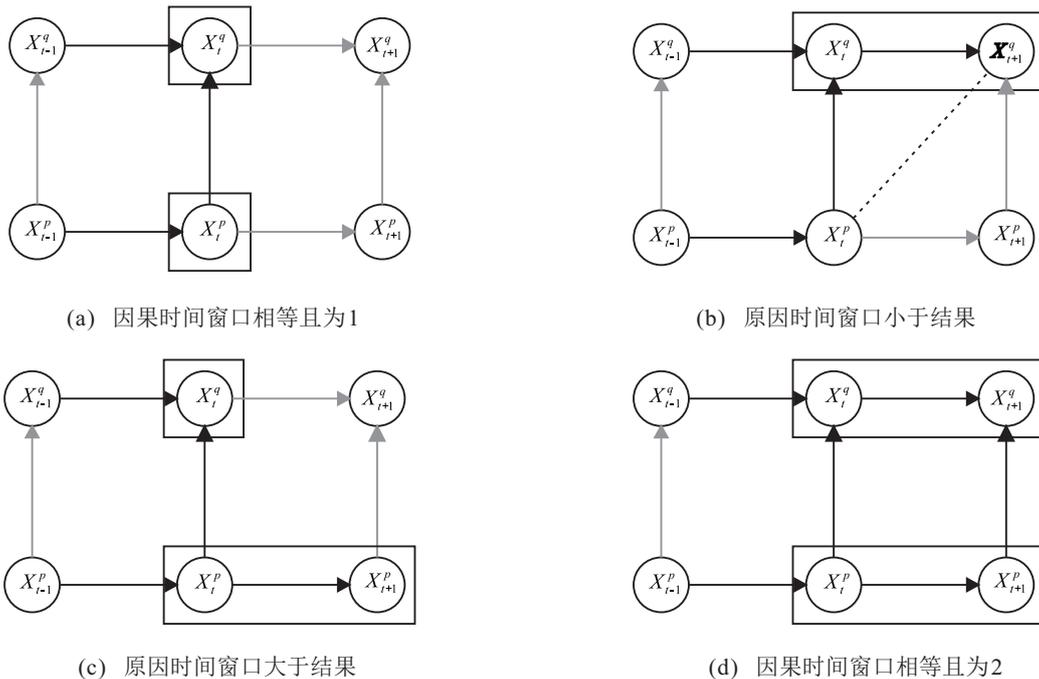


图2 CSMI值非对称性增长示意图

### 1.3 条件因果概要互信息

将因果概要互信息扩展到条件集上进行应用,可得到条件因果概要互信息.在时间序列变量因果发现过程中,常通过观察两个关联变量是否条件独立来确定因果关系.即以两个存在关联关系的时间序列变量的共因(其他时间序列变量)为条件,会导致两个变量之间条件独立,由此可确定几个变量之间的因果关系.因此,条件因果概要互信息的基本定义如下.

**定义3** 给定两个时间序列变量  $X^p$  和  $X^q$ ,当  $\gamma_{pq} = 0$  时,若变量  $X^p$  是变量  $X^q$  的原因,则CSMI的值不会随着  $\lambda_{pq}$  和  $\lambda_{qp}$  的增长而对称性增长.

定义3的示例如图2所示,其中虚线表示变量之间的相关关系(非因果关系),加粗箭头表示变量之间的因果关系.图2(a)中,由于  $\gamma_{pq} = 0$  时变量  $X^p$  是  $X^q$  的原因,CSMI为正值,即  $I(X_t^p, X_t^q | X_{t-1}^p, X_{t-1}^q) > 0$ ;图2(b)中,扩展变量  $X^q$ (结果)的时间窗口,CSMI的值相比图2(a)的值会增加,即  $I(X_t^p, X_t^{(q;2)} | X_{t-1}^p, X_{t-1}^q) > I(X_t^p, X_t^q | X_{t-1}^p, X_{t-1}^q)$ ;图2(c)中,仅扩展变量  $X^p$ (原因)的时间窗口,因此CSMI的值同图2(a)的值保持一致,即  $I(X_t^{(p;2)}, X_t^q | X_{t-1}^p, X_{t-1}^q) = I(X_t^p, X_t^q | X_{t-1}^p, X_{t-1}^q)$ ,时延  $\gamma_{pq} \in \mathbb{Z}$ ,  $\lambda_{pq}$  和  $\lambda_{qp}$  在时间窗口集合  $X^{(p;\lambda_{pq})}$  和  $X^{(q;\lambda_{qp})}$  中相关的时间窗口起始时间点通过  $\gamma_{pq}$  分隔;图2(d)中,同时扩展变量  $X^p$  和  $X^q$  的时间窗口,CSMI的值相比图2(a)的值也会增加,即  $I(X_t^{(p;2)}, X_t^{(q;2)} | X_{t-1}^p, X_{t-1}^q) > I(X_t^p, X_t^q | X_{t-1}^p, X_{t-1}^q)$ .

**定义4** 当  $\bar{\gamma}_{pq} \geq 0$  时,时间序列变量  $X^p$  和  $X^q$  在以集合  $X^R = \{X^{r_1}, \dots, X^{r_k}\}$  为条件情况下的条件因果概要互信息(CCSMI)表示为

$$\begin{aligned}
 & \text{CCSMI}(X^p; X^q | X^R) = \\
 & I(X_t^{(p;\bar{\lambda}_{pq})}; X_{t+\bar{\gamma}_{pq}}^{(q;\bar{\lambda}_{qp})} | (X_{t-\bar{\Gamma}_k}^{(r_k;\bar{\lambda}_k)}), X_{t-1}^{(p;1)}, X_{t+\bar{\gamma}_{pq}-1}^{(q;1)}). \quad (3)
 \end{aligned}$$

$\bar{\Gamma}_k$  和  $\bar{\lambda}_k$  分别为最优时延和最佳时间窗口大小.由

于  $\Gamma_k$  有多种可能值, 可将  $\bar{\Gamma}_k$  设为最大值, 将  $\bar{\lambda}_k$  设为最小值. 通过 CCSMI 可以确保, 若存在条件变量集合使得两个时间序列变量独立, 则这个条件变量集合可以被识别出来.

#### 1.4 估计和检验

时间序列数据因果发现要解决两个关键问题: 一是需要从有限长时间序列变量数据中对因果概要互信息进行估计; 二是需要有效地统计检验方法来应对样本数据较少, 总体分布未知的情况. 本文借鉴文献[26]中  $K$  近邻方法对因果概要互信息进行估计. 将  $K$  近邻方法中距离定义为两个时间窗口之间的最大距离(即两个时间窗口差值绝对值的最大值), 有

$$d((X_t^{(p; \bar{\lambda}_{pq})}, X_{t+\bar{\gamma}_{pq}}^{(q; \bar{\lambda}_{pq})}), (X_t^{(p; \bar{\lambda}_{pq})}, X_{t+\bar{\gamma}_{pq}}^{(q; \bar{\lambda}_{pq})})) = \max_{0 \leq l_p \leq \lambda_p, 0 \leq l_q \leq \lambda_q} (|(X_t^{(p; \bar{\lambda}_{pq})})_{i+l_p} - (X_t^{(p; \bar{\lambda}_{pq})})_{j+l_q}|, |(X_t^{(q; \bar{\lambda}_{qp})})_{i+l_q} - (X_t^{(q; \bar{\lambda}_{qp})})_{j+l_q}|). \quad (4)$$

因此, 对 CSMI 的估计如下:

$$\text{CSMI}(X^p; X^q) = \psi(k) + \frac{1}{n_{\gamma_{pq}, \gamma_{qp}}} \sum_{i=1}^{n_{\gamma_{pq}, \gamma_{qp}}} \psi(n_i^3) - \psi(n_i^{1,3}) - \psi(n_i^{2,3}). \quad (5)$$

其中:  $\psi$  为双伽马方程;  $n_{\gamma_{pq}, \gamma_{qp}}$  为采样数;  $n_i^{1,3}, n_i^{2,3}, n_i^3$  表示以下3个子空间

$$\begin{aligned} & (X_t^{(p; \bar{\lambda}_{pq})}, X_{t-1}^p, X_{t+\bar{\gamma}_{pq}-1}^q, (X_{t-\bar{\Gamma}_k}^{(r_k; \bar{\lambda}_k)})_{1 \leq k \leq K}), \\ & (X_{t+\bar{\gamma}_{pq}}^{(p; \bar{\lambda}_{pq})}, X_{t-1}^p, X_{t+\bar{\gamma}_{pq}-1}^q, (X_{t-\bar{\Gamma}_k}^{(r_k; \bar{\lambda}_k)})_{1 \leq k \leq K}), \\ & (X_{t-1}^p, X_{t+\bar{\gamma}_{pq}-1}^q, (X_{t-\bar{\Gamma}_k}^{(r_k; \bar{\lambda}_k)})_{1 \leq k \leq K}) \end{aligned}$$

中距离严格小于  $\epsilon_{ik}/2$  点的数量,  $k$  表示采样长度,  $\epsilon_{ik}$  表示空间中各点组成多维球体的半径. 由式(5)可推导出 CCSMI 的估计如下:

$$\text{CCSMI}(X^p; X^q | X^R) = \psi(k) + \frac{1}{n_{\gamma_{rp}, \gamma_{rq}}} \sum_{i=1}^{n_{\gamma_{rp}, \gamma_{rq}}} \psi(n_i^3) - \psi(n_i^{1,3}) - \psi(n_i^{2,3}). \quad (6)$$

针对样本数据较少、总体分布未知的情况, 可通过 CCSMI 置换检验方法确定数据的有效性, 其检验定义如下: 给定  $X^p, X^q$  和  $X^R$ , CCSMI 置换检验关联  $p$  值为

$$p = \frac{1}{B} \sum_{b=1}^B \mathbb{I} \text{CCSMI}((X^p)_b; X^q | X^R) \geq \text{CCSMI}(X^p; X^q | X^R). \quad (7)$$

其中:  $(X^p)_b$  表示  $X^p$  的置换,  $\mathbb{I}$  表示指示函数,  $B$  表示计算 CCSMI 的次数. 由其定义可知, 可将其用于 CSMI 置换检验.

## 2 无混杂因子基于 PCSMI 的因果发现

在上述理论上, 本节采用基于 PCSMI 的算法进行因果发现. 本算法假定满足因果充分性(两个可观察时间序列变量不存在混杂因子)、因果马尔可夫性和数据分布忠实性3个条件, 在此基础上对经典 PC 算法规则进行改进, 构建因果图的骨架并进行定向.

### 2.1 构建骨架

本文按照经典 PC 算法相同的步骤, 通过构建全图、删除无因果关系的变量之间连边来构建因果图的骨架. 由于本文目标为构建概要因果图, 在条件判断时允许自环以表示变量自因果性. 首先计算所有时间序列变量对的 CSMI 值以判断两个变量之间是否独立, 若独立则去除其连边; 在此基础上计算 CCSMI 值, 若条件独立则删除连边, 不独立则保留连边. 定义  $\text{Sepset}(p, q)$  表示  $X^p$  和  $X^q$  的分离集, 即  $X^p$  和  $X^q$  在给定这个集合下条件独立. 下面的定理表明, 上述过程构建的骨架是可信的.

**定理1** 设  $G = (V, E)$  为概要因果图, 假定给定子集  $S \subseteq V \setminus \{X^p, X^q\}$ , 集合  $V$  中所有变量对  $(X^p, X^q)$  的条件独立信息都是准确的, 则按上述方式构建的骨架就是  $G$  的骨架.

### 2.2 定向

定向主要是在骨架构建完成后, 按照经典 PC 算法规则确定尽可能多边的方向.

PCSMI 规则1: 在未定向  $X^p - X^r - X^q$  三元组变量中, 若  $X^r \notin \text{Sepset}(p, q)$ , 则  $X^r$  是一个对撞结构, 即  $X^p \rightarrow X^r \leftarrow X^q$ .

PCSMI 规则2: 在部分定向  $X^p \rightarrow X^r - X^q$  三元组变量中, 若  $X^r \notin \text{Sepset}(p, q)$ , 则可将未定向边定向为  $X^p \rightarrow X^r \rightarrow X^q$ .

PCSMI 规则3: 若  $X^p$  与  $X^q$  之间有连边, 且存在从  $X^p$  到  $X^q$  的路径, 则可定向为  $X^p \rightarrow X^q$ .

PCSMI 规则4: 假定有已定向的  $X^p \rightarrow X^r \leftarrow X^q$  和未定向的  $X^p - X^s - X^q$ , 则有  $X^s \rightarrow X^r$ .

除了经典 PC 算法规则之外, 本文引入两个额外定向规则: 因果概要互信息和虚假关联. 因果概要互

信息前文已经进行了描述,虚假关联是指两个变量通过他们之间的变量形成间接因果关系,其定义如下.

**定义5** 当两个变量对  $X^p$  与  $X^q$  之间存在一条路径既不包含对撞结构,也不包含  $X^p - X^q$ ,则称  $X^p$  与  $X^q$  之间存在虚假关联.

当两个变量之间不存在虚假关联时,则可通过CSMI进行定向,定向方法参考如下定理.

**定理2** 假定给定子集  $S \subseteq V \setminus \{X^p, X^q\}$ ,如果  $V$  中所有变量对  $(X^p, X^q)$  的条件独立信息都是准确的,则通过因果充分性和上述过程得到的图中每条非定向边都能表示时序变量间的因果关系.假定最大时间窗口大于因果之间最长时延  $\gamma_{\max}$ ,若变量  $X^p$  与  $X^q$  之间不存在虚假关联,则其连边的定向可通过计算  $\text{CSMI}(X^p, X^q)$  中最优时延来确定.

依据定理2以及定义3,可以在经典PC算法规则基础上增加两个额外定向规则.

**PCSMI规则5:** 对于变量对  $X^p - X^q$ ,若  $X^p$  与  $X^q$  之间没有虚假关联,且  $\bar{\gamma}_{pq} > 0$ ,则可定向为  $X^p \rightarrow X^q$ .

由于原因变量的时间窗口不能大于结果变量的时间窗口,可通过判断变量时间窗口大小来定向,如下规则.

**PCSMI规则6:** 对于变量对  $X^p - X^q$ ,若  $X^p$  与  $X^q$  之间没有虚假关联,且有  $\bar{\gamma}_{pq} = 0, \bar{\lambda}_{pq} < \bar{\lambda}_{qp}$ ,则可定向为  $X^p \rightarrow X^q$ .

基于上述6个规则,本文提出一种基于PCSMI的算法以实现无混杂因子条件下的因果发现,算法输入为时序变量集合,输出为概要因果图.

### 3 有混杂因子基于FCISMI的因果发现

混杂因子条件下因果充分性条件不满足,此时PC算法不再适用,本文采用基于FCISMI的算法进行因果发现.本文算法仍需假设因果马尔可夫性和数据分布忠实性,主要步骤与基于PCSMI的算法基本一致,也分为构建骨架和定向两个步骤.

构建骨架的方法与2.1节一致,本节不再赘述.定向需要引入  $D$  分离全集的概念,在此基础上再制定相应规则.

#### 3.1 $D$ 分离全集

$D$  分离全集中元素节点相当于有向无环图(DAG)中的父节点,因此以  $D$  分离全集中元素为条件,按照2.1节部分的策略可以识别多余的边.  $D$  分离全集的定义如下.

**定义6** 当且仅当  $X^p$  与  $X^q$  之间有一条无向路径  $U$  使得对于其每条子路径  $(X^w, X^s, X^v)$ ,  $X^r$  是  $X^p$

和  $X^q$  的  $D$  分离全集中的元素.其中  $X^s$  是子路径中的对撞结构,或有向无环图(PAG)中  $X^w$  与  $X^v$  相邻.

#### 3.2 定向规则

为构建  $D$  分离全集,参照PCSMI规则1有:

**FCISMI规则1:** 对于三元组变量  $X^p * - \circ X^r \circ - * X^q$ ,若  $X^r \notin \text{Sepset}(p, q)$ ,则可定向为对撞结构  $X^p * \rightarrow X^r \leftarrow * X^q$ .

在建立其余定向规则之前,首先需明确以下3类路径:1)判别路径,是指  $X^p$  与  $X^q$  之间的一条路径使得其中每个非端点节点  $X^r$  与  $X^q$  相邻,  $X^p$  与  $X^q$  不相邻,  $X^p$  与  $X^r$  路径中的每个节点都是对撞结构且是  $X^q$  的父节点,该路径至少包含3条边;2)暴露路径是指每个连续三元组节点都是可观测的路径;3)潜在定向路径,是指两个连续节点不指向第1个节点或指向第2个节点的路径.

参照FCI算法思路并通过上述3类路径构建其余定向规则如下.

**FCISMI规则2:**  $X^p * \rightarrow X^r \circ - * X^q$  三元组变量中,若  $X^r \notin \text{Sepset}(p, q)$ ,则可定向为  $X^p * \rightarrow X^r * \rightarrow X^q$ .

**FCISMI规则3:** 若存在三元组变量  $X^p \rightarrow X^r * \rightarrow X^q$  或  $X^p * \rightarrow X^r \rightarrow X^q$ ,且有  $X^p * - \circ X^q$ ,则可定向为  $X^p * \rightarrow X^q$ .

**FCISMI规则4:** 若存在三元组变量  $X^p * \rightarrow X^r \leftarrow * X^q$  和  $X^p * - \circ X^s \circ - * X^q$ ,且有  $X^s * - \circ X^r$ ,则可定向为  $X^s \rightarrow X^r$ .

**FCISMI规则5:** 若  $X^p$  和  $X^q$  之间存在判别路径且  $X^r \circ - * X^q$ ,则可将  $X^r \circ - * X^q$  定向为  $X^r \rightarrow X^q$ ;否则将三元组定向为  $X^s \leftrightarrow X^r \leftrightarrow X^q$ .

**FCISMI规则6:** 若有  $X^p \rightarrow X^r \rightarrow X^q$  或  $X^p - \circ X^r \rightarrow X^q$ ,且有  $X^p \circ \rightarrow X^q$ ,则可定向为  $X^p \rightarrow X^q$ .

**FCISMI规则7:** 若有  $X^p \circ \rightarrow X^q$  且  $U$  是一条从  $X^p$  到  $X^q$  的未暴露潜在定向路径,使得  $X^p$  与  $X^q$  不相邻,则可定向为  $X^p \rightarrow X^q$ .

**FCISMI规则8:** 假定有  $X^p \circ \rightarrow X^q, X^r \rightarrow X^q \leftarrow X^s, U_1$  是一条从  $X^p$  到  $X^r$  的未暴露潜在定向路径,  $U_2$  是一条从  $X^p$  到  $X^s$  的未暴露潜在定向路径.设  $\mu$  为  $U_1$  上与  $X^p$  相邻的节点 ( $\mu$  可以为  $X^r$ ),  $\omega$  为  $U_2$  上与  $X^p$  相邻的节点 ( $\omega$  可以为  $X^s$ ).如果  $\mu$  和  $\omega$  不同且不相邻,则可定向为  $X^p \circ \rightarrow X^q$  或  $X^p \rightarrow X^q$ .

以上定向规则中并未包含经典FCI算法中处理选择偏差的规则,因为选择偏差情况不会出现在本文使用的数据集中.而后类似于PCSMI算法,增加两个规则定向额外的边.

FCISMI规则9: 对于变量对  $X^p * - \circ X^q$ , 若  $X^p$  与  $X^q$  之间没有虚假关联, 且  $\bar{\gamma}_{pq} > 0$ , 则可定向为  $X^{p*} \rightarrow X^q$ .

FCISMI规则10: 对于变量对  $X^p * - \circ X^q$ , 若  $X^p$  与  $X^q$  之间没有虚假关联, 且有  $\bar{\gamma}_{pq} = 0, \bar{\lambda}_{pq} < \bar{\lambda}_{qp}$ , 则可定向为  $X^{p*} \rightarrow X^q$ .

以上规则中, 双箭头 “ $\leftrightarrow$ ” 表示存在混杂因子, 圆标记 “ $\circ$ ” 为不确定标记, 星号 “ $*$ ” 为通配符, 可以表示箭头、圆或尾标记.

基于上述10个规则, 本文提出一种基于FCISMI的算法以实现混杂因子条件下的因果发现.

### 4 实验分析

本节主要在模拟数据集和真实数据集上评估本文所提出算法的效果. 下面首先介绍本文对比的算法及其数据集设置, 而后构建实验用的模拟数据集并选用真实数据集, 最后对描述实验结果进行分析, 评价标准选用图中有向边的  $F_1$  值.

#### 4.1 对比算法及其参数设置

首先将本文提出的算法与基线算法PC和FCI进行对比, 在此基础上对照因果发现常见几类算法, 对于格兰杰系列因果发现算法, 与代表性MVGC算法进行对比; 对于基于评分的算法, 与代表性DYNOTEARS算法进行对比; 对于基于函数的算法, 与代表性TiMINo算法进行对比; 对于基于约束的算法, 与代表性PCMC和tsFCI算法进行对比; 对于深度学习因果发现算法, 与代表性CGTST算法进行对比, 其中自注意机制头设置为4个, 编码层数设置为3, transformer维度为70, epoch值为100.

所有涉及互信息的算法, 本文采用  $K$ -近邻方法进行估计, 近邻数  $K = 10$ . 此外通过显著性置换检验判断数据的有效性, 其中  $\gamma_{\max} = 5$ , 显著性水平为0.05.

### 4.2 模拟数据集构建和真实数据集选用

#### 4.2.1 模拟数据集构建

在满足因果充分性条件下, 本文生成了包含分叉结构、V-结构、菱形结构3种因果结构的模拟数据, 样本大小为1000. 在不满足因果充分性条件下, 本文生成了5节点2混杂因子结构的模拟数据. 对于每类结构, 本文按照以下过程生成6个数据集. 对于所有  $q, X_0^q = 0$ , 对于所有  $t > 0$ , 有

$$X_t^q = a_{t-1}^{qq} X_{t-1}^q + \sum_{\substack{(p,\gamma) \\ X_{t-\gamma}^p \in \text{Par}(X_t^q)}} a_{t-\gamma}^{pq} f(X_{t-\gamma}^p) + 0.1 \xi_t^q \quad (8)$$

其中:  $\gamma \geq 0; a_t^{jq}$  为  $U([-1; -0.1] \cup [0.1; 1])$  中抽取的随机系数,  $1 \leq j \leq q; \xi_t^q \sim N(0, \sqrt{15}); f$  为绝对值、tan、正弦和余弦中均匀随机选择的非线性函数.

#### 4.2.2 真实数据集选用

本文主要在雅虎网络流量数据集和纽约出租车数据集上进行验证. 雅虎网络流量数据集来自于Kaggle网站雅虎实验室 (<https://www.kaggle.com/datasets>), 由网络流量、暴力FTP、DDoS、Web攻击等67个不同的时序变量组成, 每个时序变量包含1400个左右的时间戳, 时间戳每半小时记录一次. 纽约出租车数据集来自于Kaggle网站Numenta Anomaly Benchmark, 记录的是纽约市2014-07-01~2015-01-31日出租车需求, 每半小时记录一次. 除出租车需求变量外, 还包含纽约市马拉松赛、纽约市感恩节活动、纽约市圣诞节活动等25个时序变量.

### 4.3 实验结果及分析

#### 4.3.1 模拟数据结果及分析

表1显示了无混杂因子条件下, 所有算法在模拟数据集上运行的性能, 主要通过  $F_1$  值的均值和标准差体现, 最佳性能加粗表示.

表1 无混杂因子模拟数据集上各算法性能

结构	PCSMI	PC	PCMC	TiMINo	Dynotears	MVGC	CGTST
分叉结构	<b>0.77</b> ±0.19	0.43±0.20	0.65±0.37	0.62±0.35	0.08±0.21	0.38±0.25	0.70±0.20
V-结构	<b>0.82</b> ±0.17	0.45±0.23	0.77±0.17	0.53±0.24	0.07±0.20	0.42±0.36	0.78±0.12
菱形结构	<b>0.81</b> ±0.12	0.35±0.21	<b>0.81</b> ±0.13	0.60±0.23	0.24±0.25	0.68±0.26	0.77±0.16

由表1中数据可以看出, 相比基线PC算法, 本文提出的PCSMI算法在3类结构上性能提升均十分明显(特别是在菱形结构上), 主要原因是PC算法难以解决马尔可夫等价类问题, 导致其准确率偏低; CGTST算法虽然在3类结构上相较其他算法取得了较好性能, 但仍略低于PCSMI算法; PCMC算

法在菱形结构上虽然与PCSMI算法  $F_1$  值的均值相同, 但其标准差要略高于PCSMI算法. 这主要是因为PCSMI算法中额外定向规则具有较高的因果结构发现率以及较好的稳健性.

表2显示了混杂因子条件下本文提出的FCISMI算法与基线FCI算法、tsFCI算法的性能比较. 相比

基线FCI算法,本文提出的FCISMI算法性能提升十分明显,主要原因是FCI算法对变量自因果特性考虑不足导致其准确率偏低;同时,FCISMI性能也优于tsFCI,但其 $F_1$ 值的标准差略高于tsFCI,这是因为

表2 有混杂因子模拟数据集上各算法性能

结构	FCISMI	FCI	tsFCI
5节点2混杂因子	0.45±0.10	0.25±0.18	0.38±0.09

表3 真实数据集上各算法性能

数据集	PCSMI	PC	PCMCI	TiMINo	Dynotears	MVGC	CGTST
雅虎网络流量	0.33±0.16	0.15±0.15	0.23±0.17	0.33±0.10	0.48±0.26	0.25±0.17	0.29±0.16
纽约出租车	0.41±0.31	0.22±0.16	0.26±0.30	0.61±0.13	0.00±0.00	0.37±0.16	0.40±0.18

在雅虎网络流量数据集上,所有算法均可以识别出暴力FTP、DDoS、Web攻击等时序变量是网络流量变量的原因,但存在不同程度错误和遗漏.其中Dynotears算法性能最好,其次是PCSMI算法和TiMINo算法,其原因可能是因为雅虎网络流量数据集中因果关系近似为线性关系,Dynotears算法较为适用,而本文提出的PCSMI算法性能也能做到接近最优.

在纽约出租车数据集上,除Dynotears外其余算法均可在不同程度上识别出纽约市感恩节、圣诞节活动等变量是纽约出租车需求变量的原因.TiMINo算法表现最好,其次是PCSMI算法.在对各算法生成的概要因果图进行分析后发现,TiMINo算法生成的概要因果图有60%为全连接图,并不符合实际情况.因此TiMINo算法 $F_1$ 值虚高,实际性能并不理想.综合表3中数据可以得出在真实数据集上,本文提出的PCSMI算法性能优于其他算法.

### 4.3.3 算法复杂度分析

本小节主要对基于约束算法的复杂度进行分析.在最差情况下,PC算法构建窗口因果图复杂度上限为 $(d\gamma_{\max})^2(d\gamma_{\max}-1)^{k-1}/(k-1)!$ .其中: $k$ 表示任意节点的最大度数, $\gamma_{\max}$ 表示最大时延.算法每步都需要进行条件互信息显著性检验.PCMCI算法主要通过各时序变量信息来减少检验次数.本文提出的PCSMI算法相比PCMCI算法进一步减少了测试数量,因为生成概要因果图所需检验次数少于生成窗口因果图所需检验次数,PCSMI算法复杂度上限为 $d^2(d-1)^{k-1}/(k-1)!$ .

图3显示了PCSMI和PCMCI算法在3种模拟数据结构(分叉结构、V-结构、菱形结构)中运行时间结果.可以看出,PCMCI算法识别分叉结构和V-结构耗

时约3000s,识别菱形结构耗时6500s;PCSMI算法识别分叉结构和V-结构耗时约600s,识别菱形结构耗时1000s,复杂度明显优于PCMCI算法,特别是在更复杂结构(菱形结构)中差别更明显.

### 4.3.2 真实数据结果及分析

表3显示了在雅虎网络流量和纽约出租车数据集上各算法的性能.由表3中数据可以看出,相较基线PC算法,PCSMI算法在雅虎网络流量数据集和纽约出租车数据集性能均有较大提升.

时约3000s,识别菱形结构耗时6500s;PCSMI算法识别分叉结构和V-结构耗时约600s,识别菱形结构耗时1000s,复杂度明显优于PCMCI算法,特别是在更复杂结构(菱形结构)中差别更明显.

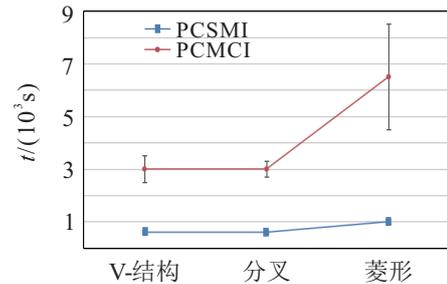


图3 PCSMI和PCMCI算法运行时间

### 4.3.4 超参数分析

本小节主要对最大时延 $\gamma_{\max}$ 和 $K$ 近邻方法中的 $K$ 值进行分析.按照4.2.1节中模拟数据集的构建方法,主要进行以下两个实验.

首先计算数据集上的CCSMI( $X^2; X^3|X^1$ )和CSMI( $X^2; X^3$ )的平均值和标准差, $\gamma_{\max}$ 在3~10之间变化,结果如图4所示,其中点代表均值,短横线代表标准差.结果表明,CSMI和CCSMI对于 $\gamma_{\max}$ 是稳健的: $\gamma_{\max}=3,4,5$ 时CSMI( $X^2; X^3$ )均值保持不变, $\gamma_{\max}=10$ 时略有下降, $\gamma_{\max}=3,4$ 时CSMI( $X^2; X^3$ )标准差保持不变, $\gamma_{\max}=5,10$ 时略有增加;对于所有 $\gamma_{\max}$ ,CCSMI( $X^2; X^3|X^1$ )的均值和标准差保持相同.由以上结果得出 $\gamma_{\max}$ 可在3~5间选用, $\gamma_{\max}$ 过小则会略掉变量间信息导致因果漏识别, $\gamma_{\max}$ 过大则变量自因果影响增大导致因果误识别.

其次,继续计算数据集上CCSMI( $X^2; X^3|X^1$ )和CSMI( $X^2; X^3$ )的平均值和标准差, $k$ 在5~100之间变化,结果如图5所示.结果表明,CCSMI( $X^2; X^3|X^1$ )均值和标准差对于 $k$ 是稳健的:CSMI( $X^2; X^3$ )均值

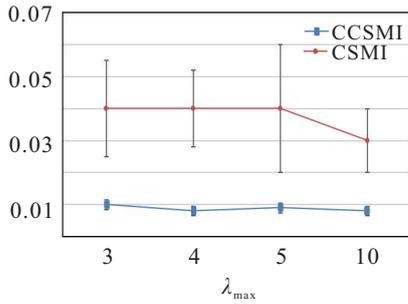


图4  $\gamma_{max}$ 变化对应CSMI和CCSMI变化

和标准差在  $k = 5$  和  $k = 10$  之间略有下降,在  $k = 10$  至  $k = 50$  之间大幅下降,在  $k = 50$  至  $k = 100$  之间恢复稳健. 由以上结果得出  $k$  可在 10 左右选用,  $k$  过小则模型过于复杂产生过拟合导致因果误识别,  $k$  越大则模型过于简单忽略了有用信息导致因果漏识别.

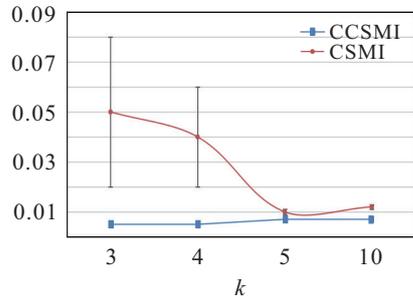


图5  $k$ 变化对应CSMI和CCSMI变化

#### 4.4 消融实验

本节主要对本文提出的 PCSMI 和 FCISMI 算法中相比基线 PC 和 FCI 算法增加的 PCSMI 规则 5、规则 6, 以及 FCISMI 规则 9、规则 10 进行消融研究, 以验证算法的有效性.

表 4 和表 5 显示了无混杂因子和有混杂因子条件下无 PCSMI 规则 5、规则 6 以及 FCISMI 规则 9、规则 10 情况下算法在模拟数据集(菱形结构和 5 节点 2 混杂因子)中的性能.

表 4 无混杂因子消融实验结果

结构	PCSMI	无 PCSMI 规则 5	无 PCSMI 规则 6
菱形结构	$0.81 \pm 0.12$	$0.65 \pm 0.18$	$0.52 \pm 0.15$

表 5 有混杂因子消融实验结果

结构	FCISMI	无 FCISMI 规则 9	无 FCISMI 规则 10
5 节点 2 混杂	$0.45 \pm 0.10$	$0.37 \pm 0.14$	$0.31 \pm 0.15$

由表 4 可以看出, 在无混杂因子条件下, 无 PCSMI 规则 5 和无 PCSMI 规则 6 两者算法性能均低于 PCSMI 算法性能, 而无 PCSMI 规则 5 比无 PCSMI 规则 6 更接近 PCSMI 算法性能. 这表明 PCSMI 规则 5 和 PCSMI 规则 6 均对算法性能有贡献, 而 PCSMI 规则 6 贡献更大, 原因是多数情况下最优时延均大于 0,

因此 PCSMI 规则 6 中通过 CSMI 值非对称性增长识别因果关系准确率更高.

由表 5 可以看出, 在混杂因子条件下, 无 FCISMI 规则 9 和无 FCISMI 规则 10 两者算法性能均低于 FCISMI 算法性能, 而无 FCISMI 规则 9 比无 FCISMI 规则 10 更接近于 FCISMI 算法性能. 这表明 FCISMI 规则 9 和 FCISMI 规则 10 均对算法性能有贡献, 而 FCISMI 规则 10 贡献更大, 原因同上述混杂因子条件下部分论述一致.

## 5 结论

本文根据是否满足因果充分性条件分别提出 PCSMI 和 FCISMI 算法, 构建概要因果图从而识别时序数据因果关系. 实验表明, PCSMI 和 FCISMI 算法具有较高的准确率和较低的复杂度, 能够较好地发现时序数据因果关系. 由于本文方法需满足 3 条假设, 对于随机性较强、采样时间不规律的应用场景具有一定局限性, 后续的研究将结合强化学习等方法, 重点解决不满足假设条件下的因果发现问题, 从而拓展方法应用范围.

### 参考文献(References)

- [1] Assaad C K, Devijver E, Gaussier E. Survey and evaluation of causal discovery methods for time series[J]. Journal of Artificial Intelligence Research, 2022, 73: 767-819.
- [2] Ana Rita Nogueira, Andrea Pugnana, Salvatore ruggieri. Methods and tools for causal discovery and causal inference[J]. WIREs Data Mining and Knowledge Discovery, 2022, 12(2): 1-39.
- [3] Jizba P, Lavička H, Tabachová Z. Causal inference in time series in terms of rényi transfer entropy[J]. Entropy, 2022, 24(7): 855.
- [4] Granger C W J. Investigating causal relations by econometric models and cross-spectral methods[J]. Econometrica, 1969, 37(3): 424.
- [5] van Leeuwen P J, DeCaria M, Chakraborty N, et al. A framework for causal discovery in non-intervenable systems[J]. Chaos: An Interdisciplinary Journal of Nonlinear Science, 2021, 31(12): 123-147.
- [6] Nauta M, Bucur D, Seifert C. Causal inference: What if[M]. MA: CRC Press, 2020: 69-80.
- [7] 乔杰, 蔡瑞初, 郝志峰. 一种基于信息瓶颈的因果关系挖掘方法[J]. 计算机科学, 2022, 49(2): 198-203. (Qiao J, Cai R C, Hao Z F. Mining causality via information bottleneck[J]. Computer Science, 2022, 49(2): 198-203.)
- [8] Amirinezhad A, Salehkaleybar S, Hashemi M. Active learning of causal structures with deep reinforcement

- learning[J]. *Neural Networks*, 2022, 154(1): 22-30.
- [9] 陈柘, 刘嘉华, 赵斌, 等. 基于GCN和TCN的多因素城市路网出租车需求预测[J]. *控制与决策*, 2023, 38(4): 1031-1038.  
(Chen Z, Liu J H, Zhao B, et al. Multi-factor taxi demand forecasting for urban road network based on GCN and TCN[J]. *Control and Decision*, 2023, 38(4): 1031-1038.)
- [10] 闫敬, 徐龙, 曹文强, 等. 基于深度强化学习的多潜器编队控制算法设计[J]. *控制与决策*, 2023, 38(5): 1457-1463.  
(Yan J, Xu L, Cao W Q, et al. Design of formation control algorithm for multiple autonomous underwater vehicles based on deep reinforcement learning[J]. *Control and Decision*, 2023, 38(5): 1457-1463.)
- [11] Pulakesh U, Zhang K, Li C, et al. Scalable causal structure learning: Scoping review of traditional and deep learning algorithms and new opportunities in biomedicine[J]. *JMIR Medical Informatics*, 2023, 15(11): 266-281.
- [12] Albers D J, Hripcsak G. Estimation of time-delayed mutual information and bias for irregularly and sparsely sampled time-series[J]. *Chaos, Solitons & Fractals*, 2012, 45(6): 853-860.
- [13] Granger C W J. Time series analysis, cointegration, and applications[J]. *American Economic Review*, 2004, 94(3): 421-425.
- [14] Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information[J]. *Physical Review E*, 2004, 69(6): 066138.
- [15] Nauta M, Bucur D, Seifert C. Causal discovery with attention-based convolutional neural networks[J]. *Machine Learning and Knowledge Extraction*, 2019, 1(1): 312-340.
- [16] Pamfil R, Sriwattanaworachai N, Aragam B. DYNOTEARS: Structure learning from Time-Series data[C]. *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics*. Chongqing, 2020: 1595-1605.
- [17] Chickering D M. Learning equivalence classes of Bayesian-network structures[J]. *Journal of Machine Learning Research*, 2002, 18(2): 445-498.
- [18] Peters J, Janzing D, Schölkopf B. Causal inference on time series using restricted structural equation models[C]. *The 27th Annual Conference on Neural Information Processing Systems*. California, 2013: 1540-1555.
- [19] Assaad C K, Devijver E, Gaussier E, et al. A mixed noise and constraint-based approach to causal inference in time series[C]. *Machine Learning and Knowledge Discovery in Databases*. Research Track. Cham: Springer International Publishing, 2021: 453-468.
- [20] Affeldt S, Isambert H. Robust reconstruction of causal graphical models based on conditional 2-point and 3-point information[C]. *Proceedings of the UAI 2015 Conference on Advances in Causal Inference*. Michigan, 2015: 1200-1219.
- [21] Spirtes P, Glymour C, Scheines R. *Causation, Prediction, and Search*[M]. Cambridge: The MIT Press, 2001.
- [22] 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.
- [23] Entner D, Hoyer P O. On causal discovery from time series data using FCI[C]. *Proceedings of the 5th European Workshop on Probabilistic Graphical Models*. North Carolina, 2010: 121-128.
- [24] Bi X T, Wu D Y, Xie D X, et al. Large-scale chemical process causal discovery from big data with transformer-based deep learning[J]. *Process Safety and Environmental Protection*, 2023, 173: 163-177.
- [25] Schreiber T. Measuring information transfer[J]. *Physical Review Letters*, 2000, 85(2): 461-464.
- [26] Frenzel S, Pompe B. Partial mutual information for coupling analysis of multivariate time series[J]. *Physical Review Letters*, 2007, 99(20): 204101.

## 作者简介

李德志(1989—), 男, 讲师, 博士生, 从事因果推断、指挥控制等研究, E-mail: 469549382@qq.com;

鲁云军(1973—), 男, 教授, 博士生导师, 从事指挥控制、指挥信息系统等研究, E-mail: lu\_yunjun@hotmail.com;

吴健平(1989—), 男, 讲师, 博士生, 从事指挥信息系统的研究, E-mail: wjp860343510@sina.com;

李强(1981—), 男, 副教授, 硕士, 从事指挥控制、指挥信息系统等研究, E-mail: jjandyoyo918618@163.com.