

控制与决策

Control and Decision

从结构推断到根因识别——工业过程故障根因诊断研究综述

赵春晖, 宋鹏宇

引用本文:

赵春晖,宋鹏宇. 从结构推断到根因识别——工业过程故障根因诊断研究综述[J]. *控制与决策*, 2023, 38(8): 2130–2157.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2023.0270>

您可能感兴趣的其他文章

Articles you may be interested in

[基于核典型相关性-熵成分分析的工业过程质量监测方法](#)

A quality monitoring method for industrial process based on kernel canonical correlation-entropy component analysis

控制与决策. 2021, 36(12): 2999–3006 <https://doi.org/10.13195/j.kzyjc.2020.0548>

[基于交叉熵的改进NPE间歇过程故障检测算法](#)

Improved NPE batch process fault detection algorithm based on cross entropy

控制与决策. 2021, 36(2): 411–417 <https://doi.org/10.13195/j.kzyjc.2019.0725>

[基于自适应混合核典型变量分析的工业过程质量相关故障检测](#)

Quality-related fault detection for industrial processes based on adaptive mixed kernel canonical variable analysis

控制与决策. 2021, 36(4): 801–807 <https://doi.org/10.13195/j.kzyjc.2020.0690>

[基于多块信息提取的AUV资源勘查系统故障检测](#)

Fault detection of AUV resource exploration system based on multi-block information extraction

控制与决策. 2021, 36(4): 790–800 <https://doi.org/10.13195/j.kzyjc.2019.0732>

[基于广义主成分分析的重构故障子空间建模方法](#)

Reconstructed fault subspace modelling method based on generalized principal component analysis

控制与决策. 2021, 36(4): 808–814 <https://doi.org/10.13195/j.kzyjc.2019.0818>

从结构推断到根因识别——工业过程故障 根因诊断研究综述

赵春晖, 宋鹏宇[†]

(浙江大学 控制科学与工程学院, 杭州 310027)

摘要: 由于现代工业过程的复杂结构, 变量间普遍存在紧密耦合, 故障往往在变量间广泛传播, 为过程运维带来挑战. 针对该问题, 工业根因诊断 (industrial root cause diagnosis, IRCD) 技术应运而生, 其从异常变量中确定故障根因, 便于针对性故障处理. IRCD 包含两个主要步骤: 结构推断和根因识别. 前者建立变量间的信息传递结构; 后者根据传递结构定位根因. 然而, 现有 IRCD 综述多侧重于结构推断, 未对根因识别步骤进行调研, 且未建立起各类 IRCD 模型与过程特性间的系统关联. 为此, 从结构推断和根因识别两个层级展开 IRCD 的研究综述. 首先, 依据推断准则的异同, 归纳 4 类经典结构推断模型; 其次, 考虑到过程的高维度、非线性、非平稳性质以及机理知识的效用, 对结构推断模型的变种及适用场景进行梳理; 随后, 对根因识别方法进行归类, 包括纯数据驱动、知识与数据融合驱动的模式, 涵盖 6 类典型方法, 并分析它们的优势与不足; 最后, 讨论 IRCD 技术中存在的挑战, 并给出未来研究方向, 为后续研究提供参考.

关键词: 工业过程; 工业根因诊断; 结构推断; 根因识别; 因果分析; 研究综述

中图分类号: TP277

文献标志码: A

DOI: 10.13195/j.kzyjc.2023.0270

引用格式: 赵春晖, 宋鹏宇. 从结构推断到根因识别——工业过程故障根因诊断研究综述 [J]. 控制与决策, 2023, 38(8): 2130-2157.

From structure inference to root cause identification: A survey on root cause diagnosis of industrial process faults

ZHAO Chun-hui, SONG Peng-yu[†]

(College of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China)

Abstract: Due to the complex structures of modern industrial processes, there are ubiquitous substantial couplings among variables, and thus faults can spread widely among them, which brings challenges to process maintenance. To address this issue, industrial root cause diagnosis (IRCD) technology has been developed, which determines the root cause of the fault from abnormal variables to facilitate targeted fault processing. IRCD consists of two main steps: structure inference and root cause identification. Structure inference establishes the information transmission structure among variables, and the root cause identification locates the root cause according to the structure. However, most of the existing IRCD reviews focus on structure inference without investigating the root cause identification step. In addition, they have not established systematic connections between various IRCD models and different process characteristics. To this end, a survey of IRCD studies is conducted in two stages: structure inference and root cause identification. First, four types of classical structure inference models are summarized according to their inference criteria. Then, considering the high-dimensional, nonlinear, and nonstationary nature of the process as well as the utility of mechanism knowledge, the variants and application scenarios of the structure inference models are sorted out. Afterward, the root cause identification methods are classified, including pure data-driven as well as hybrid knowledge-and-data-driven paradigms, covering six typical methods. The advantages and drawbacks of these methods are also analyzed. Finally, the unsolved challenges in IRCD are discussed, and future directions are put forward to provide references for follow-up research.

Keywords: industrial process; industrial root cause diagnosis; structure inference; root cause identification; causal analysis; survey

收稿日期: 2023-03-11; 录用日期: 2023-05-16.

基金项目: 国家杰出青年科学基金项目 (62125306); 广东省基础与应用基础研究基金项目 (2022A1515240003).

责任编辑: 丁进良.

[†]通讯作者. E-mail: pysong@zju.edu.cn.

0 引言

工业过程是制造业的重要组成部分,是国民经济发展的重要基础,其安全高效生产对国家而言具有重要的战略意义.随着经济的飞速发展,为了满足日益增长且呈现多样化趋势的生产需求,现代工业过程正不断向着大规模、复杂化的方向演变,在提高生产能力的同时也为工业过程的安全稳定运行带来了严峻挑战^[1-5].部分生产制造车间已经为工业设备装配了传感器及报警系统,从而监控关键过程变量的异常状况.然而,由于过程的复杂结构与庞大规模,过程变量往往数目众多,而且控制回路或物理连接会使其彼此紧密耦合.在这种情况下,故障一旦发生,故障信息可能会在不同变量间复杂传播,导致异常变量繁多,难以采取有效的故障处理措施.

为解决上述挑战,根因诊断(root cause diagnosis, RCD)技术应运而生.RCD的概念最早可追溯到20世纪90年代,当时的研究者主要依托于故障诊断或过程机理知识来推断故障的根因^[6].然而,由于知识收集成本高昂、逻辑运算复杂等问题,当时的RCD技术未能在工业过程领域得到普及.近年来,随着机器学习、因果分析等研究的发展与应用,数据驱动的RCD方法为工业过程的智能故障诊断带来全新的解决方案^[7-9],也促进了相关研究的发展.一般来说,数据驱动的RCD在故障发生后通过传感器获取故障过程变量的时间序列采样,并利用相似度量^[10-12]、因果分析^[13-15]等技术提取变量之间的信息传递网络结构,从而揭示故障传播的路径,本文将这一步骤定义为结构推断.值得注意的是,尽管工业RCD(industrial RCD, IRCD)以因果分析为代表的理论作为研究基础,但由于过程故障往往随时间传播演化,工业领域内主要关注时序上的因果分析技术.此外,由于工业数据具有高维度、非线性、非平稳等典型性质,也促使领域内的学者对因果分析技术进行改进以切合实际数据特性,而不仅仅是理论方法的简单应用.不仅如此,上述的结构推断实际上只能给出故障传播的路径,而考虑到过程内部众多的连接与回路,这种路径可能是错综复杂而难以人为分析的.由于工业现场对故障处理效率要求较高,有学者主张在结构推断的基础上对变量的信息传递进行深入分析,并结合机理知识等信息从复杂的结构中自动识别出根因变量^[16-19],本文将这一步骤定义为根因识别. IRCD方法通过结构推断和根因识别的双层级从复杂过程数据中挖掘故障根因变量,形成了具有工业特色的诊断范式,为故障处理措施的制定提供了参考依据.

为分析IRCD技术的发展趋势,在谷歌学术网站中以“root cause diagnosis”或“root cause tracing”结合“process”或“industrial”作为组合关键词进行文献检索,可以获取领域中历年发表文献的数目,如图1所示.不难发现,关于IRCD的研究数目在近20年间呈现显著的上升趋势,值得注意的是,自2011年开始,IRCD相关的文献数目的增长速率明显加快,这得益于同时期人工智能与大数据技术的发展,一定程度上反映了数据驱动的IRCD方法正逐步成为领域内的研究热点.同时,以“causal inference”或“causal analysis”进行检索,获取IRCD的理论基础,即因果分析相关研究的历年文献发表数目.可以发现,因果分析的研究同样呈现整体上升趋势,且增长热潮早于IRCD领域,这也从侧面反映了因果相关的研究为IRCD的发展与进步奠定了理论基础.此外,通过同阶段的研究数目可以发现,因果理论研究的数目显著多于IRCD,说明因果分析相关技术在工业领域还有广阔的研究与应用空间.然而,相比于国际上IRCD研究的迅猛增长趋势,国内的相关研究并不多见,在中国知网中搜索“根因诊断”或“溯源诊断”,所获取的相关文献总数不足70篇,这说明我国的IRCD技术发展相对较晚,尚存广阔的探索空间.

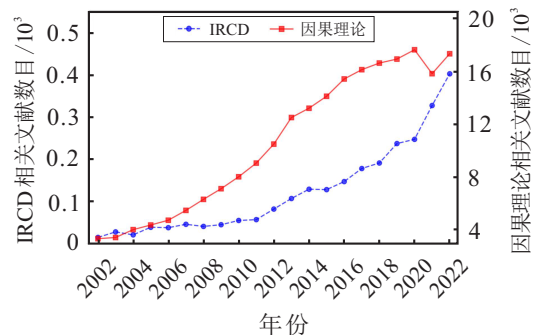


图1 谷歌学术网站中对近20年IRCD以及因果分析领域相关文献的组合关键词检索数目统计

尽管工业过程故障的RCD已经得到了一定发展,目前领域内仍缺乏对RCD技术系统的、全面的研究综述.具体而言,现有的IRCD综述^[20-22]主要侧重于结构推断步骤,未对根因识别进行深入讨论;此外,现有综述文章通常关注于经典结构推断模型的理论分析,事实上随着数据分析技术的不断进步,这些经典模型已经衍生出诸多改进版本以处理更加复杂多样的数据模式,而这些改进方法与不同过程数据特性之间的适配关系尚未被系统地梳理挖掘.

针对上述问题,本文从结构学习与根因识别两个层级依次展开,对工业过程的RCD研究进行系统性研究综述,文章整体架构如图2所示.首先,依据推断

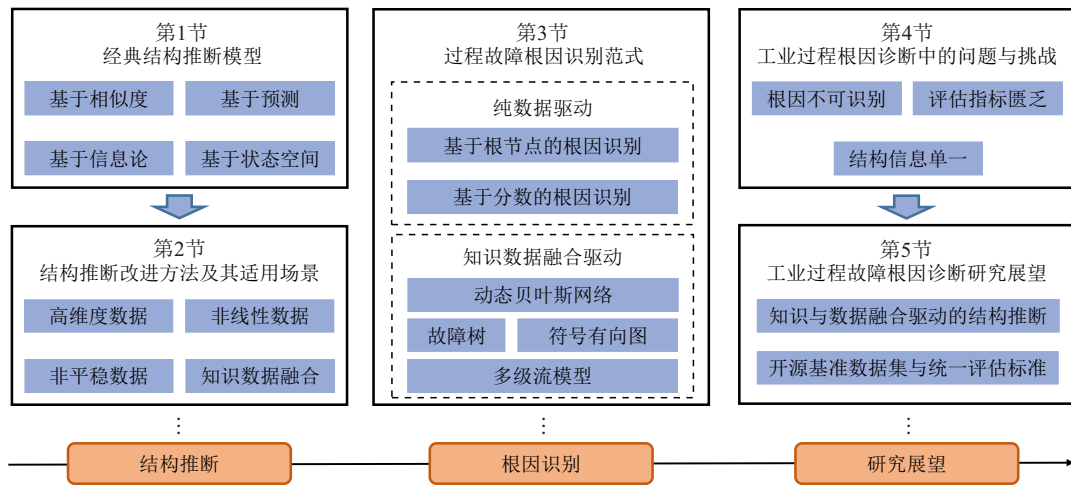


图2 本文的组织架构

准则的异同,本文在第1节中归纳了4类主要的经典结构推断模型,包括基于相似度、预测、信息理论,以及状态空间的方法.随后,考虑到过程中高维度、非线性、非平稳等典型数据特性以及过程机理知识的效用,在第2节中梳理了上述4类方法各自的改进版本及适用场景,实现结构推断理论与过程特性间的关联匹配.在第3节中,本文对现有的根因识别范式进行归类总结,包括纯数据驱动和数据与知识融合驱动两种范式,涵盖基于根节点、分数以及动态贝叶斯网络、故障树、符号有向图、多级流模型的识别方法,并对比分析了这些方法的优势与不足.在第4节中,从根因变量的可识别性、IRCD方法性能的评估指标,以及IRCD中的信息传递结构3个方面探讨工业过程故障RCD中尚未解决的问题与挑战.最后,本文基于上述章节的讨论结果,引出潜在的未来研究趋势,包括知识与数据融合驱动的诊断策略以及开源基准数据集与统一评估标准,为推动后续研究提供参考.

本文的主要贡献点总结如下:

1) 本文从结构推断与根因识别两个层级递进地对工业过程故障的RCD研究进行全面综述.对于上述两个层级中的众多方法,本文根据模型原理和应用场景的不同进行了系统地梳理分类,并对不同方法的性质与适用工业对象给予直观的展示与对比.

2) 本文通过对IRCD领域研究现状的梳理与分析,对现有研究的问题与挑战进行了归纳总结,并在此基础上进行展望,提出了若干潜在的未来方向,为后续的研究提供有力的依据和参考.

1 经典结构推断模型

结构推断是RCD任务中的首要步骤,用于获取故障变量间的信息传递关系.给定 N 个故障变量

$\{\mathbf{x}_1(t), \dots, \mathbf{x}_N(t)\}$,其中 $\mathbf{x}_i(t)$ 表示第 i 个变量的时间序列.结构推断的目的是获取一个具有 N 个节点的有向图,其中图节点用于表示变量,而节点之间的连边用于指示两个变量间信息传递关系的有无及方向,通过这样的图结构反映故障在不同变量之间的传播的路径.本文依据结构推断的准则不同,将现有IRCD方法分为两个大类:非因果模型和因果分析模型.前者主要指基于相似度的方法;后者则包括3个主要类别:基于预测、信息论以及状态空间的方法.本节将对这些方法的原理和性能进行详细介绍与分析.

1.1 基于相似度的非因果结构推断模型

在因果分析技术得到广泛发展和关注之前,IRCD领域内主要基于时序相似度这种非因果的方式来分析两个变量间的时延关系.时延反映了两个变量跨时间的相似性.具体而言,给定两个时间序列 $\mathbf{x}(t)$ 和 $\mathbf{y}(t)$,以及相似度度量指标 S ,则两个序列的时延可以定义为

$$\tau = \arg \max_k S(\mathbf{x}(t), \mathbf{y}(t+k)), \quad (1)$$

即两个序列错位 τ 个采样点后最为相似.值得注意的是,这里的时延 τ 的正负能够表示错位的方向,这使得时延可用于描述信息传递的方向.若 $\mathbf{x}(t)$ 和 $\mathbf{y}(t+\tau)$ 足够相似且 τ 取值为正,则说明信息从变量 \mathbf{x} 流向 \mathbf{y} ;反之,若 τ 取值为负,则表明信息是从 \mathbf{y} 流向 \mathbf{x} .基于这种思想,变量间的信息传递结构可以被确定下来.不同的基于相似度的推断模型本质上都遵循这种思想,区别在于相似度度量指标 S 不同.

1.1.1 互相关函数

互相关函数(cross-correlation function, CCF)^[10]是错位采样点数目 k 的函数,其本质上计算了两个时间序列错位后的皮尔逊相关系数.给定时间序列 \mathbf{x}

流向 \mathbf{y} , 二者的 CCF 按如下方式计算:

$$\rho(k) = \frac{1}{T-k} \sum_{j=1}^{T-k} \frac{[\mathbf{x}(j) - \mu_x][\mathbf{y}(j) - \mu_y]}{\sigma_x \sigma_y}. \quad (2)$$

其中: T 为采样点总数, k 为错位时间点数目, $\mathbf{x}(j)$ 和 $\mathbf{y}(j)$ 为在第 j 个时间点上的采样, μ_x 和 μ_y 分别表示 $\mathbf{x}(t)$ 和 $\mathbf{y}(t)$ 的均值, σ_x 和 σ_y 分别表示 $\mathbf{x}(t)$ 和 $\mathbf{y}(t)$ 的标准差.

将 $\rho(k)$ 的最大值和最小值分别记为 ρ_M 和 ρ_m , 其取最大和最小值时的错位采样点数目记作 k_M 和 k_m . 根据上述时延的定义, 基于 CCF 的时延按如下方式计算:

$$\tau = \begin{cases} k_M, & \rho_M + \rho_m \geq 0; \\ k_m, & \rho_M + \rho_m < 0. \end{cases} \quad (3)$$

这种计算方式的设计是考虑到相关系数可能为负值, 本质上是选取了使得相关系数绝对值最大的错位采样点数目作为时延.

为确定信息传递关系是否显著, Bauer 等^[23] 设计了一种方向性指数 ψ , 即

$$\psi = 2 \frac{|\rho_M + \rho_m|}{\rho_M + |\rho_m|}. \quad (4)$$

经过统计检验, Bauer 等^[23] 推导出当 $\psi \geq 0.46T^{-0.16}$ 时, 可认为变量间存在显著的信息传递关系, 其方向可根据时延的正负来确定.

Bauer 等^[23] 于 2008 年首次将 CCF 应用于 RCD 任务. 然而, Landman 等^[24] 在后续的研究中发现, 这种方法存在显著的缺陷: 首先, 其难以适用于振荡型故障, 因为振荡通常导致过程变量具有周期性, 从而使得 CCF 也随着错位样本数变化而呈现周期性波动, 难以确定真正的时延; 此外, 这种方法只能衡量两个变量间的影响, 忽略了多过程变量间的相互作用, 因此容易给出冗余的信息传递关系.

1.1.2 动态时间规整

动态时间规整 (dynamic time warping, DTW)^[11] 是一种典型的时间序列距离度量手段, 其通过动态规划实现序列间相似点的自动匹配, 从而衡量两个时间序列间的趋势相似性, 其取值越小说明序列越相似. Li 等^[7] 将 DTW 应用于工业过程故障的 RCD 中, 对于时间序列 $\mathbf{x}(t)$ 和 $\mathbf{y}(t)$, 设计了一种基于 DTW 的时延指标 D 来确定二者的信息传递关系及方向, 即

$$D = \frac{\min \text{DTW}(\mathbf{x}(t), \mathbf{y}(t+k))}{\text{DTW}(\mathbf{x}(t), \mathbf{y}(t))}. \quad (5)$$

当 D 取值大于一定阈值时, 可以认为 $\mathbf{x}(t)$ 与 $\mathbf{y}(t)$ 之间存在显著的信息传递关系, 传递方向同样依据时延的正负来确定. 时延可按如下方式计算:

$$\tau = \arg \min_k \text{DTW}(\mathbf{x}(t), \mathbf{y}(t+k)). \quad (6)$$

Li 等^[7] 发现 DTW 能够处理非平稳过程中的结构推断问题. 然而, DTW 本质上衡量的是时间序列的形状相似度, 而某些变量之间虽具有显著信息传递关系, 但形状未必相似, 例如 $\mathbf{x}(t) = t$ 和 $\mathbf{y}(t) = \mathbf{x}^2(t-1)$. 因此, 基于 DTW 的 IRCD 方法并未在此基础上得到更进一步的研究.

1.1.3 K 近邻方法

K 近邻 (K nearest neighborhoods, KNN)^[12] 是一种经典的数据挖掘算法, 其通过寻找最近邻样本的方式来度量两个变量间的非线性距离. Stockmann 等^[25] 于 2012 年将其应用于 IRCD 任务. 具体而言, 给定两个具有 T 个采样点的变量 \mathbf{x} 和 \mathbf{y} , 对于 \mathbf{x} 的每个采样点 $x(t)$, 该方法利用 KNN 找出 \mathbf{x} 中与之最为接近的 K 个其他采样点 $\{x(t_1), \dots, x(t_K)\}$. 进而获取变量 \mathbf{y} 中与这 K 个近邻点相对应的错位 k 个采样点的样本, 并利用它们的均值作为对于 $y(t+k)$ 的拟合结果. 通过这种方式, \mathbf{x} 和 \mathbf{y} 在错位 k 个样本点后的距离可按如下方式计算:

$$s_k = \sum_{t=1}^T \left[y(t+k) - \frac{1}{K} \sum_{i=1}^K y(t_i+k) \right]^2. \quad (7)$$

在此基础上, Stockmann 等^[25] 设计了一种归一化的非线性相似度指标

$$\rho = \left(1 + \sqrt{\frac{K s_k}{T(K+1)}} \right)^{-1}, \quad (8)$$

使得 ρ 最大的 k 被确定为时延. 这里信息传递关系的确定与前述方法类似, 不再赘述.

综上所述, 基于 KNN 的方法通过对不同变量间近邻点的拟合实现了相似度的度量. 然而, 与上述基于 CCF 和 DTW 的方法类似, 这种方法仅适用于双变量场景, 且对于非线性关系的拟合能力有限. 随着工业过程的大规模化与复杂化, 过程数据通常呈现高维度, 且变量间可能具有强非线性关系, 这为基于相似度的结构推断方法带来挑战. 实际上, 随着近年来人工智能技术的不断发展, 一些更为先进的数据分析手段可以用于 IRCD 中的相似度度量. 针对变量的高维特点, 可以采用考虑了协方差的马氏距离^[26] 进行相似度计算, 从而将高维数据的分布特点纳入分析. 此外, 针对强非线性关系, 可以利用核方法^[27-28] 将低维空间中的数据映射到高维核空间, 从而更加充分地刻画复杂的非线性特性. 类似地, 深度学习模型也可用于拟合强非线性关系. 值得注意的是, 尽管这些方法目前尚未被应用于基于相似度的结构推断中, 但不同

的结构推断方法的框架是类似的,区别仅体现在相似度指标选取的不同.因此,上述的马氏距离、核方法等均可替换现有的相似度度量手段,从而灵活地考虑过程数据的高维度、非线性特点.

1.2 基于因果的结构推断模型

尽管上述基于相似度的结构推断方法在IRCD中展示出一定的有效性,但这些模型存在一个共同的缺陷,即仅考虑成对关系.一个工业过程的典型设备上通常有几十到数百个测点,这种逐对变量分析的方法不仅效率低下,而且忽略了多变量间的相互作用,导致分析结果不准确.随着数据分析与挖掘理论的不断发展与推广,因果分析方法^[29-32]逐渐在各个领域崭露头角,也为工业过程故障的RCD带来了全新的研究思路.相比于相关性、相似性,因果关系是一种更为严格且精准的约束^[33-34],更加适用于描述动态系统中的信息传递过程^[35].在近5年内,因果分析方法逐步替代了基于相似度的方法,已成为IRCD任务中结构推断模型的主流.基于对因果关系的刻画方式不同,本文将基于因果的结构推断模型分为3类,包括基于预测、信息理论以及状态空间的方法.下面对这3类方法逐一介绍.

1.2.1 基于预测的方法

基于预测的方法主要指格兰杰因果(Granger causality, GC)^[13],该方法通过时序上的预测关系来揭示因果效应.具体而言,给定两个变量 x 和 y ,如果相比于仅采用 y 的过去信息来预测 y 的未来取值,同时引入 x 和 y 的过去观测能够显著提高对 y 的预测精度,则可认为 x 是 y 的格兰杰原因,而 y 是 x 的结果.换言之,在GC的框架下,因果可以由显著的预测贡献来定义.

具体而言,在GC方法中构建两个自回归(autoregression, AR)模型

$$y(t) = \sum_{i=1}^l a_{i,1}y(t-i) + \sum_{i=1}^l b_{i,1}x(t-i) + \varepsilon_{xy}, \quad (9)$$

$$y(t) = \sum_{i=1}^l a_{i,2}y(t-i) + \varepsilon_y. \quad (10)$$

其中: $a_{i,1}$ 、 $b_{i,1}$ 和 $a_{i,2}$ 均为回归系数, l 为时延, ε_{xy} 和 ε_y 为预测误差.

在上述两个预测模型中,一个使用了 x 和 y 的过去观测值,而另一个仅利用了 y 自身的过去信息.在GC中建立如下的因果度量指数来比较两个AR模型的预测误差大小,评估 x 对 y 的预测能力,以便确定 x 和 y 间是否存在因果关系:

$$\eta_{x \rightarrow y} = \ln \frac{\text{var}(\varepsilon_y)}{\text{var}(\varepsilon_{xy})}, \quad (11)$$

其中 var 表示方差.当该指数显著大于0时,可认为 x 是 y 的格兰杰原因.

上述的GC方法仅分析了双变量间的因果关系,为了考虑多元过程变量间的影响,条件格兰杰因果(conditional Granger causality, CGC)^[36-38]和向量自回归格兰杰因果(vector autoregression Granger causality, VAR-GC)^[39-41]被相继应用于IRCD. CGC的形式与传统的GC十分相似,对于 N 个变量,其中待分析的变量为 x 和 y ,其他变量记作 $\{z_1, \dots, z_{N-2}\}$ (称之为条件变量),CGC中同样构建两个自回归模型,即

$$y(t) = \sum_{i=1}^l a_{i,1}y(t-i) + \sum_{i=1}^l b_{i,1}x(t-i) + \sum_{j=1}^{N-2} \sum_{i=1}^l c_{ij,1}z_j(t-i) + \varepsilon_{xy|z}, \quad (12)$$

$$y(t) = \sum_{i=1}^l a_{i,2}y(t-i) + \sum_{j=1}^{N-2} \sum_{i=1}^l c_{ij,2}z_j(t-i) + \varepsilon_{y|z}. \quad (13)$$

其中: $c_{ij,1}$ 和 $c_{ij,2}$ 为CGC中条件变量的回归系数, $\varepsilon_{xy|z}$ 和 $\varepsilon_{y|z}$ 为条件变量引入后的预测误差.与GC类似,变量 x 与 y 之间的因果关系可通过对两个预测误差的方差进行分析来确定.

CGC虽然能够考虑多变量的影响,但对于每个条件变量都要引入其在若干个时间点上的采样,会显著提高回归模型中自变量的数目,甚至导致维数灾难^[42].VAR-GC通过建立如下式所示的向量自回归模型来考虑多变量间的影响,同时可以避免维度过高:

$$\begin{bmatrix} x_1(t) \\ \vdots \\ x_N(t) \end{bmatrix} = \sum_{i=1}^l \mathbf{A}_i \begin{bmatrix} x_1(t-i) \\ \vdots \\ x_N(t-i) \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{bmatrix}. \quad (14)$$

其中: \mathbf{A}_i 为第 i 个延迟上的系数矩阵, ε_i 为变量 x_i 的预测误差.

VAR-GC有两种经典的因果判别方式.第1种是时域上的,这种方法与传统GC类似,若要分析变量 x_i 是否为变量 x_j 的格兰杰原因,只需要在式(14)的基础上去除变量 x_i ,再通过对两个模型 x_j 中的预测误差来判别即可;另一种是基于频域的判别方法^[43],该方法通过对式(14)所示的VAR-GC模型进行 Z 变换,将时域模型转换到频域上,可以推导出转换后VAR-GC模型在频域上的传递函数为

$$\mathbf{H}(f) = \left(\mathbf{I}_N - \sum_{i=1}^l \mathbf{A}_i e^{-j2\pi i f} \right)^{-1}, \quad (15)$$

其中 \mathbf{I}_N 是维度为 $N \times N$ 的单位矩阵。

基于如上推导,变量 \mathbf{x}_i 对 \mathbf{x}_j 在频率 f 上的因果度量指数可按下式计算:

$$\eta_{\mathbf{x}_i \rightarrow \mathbf{x}_j} = \frac{H_{i,j}(f)}{\sqrt{\sum_{n=1}^N |H_{i,n}(f)|^2}}, \quad (16)$$

其中 $H_{i,j}(f)$ 表示 \mathbf{H} 中第 i 行第 j 列的元素。

近年来,GC 相关的方法被广泛应用于工业过程故障的 RCD 任务中。Yuan 等^[8,44] 提出了基于 GC 的 RCD 模型,并通过仿真系统和化工过程中的实验验证了 GC 模型的有效性。同时,该文指出对于振荡型异常,由于其周期性较强,基于频域的 VAR-GC 方法具有较好的性能。Landman 等^[45] 将频域的 VAR-GC 方法与过程变量间的连接关系相结合,实现了对因果关系的检测与振荡源的识别。对于非振荡型故障,时域的 GC 方法应用更为普遍。Chen 等^[46] 将 CGC 方法应用于 RCD 任务,并通过田纳西-伊斯曼过程 (Tennessee Eastman process, TEP) 基准数据集^[47] 中的阶跃故障案例验证了该方法的有效性。He 等^[16] 通过对比故障前后变量间格兰杰因果关系的变化定位异常因果,从而识别故障根因变量。该方法在后续研究中被应用于热轧带钢头部拉窄故障的 RCD 中^[48]。此外,GC 方法也与典型的特征提取方法相结合来提升诊断性能^[49-52]。

尽管 GC 方法已在工业过程故障的 RCD 中有诸多研究与应用,然而 GC 模型本质上基于线性的 AR 或 VAR 模型,无法有效应用于非线性过程。此外,GC 方法假设过程是平稳的,即过程数据的分布不随时间而改变,这一假设由于过程工况的不断变化可能难以满足。这些挑战限制了 GC 方法的应用场景,同时也促使了对 GC 方法的改进。本文将于第 2 节介绍 GC 的改进模型,并梳理各模型的适用场景。

1.2.2 基于信息论的结构推断模型

基于信息论的结构推断方法主要指转移熵 (transfer entropy, TE)^[14]。该模型与 GC 类似,通过度量时序上变量间的影响来分析因果关系。区别在于,该方法不再通过预测误差的变化来进行因果,而是通过信息量的变化。信息量可以通过熵值来度量,对于任一随机变量 \mathbf{x} ,若其在 t 时刻具有 Q 个潜在状态,且处于第 i 个状态上的概率为 p_i ,则其熵函数 H 取值可以按下式计算:

$$H[\mathbf{x}(t)] = - \sum_{i=1}^Q p_i \log p_i. \quad (17)$$

由式(17)可知:分布越散乱,即不确定度越强,熵值越大;而当变量所处状态确定时,即 \mathbf{x}_t 在某个状态上的取值无限接近于 1 时,熵值趋近于 0。

在 TE 模型中,对于两个变量 \mathbf{x} 和 \mathbf{y} ,如果 \mathbf{x} 的过去信息的引入能够显著减少 \mathbf{y} 未来信息的不确定性,则认为 \mathbf{x} 的引入对 \mathbf{y} 的未来状态产生了显著的影响,即 \mathbf{x} 导致了 \mathbf{y} 。具体地,TE 的计算方式为

$$\begin{aligned} \text{TE}_{\mathbf{x} \rightarrow \mathbf{y}} = & \sum_{\mathbf{y}_{t+1}, \mathbf{x}_t, \mathbf{y}_t} p(\mathbf{y}(t+h), \mathbf{x}_t, \mathbf{y}_t) \log \frac{p(\mathbf{y}(t+h)|\mathbf{x}_t, \mathbf{y}_t)}{p(\mathbf{y}(t+h)|\mathbf{y}_t)} = \\ & H(\mathbf{y}(t+h)|\mathbf{y}_t) - H(\mathbf{y}(t+h)|\mathbf{x}_t, \mathbf{y}_t). \end{aligned} \quad (18)$$

其中: \mathbf{x}_t 和 \mathbf{y}_t 分别为变量 \mathbf{x} 和 \mathbf{y} 从 $t-l+1$ 时刻到 t 时刻的观测值所构成的向量, h 为时间步长, $p(\cdot|\cdot)$ 表示条件概率分布的概率密度函数。

由上述定义可知,当 $\text{TE}_{\mathbf{x} \rightarrow \mathbf{y}}$ 大于 0 时,可认为 \mathbf{x} 向 \mathbf{y} 传递了信息,且 $\text{TE}_{\mathbf{x} \rightarrow \mathbf{y}}$ 值越大,说明信息传递关系越强。此外,为了简化因果结构,变量间的因果关系可通过不同信息传递方向上 TE 的差值来度量,即

$$\eta_{\mathbf{x} \rightarrow \mathbf{y}} = \text{TE}_{\mathbf{x} \rightarrow \mathbf{y}} - \text{TE}_{\mathbf{y} \rightarrow \mathbf{x}}. \quad (19)$$

当该值显著大于 0 时,可认为 \mathbf{x} 导致了 \mathbf{y} 。

TE 的运算是基于概率密度的,因此其可以处理非线性的因果关系。此外, Barnett 等^[53] 证明了当变量服从高斯分布时,TE 与传统的 GC 方法完全等价,从而建立起两种经典因果分析方法间的联系。目前,TE 在 IRCD 任务中已经有了一定的应用。Bauer 等^[9,54] 率先将 TE 方法用于故障传播路径的分析,并给出了 TE 方法中变量概率函数的估计方法,即因果显著性检验策略,同时指出,TE 方法中对于概率密度的估计步骤会引入显著的运算开销。为降低 TE 的运算复杂度, Duan 等^[55] 提出了一种无需概率密度估计的转移熵方法,该方法无需对概率空间进行假设,从而提高了运算效率。Rashidi 等^[56] 设计了一种基于符号动态性的归一化转移熵,通过对原始数据的符号化来简化运算过程。

除了运算复杂度的挑战,TE 的参数设定也备受研究者的关注。传统 TE 模型中需要人为设定时延参数 l ,而不同的参数设定下可能会获得不同的因果分析结果。Bauer 等^[9] 指出在实际应用场景中变量间的时延通常是未知的,并建议采取不同时延下变量转移熵的平均值进行分析,然而,这种平均操作可能使得真正的因果关系被某些时延设定不合适条件下所

获取的分析结果所影响. Naghoosi等^[57]设计了一种基于互信息的方法来确定变量间的时延关系,并指导TE中时延参数的选取. Lindner等^[58]对过程动态性和TE最佳参数之间的关联进行深入分析,给出了TE的参数设定规则,并总结出了基于TE的工业过程故障RCD的系统流程. 除了对于 l 参数的讨论, Zope等^[59]同样对时间步长参数 h 的选取进行了分析.

除了上述的运算效率与时延参数设定问题,与GC方法类似,如何考虑多变量间的影响也是TE研究的重要方向. 偏转移熵(partial transfer entropy, PTE)^[60]是一种最为经典的多变量TE模型. 在给定其他变量 Z 的条件下, PTE按下式计算变量之间的信息传递关系:

$$\text{PTE}_{x \rightarrow y} = \sum_{y_{t+h}, \mathbf{x}_t, \mathbf{y}_t, \mathbf{Z}_t} p(y(t+h), \mathbf{x}_t, \mathbf{y}_t, \mathbf{Z}_t) \times \log \frac{p(y(t+h) | \mathbf{x}_t, \mathbf{y}_t, \mathbf{Z}_t)}{p(y(t+h) | \mathbf{y}_t, \mathbf{Z}_t)}. \quad (20)$$

可见, PTE方法在传统TE的基础上引入了其他变量的影响. 然而, Duan等^[61]指出, PTE通常将除待分析变量之外的所有变量作为 Z 引入到模型中, 这将带来巨大的运算开销. 为解决该问题, 他们提出了一种直接转移熵(direct transfer entropy, DTE)方法. DTE无需将所有变量引入到TE模型中, 而是在TE所获取的因果结构的基础上寻找存在间接因果路径的变量组, 并将中间变量引入到TE中进行分析以判断二者是否存在直接因果关系, 从而大大简化了运算流程. 目前, DTE已被应用于工业过程故障的RCD中^[62-63].

综上, TE及其改进方法在RCD领域得到了一定的应用与发展. 不仅如此, 基于TE的RCD的方法也逐渐与过程监测模型结合起来以构成完整的故障检测与诊断系统^[64-65]. 相比于前述的GC模型, TE对于非线性过程数据具有更好的处理能力, 但同时对于数据也提出了更高的要求并引入了更高的运算复杂度^[66]. Bauer等^[54]对TE模型所需的样本数进行了探究, 发现TE一般需要400~2000个样本才能够表现出相对好的性能. 对于某些实时性要求较强的故障诊断场景, 这样的数据需求可能是难以满足的. 此外, 与GC方法类似, TE模型也要求过程数据是平稳的, 通常无法应用于非平稳过程. 本文将在第2节对此进行展开介绍.

1.2.3 基于状态空间的结构推断模型

基于状态空间的方法对于具有强非线性耦合的变量间的因果推断有着显著优势, 其将时间序列嵌入到高维空间中, 通过分析不同时间序列在高维空间中的非线性重构关系^[67]来确定因果关系是否

存在. 在IRCD领域, 收敛交叉映射(convergent cross mapping, CCM)^[15]是应用最为广泛的一种基于状态空间的结构推断模型. CCM的核心思想是: 如果变量 x 是 y 的原因, 则 y 对应的状态空间中包含 x 的演化信息, 能够通过高维流形的重构关系来反映. 具体地, 建立 x 和 y 的状态空间如下:

$$\mathbf{x}_s(t) = [x(t), x(t-\tau), \dots, x(t-(m-1)\tau)], \quad (21)$$

$$\mathbf{y}_s(t) = [y(t), y(t-\tau), \dots, y(t-(m-1)\tau)]. \quad (22)$$

其中: τ 为嵌入时延, m 为嵌入维度. 两个状态空间分别构成了系统 X 和 Y . CCM在系统 X 中对于每个样本点寻找 $\mathbf{x}_s(i)$ 的 m 个近邻点, 按近邻程度顺序排列, 获得 $\{\mathbf{x}_s(i, j)\} = \{\mathbf{x}_s(i, 1), \dots, \mathbf{x}_s(i, m)\}$, 将其映射到系统 Y 中, 对应的点定义为 $\{\mathbf{y}_s(i, j)\}$. 基于此可以计算 $\hat{\mathbf{y}}_s(i)$ 的估计值为

$$\hat{\mathbf{y}}_s(i) = \sum_{j=1}^m w_j \mathbf{y}_s(i, j). \quad (23)$$

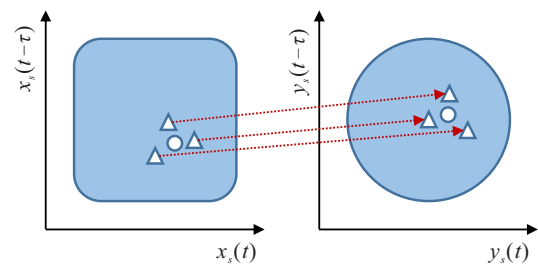
其中

$$u_j = \exp\left(-\frac{\|\mathbf{x}_s(i) - \mathbf{x}_s(i, j)\|_2}{\|\mathbf{x}_s(i) - \mathbf{x}_s(i, 1)\|_2}\right), \quad (24)$$

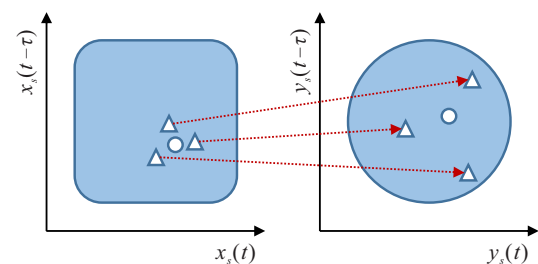
$$w_j = u_j / \sum_{k=1}^m u_k. \quad (25)$$

定义 $\hat{\mathbf{y}}_s(t)$ 为 $\mathbf{y}_s(t)$ 从 X 到 Y 的交叉映射, 计算 $\hat{\mathbf{y}}_s(t)$ 与 $\mathbf{y}_s(t)$ 的皮尔逊相关系数 ρ , 若 ρ 取值在0~1之间, 则说明二者在高维空间中的演化信息正相关, 即存在 y 到 x 的因果关系.

图3展示了CCM方法的核心概念. 其中: 圆形点代表系统中的样本点, 三角形点代表近邻点, 虚线箭头代表系统间的映射关系. 图3(a)中展示了存在 y 到



(a) 存在 y 到 x 的因果关系



(b) 无因果关系

图3 CCM方法的因果推断原理示意

\mathbf{x} 的因果关系的情况,可以发现,对于系统 X 中的一个样本,其近邻点在映射后仍收敛于其在系统 Y 中所对应的点,这说明系统 \mathbf{x} 和 \mathbf{y} 存在相似的流形结构;图 3(b) 则展示了因果关系不存在的情况,可以看出近邻点在映射后发散.

CCM 方法于 2012 年被提出,近年来由于其其在处理强非线性数据上的优势逐渐被应用于工业过程故障的 RCD 中. Luo 等^[68] 为 CCM 提出了一种快速映射与重构方法,从而加速了 CCM 的因果推断效率. Wang 等^[69] 设计了一种改进的 CCM 方法,该方法基于伪最近邻理论确定 CCM 的最佳嵌入维数,并给出了一种时延检测函数来去除虚假因果关系. Tian 等^[70-71] 提出了一种单尾显著性指标来确定 CCM 中的时延参数及因果关系可靠性. 然而,由于 CCM 方法提出较晚,且主要针对双变量场景,在 IRCD 中仍存在较大的改进空间.

2 结构推断改进方法及其适用场景

本节针对过程数据的高维性、非线性以及非平稳性 3 种特性,结合过程机理知识的效用,介绍结构推断模型的改进方法. 值得注意的是,在如上介绍的 4 类结构推断模型中,非因果的方法在近年来的研究中逐渐被基于因果的方法所替代,而基于状态空间的方法由于提出时间较晚,尚未在 IRCD 任务中广泛应用,因此本节主要关注基于预测及信息论的结构推断模型,即 GC 和 TE 方法的改进.

2.1 针对高维过程数据的结构推断改进方法

通常情况下,一个工业过程的测点在几十个到数百个之间,在故障发生后,由于故障信息的大范围传播,可能会引入较多的故障变量^[72-73]. 上述介绍的多数结构推断模型需要对变量进行成对分析,这将会引入显著的计算开销. 具体而言,对于一个具有 N 个变量的过程数据集,成对的推断方法需要执行 $(N^2 - N)$ 次推断操作. 为减少运算次数,更高效地处理高维数据,一些改进的结构推断模型被相继提出. 由于 GC 方法基于预测模型来推断因果关系,而单变量的预测模型很容易被灵活地推广到多变量场景,GC 方法在处理高维数据上具有显著优势. 此处介绍两种 GC 的高维改进模型,包括套索格兰杰 (lasso Granger, LG)^[74] 和稀疏邻接预测 (sparse adjacency forecasting, SAF)^[75].

LG 在多变量预测模型中引入了一种 L_1 范数稀疏化约束^[76] 来选择显著的预测关系,从而表征因果性. 对于变量 $\mathbf{x}_i (i = 1, \dots, N)$, LG 中构建如下的预测模型:

$$\min_{w_{j,i,p}} \left\| x_i(t) - \sum_{j=1}^N \sum_{p=1}^l w_{j,i,p} x_j(t-p) \right\|_2^2 + \lambda \sum_{i=1}^N \sum_{p=1}^l |w_{j,i,p}|. \quad (26)$$

其中: $w_{j,i,p}$ 为 $x_j(t-p)$ 对 $x_i(t)$ 的回归系数, λ 为 L_1 稀疏化约束的系数.

L_1 范数约束具有稀疏变量选择的功能,图 4 中给出其原理的几何解释,展示了一个双变量回归问题的几何表示. 其中: w_1 和 w_2 为回归系数,蓝色矩形表示 L_1 范数约束,而红色椭圆则表示回归误差. 可见当同时对回归误差和 L_1 范数约束进行最小化时,最终的解倾向收敛于坐标轴,即两个回归系数中的其中一个被稀疏化,仅有对回归最为关键的变量能够具有非零的回归系数. 为了进一步说明 L_1 范数的独特功效,图 4 中给出了同样常用的 L_2 范数正则化约束的几何解释. 可发现 L_2 范数在二维坐标平面中表现为圆形,因此与回归误差难以相交于坐标轴,从而无法产生稀疏解.

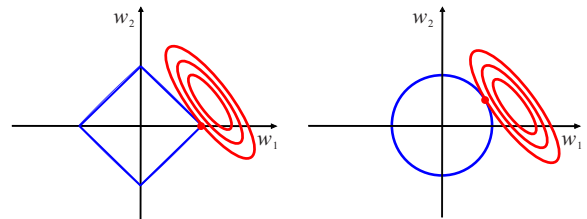


图 4 L_1 以及 L_2 范数约束的几何解释

LG 方法中利用了 L_1 范数的稀疏变量选择性质,由于 GC 可以解释为显著的预测关系,如果 LG 模型的优化结果中,对于任意 p , $w_{j,i,p}$ 取值均为 0,则说明变量 \mathbf{x}_j 对 \mathbf{x}_i 没有显著的预测贡献,因此其对应系数全部被稀疏化,从而 \mathbf{x}_j 对 \mathbf{x}_i 不存在因果关系. 反之,若存在 p 使得回归系数 $w_{j,i,p}$ 非零,则 \mathbf{x}_j 是 \mathbf{x}_i 的格兰杰原因.

LG 方法通过引入 L_1 范数项避免了成对建模,而仅需对每个变量建模一次即可,显著提高效率的同时考虑了多变量的影响. 尽管如此, LG 仍需要重复地建模. 为进一步简化建模流程, Song 等^[75] 提出了 SAF 方法,该模型设计了多变量间的邻接矩阵结构,实现了多变量的一次性因果推断. 该方法的优化目标为

$$\min_{\mathbf{A}, \mathbf{W}} \left\| \mathbf{x}(t) - \text{diag}(\mathbf{A} \mathbf{X}_l \mathbf{W}) \right\|_2^2 + \lambda \sum_{i=1}^N \sum_{j=1}^N |a_{ij}|. \quad (27)$$

其中: $\mathbf{x}(t) = [x_1(t), \dots, x_N(t)]^T$; \mathbf{X}_l 为时延矩阵,计算方式为 $\mathbf{X}_l = [\mathbf{x}(t-1), \dots, \mathbf{x}(t-l)]$; \mathbf{A} 为待求的邻接矩阵,其第 i 行第 j 列元素记作 a_{ij} ; \mathbf{W} 为待求的

权重矩阵,且 $\mathbf{W} \in \mathbb{R}^{l \times N}$; $\text{diag}(\cdot)$ 表示取矩阵的对角线元素. 为便于理解,可将式(27)的矩阵运算改写为变量级别的形式

$$\min_{a_{ij}, w_{pi}} \left\| x_i(t) - \sum_{j=1}^N a_{ij} \sum_{p=1}^l w_{pi} x_j(t-p) \right\|_2^2 + \lambda \sum_{j=1}^N |a_{ij}|, \quad (28)$$

其中 w_{pi} 为矩阵 \mathbf{W} 第 p 行第 i 列的元素. 可见 SAF 的形式与 LG 极为类似,区别在于其使用邻接矩阵 \mathbf{A} 统一表示所有变量间的预测关系,变量 \mathbf{x}_j 对 \mathbf{x}_i 的预测关系可用元素来描述. 由于 SAF 可以通过矩阵形式直接描述多变量的预测过程,无需重复为每个变量建模,进一步提高了推断效率.

尽管 L_1 范数约束在多变量格兰杰因果推断中获得了广泛的应用,但这种方法通常对系数 λ 较为敏感. 当 λ 取值过小时,会导致稀疏化力度不足,产生冗余因果;当 λ 取值过大时,会导致推断结果过于稀疏,丢失了重要的因果关系. 然而,与有监督学习不同,因果推断任务不具备标签,无法通过交叉验证来确定使得因果提取性能最佳的稀疏化系数,因此,现有方法^[74]往往以预测误差作为交叉验证的指标以进行参数设定,可能无法客观、真实地反映因果推理的合理性. 到目前为止,该问题尚未得到充分解决.

2.2 针对非线性过程数据的结构推断改进方法

由于复杂的过程机理和运行条件,过程数据中往往存在非线性关系,无法使用线性模型进行拟合. 在如上介绍的结构推断模型中,TE、CCM 等方法能够

有效处理非线性特性,而 GC 则为线性方法. 考虑到 GC 结构灵活、易于拓展到高维场景的优势,研究者们也开始逐渐将目光转向 GC 的非线性拓展. GC 方法的本质是时间序列预测模型,在传统 GC 中使用线性的 AR 模型实现预测,而非线性的 GC 方法通常将 AR 替换为其他非线性的回归模型.

Chen 等^[77]使用高斯过程回归(Gaussian process regression, GPR)^[78]进行时序预测,该方法将 CGC 中的 AR 模型替换为 GPR 以提取非线性、多变量的因果性. 由于 GPR 本身运算复杂度较高,且 CGC 需要以成对的形式推断,该方法在因果推断之前先使用非线性检验来筛选非线性变量,仅在与这些变量相关的推断中使用 GPR 以降低运算量.

为提高时序预测的准确性并减少重复建模次数,Chen 等^[79-80]提出了一种稀疏因果残差神经网络(sparse causal residual neural network, SCRNN)模型,该方法可看作是 LG 模型的非线性拓展. SCRNN 的结构如图 5 所示,该模型中为每个变量建立一个预测模型. 首先使用层次化分组套索(hierarchical group lasso, HG-Lasso)^[81]对变量和时延进行稀疏约束,再利用多层神经网络对变量进行非线性预测,同时引入线性残差连接以促进网络收敛. 通过最小化预测误差,SCRNN 能够提取变量间的线性与非线性因果关系,并为每个变量自动设定合适的时延参数. 然而,该方法依旧需要为每个变量建立一个神经网络模型,运算复杂度较高.

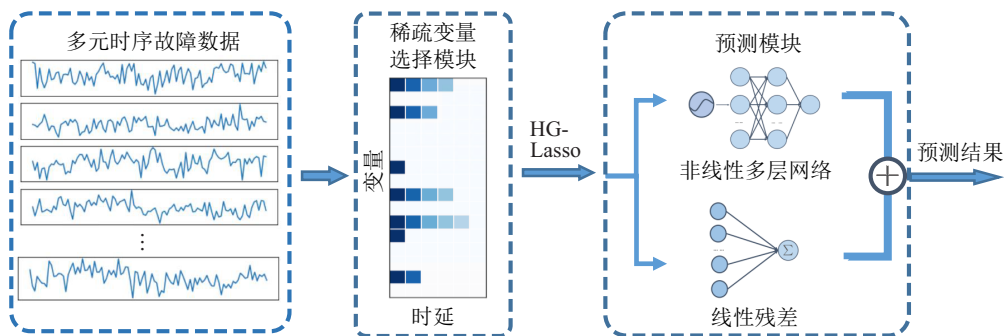


图 5 SCRNN 模型结构

为了避免重复建模, Song 等^[17]提出了一种多阶预测图提取(multi-level predictive graph extraction, MPGE)模型, MPGE 的结构如图 6 所示. 该模型首先将输入的多元过程数据输入二维卷积神经网络(convolutional neural network, CNN), 从而实现非线性特征提取. 值得注意的是, 此处的 CNN 在变量维度的核尺寸设置为 1, 从而避免了特征提取阶段不同变量

间的交互. 随后, 各变量的特征在邻接矩阵的指导下进行多次信息融合, 最终得到预测结果, 此处邻接矩阵的作用与 SAF 模型中的类似. 通过在最小化预测误差的同时引入对邻接矩阵的逐层稀疏化操作, 能够提取到变量间直接或间接的显著预测关系, 从而表征非线性 GC.

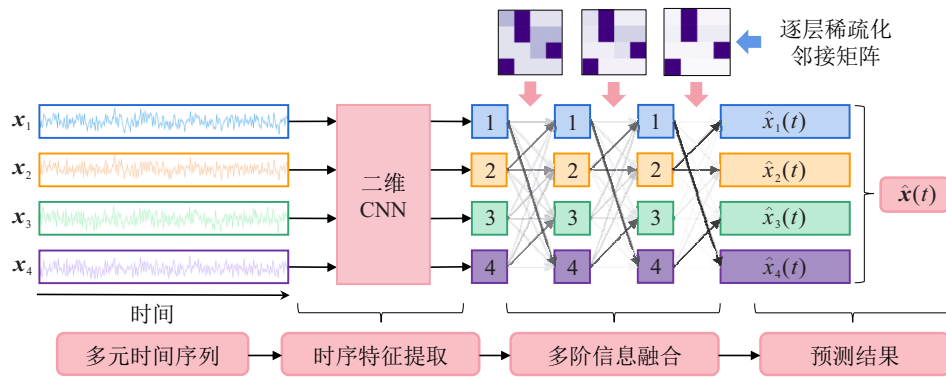


图6 MPGE模型结构(以4个过程变量为例)

随着近年来深度学习的不断发展,越来越多的先进但较为复杂的时序特征提取网络结构被提出并应用于非线性GC提取与RCD任务中,这些方法本质上修改了时序预测的模型结构,使其能够适应非线性更强、时序关系更为复杂的过程数据^[82-86]。值得注意的是,学习能力越强的神经网络结构往往具有越多的待训练参数,从而需要更充足的故障数据进行模型训练。然而,工业过程中的故障数据在多数场景下难以获取,持续的故障往往会导致设备受损、生产暂停,因此需要尽快采取检修维护工作,这也大幅缩短了可供采集故障数据的时间。因此,基于复杂神经网络的非线性GC方法虽然理论上正确,但在实际应用中存在着诸多限制。如何将基于复杂神经网络的非线性GC模型应用到实际场景中,仍然是亟待解决的问题。

2.3 针对非平稳过程数据的结构推断改进方法

在结构推断过程中,过程数据的非平稳特性挑战也引发了学者们的广泛关注。非平稳是指过程数据的分布随时间不断变化,数学上可表现为均值和协方差的变化^[87-88],这将会导致过程变量间的时序依赖关系不断改变,难以用统一的模型来描述。传统的基于AR以及VAR方法的GC模型,以及上文展示的TE、PTE和DTE模型,均要求数据是平稳的,将它们应用于非平稳数据时会引入虚假的因果关系。目前,GC和TE模型均被改进以适应过程的非平稳特性。

对于GC方法,最典型的非平稳改进模型包括向量误差纠正(vector error correction, VEC)^[89-90]模型和时变参数方法^[91]。VEC方法假设非平稳时间序列在一次差分后能够转化为平稳时间序列,并同时使用原始数据和差分数据建立预测模型来考虑变量间的长期和短期因果关系。Zhang等^[90]利用VEC模型提取过程变量间的非平稳格兰杰因果关系,并将其应用于热轧过程的RCD。时变参数方法中模型参数不是固定的,而是随着新的过程数据而不断自适

应更新,从而捕捉数据的变化特性。Song等^[91]对前述的SAF方法进行改进,提出了一种稀疏时变预测关系提取(sparse and time-varying predictive relation extraction, STPRE)方法,通过引入次梯度和泰勒展开技巧,实现模型中邻接矩阵参数的不断更新,并求解了更新过程的闭式解,保证了因果推断过程的可靠性。

对于TE方法,对其进行非平稳扩展的主要手段是符号化。Staniek等^[92]提出了一种符号转移熵(symbolic transfer entropy, STE)方法,该方法依照原始时间序列在每个时间窗口内各采样点相对大小的排列顺序对原始序列赋予符号,以表示短期的时序变化模式,从而去除非平稳趋势。通过对不同变量的符号序列之间计算转移熵,能够获取变量间的短期因果依赖性,避免虚假因果。STE同样被拓展到多变量场景,类似于将TE改进为PTE的过程,Kugiumtzis^[93]在STE中引入多变量的信息,从而将其改进为偏符号转移熵(partial symbolic transfer entropy, PSTE)。目前,STE和PSTE已被应用于IRCD中。Wen等^[94]将STE方法引入IRCD中,并将其与多元统计监测理论中的控制图模型相结合,基于控制图的信息实现符号化,从而提高模型鲁棒性。Duan等^[95]提出了一种多尺度的PSTE方法,通过比较不同时延参数下PSTE的取值大小确定因果信息传递的时延及方向,并设计一种因果判据以确定因果关系的显著性。值得注意的是,STE和PSTE的符号化操作能够在一定程度上对数据进行简化,因此能够提高运算效率,然而,这种操作同样会不可避免地造成信息损失,且无法提取长期因果关系,因此,仍具有一定的改进空间。

2.4 引入知识的结构推断改进方法

上述介绍的结构推断方法均为纯数据驱动的,然而在实际的工业场景中,所获得故障传递结构的可解释性也极为重要。如果数据分析的结果在过程机理

上难以解释或明显与机理相悖,则会使操作人员感到困惑甚至直接放弃对RCD模型的信任.为了解决上述问题,部分学者主张利用过程机理知识对结构推断模型进行约束或者修正,从而避免与机理知识相悖的推断结果.目前用于结构推断模型的知识主要包含两种:测点连接关系和测点分布信息.

基于测点连接关系的方法根据工艺流程图或管道仪表图等知识提前人为总结出不同测点间是否存在物质、能量或信息层面的直接交互,从而确定两个过程变量之间是否可能存在因果.这种方法认为,如果两个变量在实际过程中没有直接的交互或影响关系,则二者不可能具有因果等信息传递关系.因此,在结构推断过程中可以直接略去这些没有交互关系的变量之间的分析,一方面能够显著减少计算量,另一方面能够保持与机理的一致性.目前这种思想理念已经应用于KNN^[96]、GC^[45]、TE^[62]方法以进行推断结果的约束与改进.

基于测点分布信息的方法的思想是:一个工业过程中往往存在多个设备,尽管过程故障影响范围可能较广,但其故障传播路径也应遵循测点的分布关系,即故障信息先在根因变量所处的设备内部进行传播,再通过设备之间的连接传递到其他设备上.基于这一认识,不难发现,同一设备上的测点间应存在较强的信息传递关系,而来自不同设备的测点间的信息传递关系应相对较弱.因此,如图7所示,过程测点应以设备为单位进行分组,通过组内、组间两个层级进行结构推断和根因识别.这种推断方式不仅能够增强推断过程和过程机理的吻合度与可解释性,还能够显著减少待推断的关系,提高诊断效率.目前,这种组别关系知识的引入已在IRCD领域展现出有效性.Zhu等^[97]提出了分块TE模型,该方法利用过程知识将不同变量分配到多个子模块中,并在每个子模块中利用TE模型进行结构推断.在此基础上,Luo等^[98]

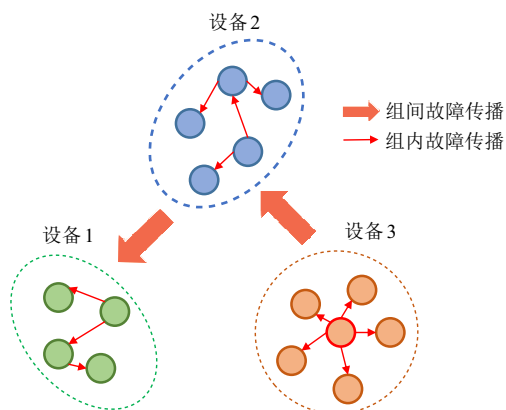


图7 引入测点分布知识的分组结构推断框架

通过过程机理知识建立了不同子模块的中间变量,打通了不同模块间的因果传递路径,并利用DTE方法进行结构推断.此外,分组推断的有效性也在GC的建模框架中得到了验证^[99-100].另外,考虑到过程结构复杂难以分组的问题,部分研究主张利用数据驱动的方法进行聚类分组^[52],然而,这种方法往往基于变量相似度等指标进行聚类,可能与机理并不一致.

上述研究表明,融入了知识的结构推断方法能够给出与过程机理更加吻合、可解释性更强的信息传递结构,然而,由于过程的复杂性和知识质量的限制,相关研究并不多见,纯数据驱动的方法仍旧是领域内的主体研究方向.

2.5 结构推断方法的性能对比分析

综上,本文梳理了不同的结构推断方法及其适用场景.表1中呈现了不同模型的性能对比以及相应的工业应用领域.值得说明的是,由于大量已有方法并未给出定量的运算复杂度分析,本文参考已有IRCD综述中的定性评价方式^[21],将运算效率分为低、中、高三档,分别对应计算时间极快的线性模型(如CCF和GC等)、需要概率估计或收敛性分析等操作的非线性模型(如TE和CCM等)、建模过程中多轮反复迭代且需要GPU等运算加速器参与的方法(如基于神经网络的方法).此处STE方法虽然属于TE类别,但运算效率较高,这是因为其符号化操作对原始数据进行了简化,显著提高了运算效率.运算效率是从单次建模的角度进行评估.为度量多变量建模的性能,此处引入建模次数,即模型针对多变量场景下(N 个变量)需要的反复建模次数.通过这两个指标综合评价计算量大小.不仅如此,本文还引入了不确定性评价指标,衡量建模过程中模型的结构推断结果受超参数的不同设定、不同的初始化等因素的影响程度.同样定量地分为低、中、高三档.具体而言:CCF、DTW、KNN等不存在或仅有少量超参数的方法被认定为不确定性低;线性的GC类方法和CCM对时延、稀疏化系数等超参数较为敏感,因此被认定为具有中等不确定性;TE类方法对时延等超参数极为敏感,不同的设定可能会给出截然不同的推断结果^[21],且基于神经网络的方法具有较多超参数,训练过程受随机性影响显著,因此被归类为高不确定性方法.需要说明的是,引入知识的结构推断方法并未对模型本身进行修改,而是基于知识避免不合理的推断结构,因此在表1中没有单独列出分析.

对表1中信息进行分析,可以发现,尽管现有方法已在化工、能源、石油、造纸等诸多工业领域有所

表1 结构推断模型的特点及适用范围比较

类型	模型	多变量	高维度	非线性	非平稳	不确定性	运算效率	建模次数	已有工业应用领域
非因果方法	CCF ^[23-24]	×	×	×	×	低	高	$N^2 - N$	化工过程 ^[23] 、造纸过程 ^[24]
	DTW ^[7]	×	×	×	√	低	中	$N^2 - N$	化工过程 ^[7]
	KNN ^[25]	×	×	√	×	低	中	$N^2 - N$	化工过程 ^[25]
传统GC模型	GC ^[13]	×	×	×	×	中	高	$N^2 - N$	化工过程 ^[49,51] 、制药过程 ^[16] 、热轧钢过程 ^[48]
	CGC ^[36-38]	√	×	×	×	中	高	$N^2 - N$	化工过程 ^[46]
	VAR-GC ^[31-33]	√	×	×	×	中	高	$N^2 - N$	化工过程 ^[8,44,50] 、造纸过程 ^[45]
稀疏GC模型	LG ^[74]	√	√	×	×	中	高	N	暂无工业应用
	SAF ^[75]	√	√	×	×	中	高	1	化工过程 ^[75]
非线性GC模型	基于GPR的GC ^[77]	√	√	×	×	高	中	$N^2 - N$	化工过程 ^[77]
	SCRNN ^[79-80]	√	×	√	×	高	低	N	化工过程 ^[79] 、卷烟生产过程 ^[80]
	MPGE ^[17]	√	√	√	×	高	低	1	化工过程、卷烟生产过程 ^[17]
	基于复杂神经网络结构的GC ^[82-86]	√	×	√	×	高	低	N	化工过程 ^[83,85,86] 、热轧钢过程 ^[84] 、炼铁过程 ^[82]
非平稳GC模型	VEC ^[89-90]	×	×	√	×	中	高	$N^2 - N$	热轧钢过程 ^[90]
	STPRE ^[91]	√	×	√	×	中	高	1	火力发电过程 ^[91]
传统TE模型	TE ^[14]	×	×	√	×	高	中	$N^2 - N$	化工过程 ^[9,54,58-59,65,98] 、石油工业 ^[57] 、选矿过程 ^[59] 、热轧钢过程 ^[64] 、火电过程 ^[59]
	PTE ^[60]	√	×	√	×	高	中	$N^2 - N$	暂无工业应用
	DTE ^[61-63]	√	×	√	×	高	中	$N^2 - N$	化工过程 ^[61] 、造纸过程 ^[62] 、热轧钢过程 ^[63]
符号化TE模型	STE ^[92,94]	×	×	√	√	高	高	$N^2 - N$	化工过程 ^[94]
	PSTE ^[93,95]	√	×	√	√	高	中	$N^2 - N$	火电过程 ^[95]
	CCM ^[115]	×	×	√	×	中	中	$N^2 - N$	化工过程 ^[68-69] 、火力发电过程 ^[70-71]

应用,但目前尚无方法能够同时处理多种复杂的过程特性.另外,总结表1中信息可知,目前TE和GC相关的改进方法在工业过程故障的RCD任务中具有较为显著的优势,GC通常结构较为灵活,易于拓展到高维非线性数据中,而TE及其改进方法复杂度通常较高,一般适用于变量较少的场景.

3 过程故障根因识别范式

以上章节介绍了过程变量间的结构推断方法,这些方法能够揭示变量间的信息传递关系,从而描述故障传播的路径,为根因变量的识别奠定了基础.本节对RCD中的根因识别方法进行回顾,并将现有根因识别方法凝练为两大范式:纯数据驱动方法以及知识数据融合驱动方法.

3.1 纯数据驱动根因识别方法

纯数据驱动根因识别方法包含基于根节点的方法和基于分数的方法.这两类方法利用一定数目

的在线故障数据推断出故障变量之间的信息传递关系.此外,这些方法通常需要使用故障隔离^[101-102]作为前序步骤.具体而言,一个故障可能无法在短时间内波及所有设备和回路,因此需要筛选出受故障显著影响的过程变量.如果直接对全部变量进行推断,则某些尚未受到故障传播影响的变量间的因果关系将会影响RCD的性能^[16].

3.1.1 基于根节点的根因识别方法

基于根节点的根因识别是目前应用最为广泛的根因识别方法.对于 N 个故障变量,利用上文介绍的结构推断方法可以确定每一对变量间的信息传递关系,从而形成一个有向图结构,图中各节点表示不同的变量,节点间的连边及其方向用于描述故障传播路径的存在性和方向.可以用一个维度为 $N \times N$ 的邻接矩阵来描述变量间的信息传递关系,与图结构相对应,若其第 i 行第 j 列元素取值为1,则表示信息由变量 x_j 传递到 x_i ;若取值为0,则说明二者之间不存在

故障传播路径. 通过上述图结构及其邻接矩阵, 能够找出故障传播的根节点以实现根因变量的识别. 具体地, 若在因果图中某个变量向其他变量传递了信息, 而没有任何其他变量对该变量传递信息, 则该变量为图中的一个根节点, 从而识别为根因变量. 同理, 若某变量在邻接矩阵中所对应的行向量中各元素全部为0, 而其对应的列向量不全为0, 则说明该变量为根因变量. 在如上的根因识别过程中并没有利用或引入额外的过程机理或专家知识, 而是完全基于结构推断步骤所获取的图数据, 根据对于数据驱动以及知识驱动方法的一般理解^[103-104], 将其归类于纯数据驱动方法. 图8给出了一个信息传递图结构及其对应邻接矩阵的示意图. 图8中给定了5个变量 $\{a, b, c, d, e\}$ 及这些变量间的信息传递路径, 白框代表不存在路径. 由图结构可以确定节点 a 为故障传播路径的根节点, 因此变量 a 为根因变量. 在邻接矩阵中可以发现, 变量 a 所对应的行取值均为0, 说明没有任何其他变量将信息传递到 a , 同样体现了 a 为根因变量.

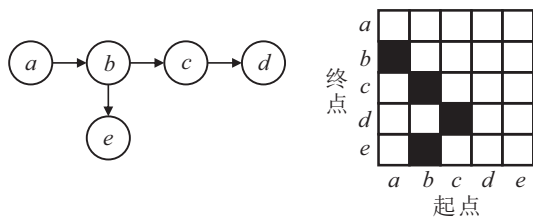


图8 信息传递图结构及邻接矩阵示意

目前, 大多数 IRCD 的研究工作中均采用基于根节点的范式实现根因识别. 然而, 在结构推断步骤中通常关注于每对变量间的信息传递关系, 而并不考虑所有变量共同形成的传播路径, 因此可能会出现双向或环形信息传递结构, 难以给出确定的根因识别结果^[17]. 图9给出了一个无法识别根因的信息传递结构, 其中红色箭头表示导致根因变量无法识别的信息传递关系, 该关系导致了变量存在环形结构 $\{a, b, e\}$, 相应地, 邻接矩阵中并不存在某个变量所对应的行向量元素均为0.

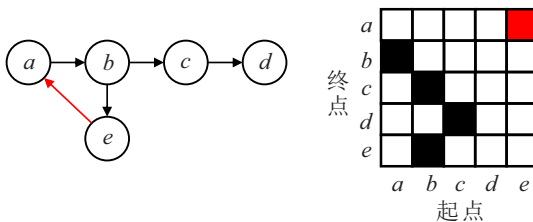


图9 根因无法识别案例的信息传递及邻接矩阵示意

事实上, 由于过程中广泛存在的控制回路, 双向及环形的信息传递结构并不少见, 如果根因变量恰巧处于这样的结构中, 则故障传播路径中不存在根节

点, 无法给出根因变量. 为解决根因无法识别的问题, 已有学者试图对信息传递结构进行简化. He等^[16]提出了一种因果关系对比的方法用于 IRCD, 该方法中并不直接使用故障变量间的因果关系来识别根因变量, 而是通过对比正常与故障工况下因果关系的变化来确定由于故障传播而触发的信息传播路径. 这种方法能够显著地简化信息传递的图结构, 一定程度上避免了双向和环形因果关系的影响. Liu等^[18]提出了一种基于最大生成树的信息传递结构的简化模型, 并在基于 CGC 的因果推断框架下予以实现, 该方法不再以二元的有或无来描述变量间的因果关系, 而是以式(11)中所示的统计量定量化描述因果的显著性, 从而将信息传递图结构转变为带权图. 通过求解该图结构的最大生成树, 能够保留最重要的信息传递路径而去除一些显著性不高的因果关系, 从而避免双向和环形结构, 同时将所获取的树结构的根节点作为根因变量. 然而, 上述的简化方法可能会去除部分重要的因果关系, 从而导致根因识别不准确. 因此, 基于根节点的根因识别方法在保证根因变量可识别性方面仍存在一定的不足.

3.1.2 基于分数的根因识别方法

近年来, 考虑到基于根节点的方法中复杂信息传递关系所带来的根因无法确定问题, 基于分数的根因识别方法逐渐得到研究与应用. 图10展示了基于分数的根因识别范式的示意图, 这类方法旨在将复杂的信息传递结构转变为每个变量的根因分数, 而根因变量可以被直接确定为分数最高的变量, 从而保证了根因变量的可识别性. 目前, 基于分数的根因识别方法可以分为两类: 因果信息量化和因果关系对比.

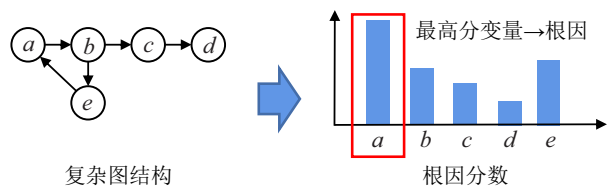


图10 基于分数的根因识别范式示意

因果信息量化方法的核心思想是: 根因变量为故障传播过程中的信息源头, 因此在信息传递结构中应输出最多的信息量, 通过对这种信息量的量化, 即可计算出根因分数.

Yuan等^[8]提出了一种称为因果流量的根因分数. 该分数通过计算因果图中各变量节点出度与入度之差来表示每个变量输出的因果信息量. 该指标核心理念简洁明了, 因果图中节点的出度表示变量向外对其他变量输出信息的次数, 入度则表示改变了

接收其他变量信息的次数. 因此, 二者之差能够表征各变量向外输出的信息量, 其取值越高说明变量越有可能是根因. Chen等^[80]在此基础上提出了一种用于根因识别的可达性根因分数. 该方法首先计算信息传递图结构的可达性矩阵, 用于表征每对变量间是否存在信息通路. 给定邻接矩阵 A , 可达性矩阵 R 的计算方式如下:

$$R = A \cup A^2 \cup \dots \cup A^N, \quad (29)$$

其中 \cup 表示或运算.

定义矩阵 R 第 i 行第 j 列的元素为 r_{ij} , 第 i 个变量 x_i 的可达性根因分数按下式计算:

$$R(x_i) = \sum_{j=1}^N r_{ji} - \sum_{j=1}^N r_{ij}. \quad (30)$$

该分数本质上是可达性矩阵中各变量出度与入度之差, 思想与前述的因果流量指标类似.

Song等^[17]提出了一种用于根因数量化的Root-Rank算法, 该方法构建一个多阶预测贡献矩阵 P 来描述每对变量间的信息传递关系, 并在理论上证明了各变量所输出的信息量能够通过 P 矩阵的特征值分解来计算. 该方法从数学上保证了了解的存在性, 可靠性较强, 其性能在实际卷烟切割过程的故障案例中得到了验证.

不同于基于因果信息量化的方法, 基于因果关系对比的根因评分方法通过对比故障前后因果关系的改变来确定根因. 图11为该类方法的示意图. 如图11所示, 这类方法认为, 根因变量在故障发生后作为故障源头向其他变量传播信息, 因此与其相关的因果关系均被故障传播所主导, 相比于故障前变化最为显著, 从而可以通过量化每个变量在故障前后与其他变量因果关系的改变来计算根因分数. 这一思想最先起源于对于不变网络的故障根因节点识别^[105-107], 近年来其有效性在IRCD任务上也得到证实. Song等^[91]设计了一种因果稳定性指标来描述变量之间的因果关系强度, 通过对比故障前后该指标的变化量实现根因评分. Zhou等^[83]利用时序预测任务中的注意力系数来表示变量间的GC, 并基于故障前后注意力系数的差值来计算根因分数. 实际上, 不受故障影响的因果关系往往源自正常的过程运行机理而非故障传播, 而基于因果关系对比的方法恰好能够排除这种冗余的信息, 使模型更加关注于故障相关的因果关系, 相比于基于因果信息量化更为合理. 然而, 这种方法通常需要对正常数据中的信息传递关系进行分析, 引入了额外的计算开销.

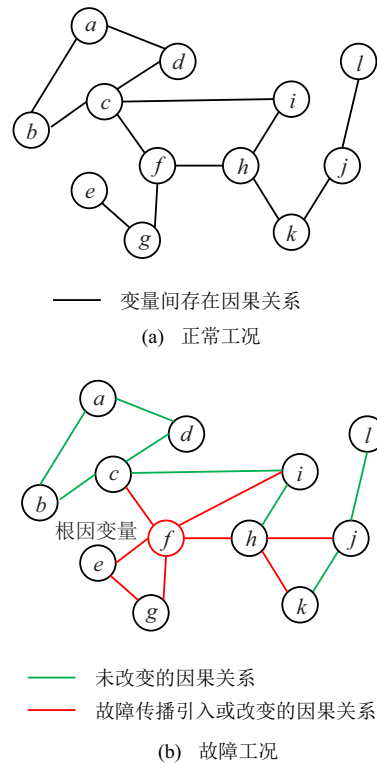


图11 基于因果关系对比的根因评分方法

以上介绍了基于分数的根因识别范式, 该方法能够保证根因变量的可识别性, 具有较强的实用性. 然而, 由于工业场景中对于可解释性的要求, 该方法仅能提供分数, 而不能展现故障传播的路径, 从而并不能完全替代基于根节点的分析方法. 因此, 一种更好的分析思路是将二者结合, 一方面, 通过观察信息传递结构中的故障传播路径来分析RCD结果与过程机理是否吻合; 另一方面, 结合根因分数对复杂的含双向及环形的结构进行分析. Yuan等^[8]在针对由GC方法所获取的因果图结构的分析中引入了这种思想. 在该项研究中, 针对具有双向、环形的复杂因果结构, 首先通过分析因果图中的故障传播路径来初步确定振荡所在的控制回路, 随后引入因果流量指标获取根因分数以识别振荡源. Chen等^[80]将因果图结构与根因分数的协同分析用于TEP中故障案例的根因分析, 通过所获取的因果图分析故障传播路径与机理的吻合性, 随后利用所设计的可达性分数识别位于双向因果结构中的根因变量. 上述两项研究均采用了基于因果信息量化的根因分数, 实际上类似的思想也可以应用于基于因果关系对比的根因评分上. 由于这种根因评分手段并不完全基于故障传播路径, 而是一种对故障前后因果关系改变程度的量化, 该方法所获取的根因评分与故障传播的图结构涵盖了不同信息, 能够从不同的角度印证根因识别结果的可靠性^[83]. 然而, 目前关于两种方法结合的相关研究并不

多见,且通常停留在结果分析阶段,并未在根因分数计算等步骤中实现二者的有机融合。

3.2 知识与数据融合驱动的根本识别方法

通常,工业过程的故障数据较难收集^[108-111]。一方面,在大多数情况下过程都运行在正常状态以完成生产工作;另一方面,出于安全性考虑,在故障发生后应尽快采取故障处理维修措施,因此可能没有充足的时间进行故障数据采集。此外,由于完全基于数据而未考虑过程测点间的连接关系等过程信息,纯数据驱动的根本识别方法所获取的故障传播路径可能并不符合过程机理。因此,一些学者试图将过程的机理知识引入根本识别中。近年来,动态贝叶斯网络(dynamic Bayesian network, DBN)^[112]因其结构灵活、易于与知识以及数据结合的特点,成为知识与数据融合驱动的根本识别的主要方法。此外,经典的知识与数据融合驱动的根本识别方法,包括故障树^[113]、符号有向图(signed directed graph, SDG)^[114]、多级流模型(multilevel flow model, MFM)^[115]也将在下面进行展开介绍。

3.2.1 基于DBN的根本识别模型

DBN使用概率模型描述变量间同时刻以及跨时刻的因果关系。图12以3个变量的情况为例,呈现了DBN的主要结构。如图12所示,DBN模型由图结构和概率参数两部分组成。在图结构中,每个变量视作一个图节点,而变量间的连边及其方向则表示因果信息传递关系,除了同时刻变量间的因果,在相邻采样点上也具有跨时间的因果关系,从而考虑数据的动态性。此外,条件概率则描述了变量间状态的相互作用关系。在RCD任务中,每个变量通常设定为可处于两种状态,即正常和故障,从而描述故障信息的传播。

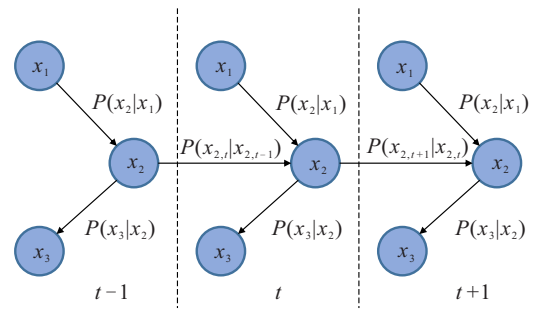


图12 DBN模型结构示意图

DBN的建立通常包含结构确定和参数学习两个步骤。与上述纯数据驱动的根本识别方法不同,DBN包括离线建模和在线诊断两个步骤。Yu等^[19]构建了如图13所示的基于DBN的根本识别框架,该框架在后续研究中被广泛参考和使用。在基于DBN的根本识别中,离线建模不基于在线故障数据,而是基于历史上采集的正常数据和过程机理知识来构建图结构并求解相应的条件概率参数。在该建模阶段,知识和数据实际上分别负责DBN的结构确定和参数学习,二者的融合共同构建了完整的DBN。首先利用机理确定不同变量测点在实际过程结构中是否有连接关系,并依照这种连接关系建立DBN的图结构。这一步骤实际上将过程知识纳入了考虑,具体而言,如果两个变量在过程结构中毫无关联,则二者不存在直接的物理化学作用或者控制关系,因此是不可能具有直接因果关系的。在此之后,由于需要推断不同变量节点的条件概率参数以量化变量间的依赖关系,应引入过程数据,从而基于由机理知识确定的图结构使用极大似然估计来求解概率参数。在线诊断时,由于DBN结构和参数已经确定,只需要通过故障数据来定位异常的变量,并在DBN的网络结构中回溯以确定根因变

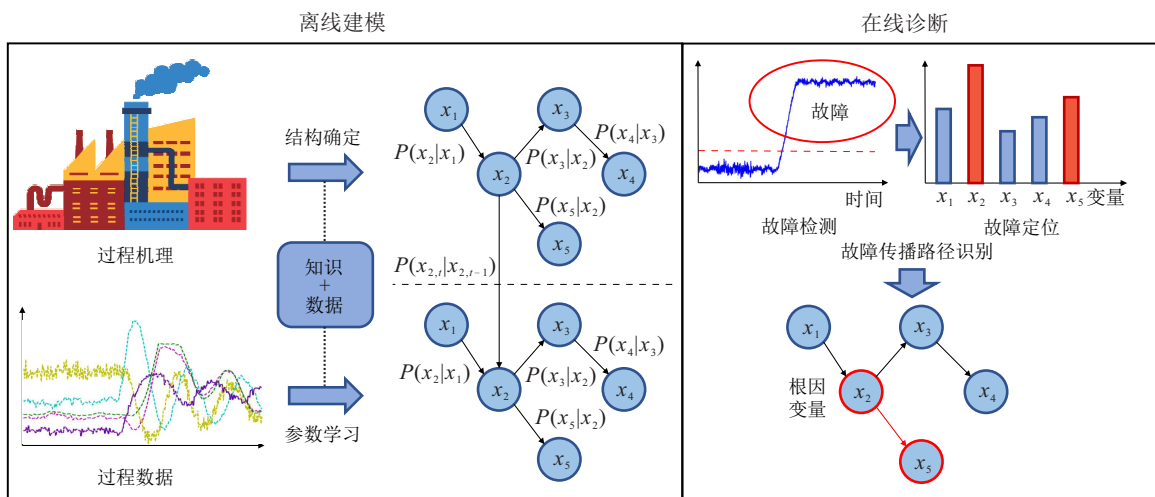


图13 基于DBN模型的知识数据融合驱动的根本识别框架

量即可.值得注意的是,虽然在这一步骤中表面上没有直接利用过程机理知识,但故障根因的回溯需要依照DBN的图结构进行,而图结构是在离线建模阶段由知识所确定的.因此,在线诊断仍旧在数据与知识的融合下进行,并非纯数据驱动.通过这种方式,故障传播路径能够符合过程机理,给出更加可解释、可靠的诊断结果.

尽管DBN拥有上述优势,但仍在理论层面存在缺陷.在传统的DBN中涉及概率参数估计,通常会在这一步骤中假设变量服从高斯分布,从而导致模型对非高斯数据难以适应.另外,DBN具有马尔科夫性假设,即假设变量的状态仅受到上一时刻状态的影响,这可能会使得模型无法合理地处理具有高阶时序相关性的过程数据.此外,因DBN中每个时刻的因果图结构是一致的,故无法处理非平稳过程中时变的因果关系.不仅如此,由于过程机理的复杂性,仅依靠过程知识建立起的图结构可能不够准确,从而影响诊断性能.

近10年来,基于DBN的根因识别方法得到了长足发展,一些关于DBN的改进方法逐步被提出.针对DBN的高斯性和马尔科夫性假设,Zhang等^[116]提出了多时间片的DBN方法,并利用高斯混合模型(Gaussian mixture model, GMM)拟合非高斯分布.Yue等^[117]设计了一种模糊DBN,基于模糊推理准则对变量的异常状态进行精细刻画,避免了高斯分布假设,并将其应用于铝电解槽故障的RCD中.针对DBN中利用知识构建图结构的不准确性,Zhu等^[97]利用工业知识将整个复杂的过程分解为多个子系统,再用TE方法对每个子系统上的信息传递结构进行推断以作为DBN的结构,实现了大规模系统的分解简化.Gharahbagheri等^[118]针对DBN中可能出现的环形结构,设计了一种伪节点方法对结构进行简化,并分析了DBN与常规过程监测方法的适配性.此外,针对环形因果问题,Kumari等^[119]设计了一种分解方法,通过将环形结构中较弱的因果关系转换为时序关系,从而将环形的结构简化为非环形.针对DBN的平稳性假设,Gao等^[120]利用协整分析方法获取非平稳变量间的因果依赖关系,再将其整合到DBN结构中,从而使得DBN能够处理非平稳过程数据.除上述研究外,Zhang^[121]指出DBN的待学习参数随着变量数目呈指数增加,具有高复杂度,并为此提出了一种动态的不确定因果图模型,该模型能够显著提高运算效率,并通过引入逻辑门结构提高了可解释性.

3.2.2 基于故障树的根因识别模型

故障树^[113]是一种典型的知识与数据融合驱动的RCD方法,其通过具有逻辑运算功能的树结构来表示过程的故障模式.一个典型的故障树结构如图14所示.其中: $\{x_1, \dots, x_4\}$ 表示底层部件或元件的异常是否发生, $\{s_a, \dots, s_d\}$ 表示中间故障事件,T表示整个过程或设备是否出现目标故障类型.通过不同底层故障事件与中间故障事件的逻辑运算,故障树能够直观清晰地表示从底层异常到整个过程的系统级故障在这一过程中的传播规律.

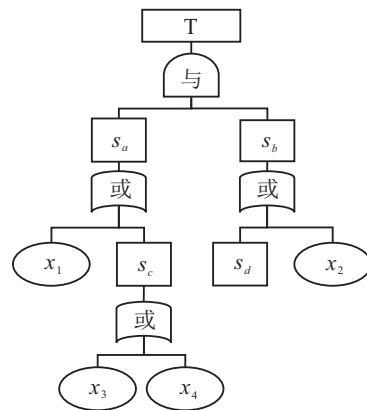


图14 故障树模型结构示意图

故障树的分析主要包含树结构建立、定性分析以及定量分析3个步骤^[122],与DBN类似,这一过程同样体现了知识与数据的融合.为了构建树结构,需要引入专家知识或过程机理分析的指导.随后,在定性分析中利用数学语言描述底层故障事件与顶层事件T之间的逻辑关系,并采取最小割集算法等方法对故障树进行简化^[122].上述两个步骤是纯知识驱动的,而在定量分析步骤中则需要对历史故障数据进行统计分析,从而获取各个底层故障事件的发生概率及其对顶层事件T发生概率的影响关系,以确定各个底层事件的重要度.需要说明的是,对于一个工业过程而言,故障树并不是单一的,考虑到过程中可能存在不同的子系统和多样的故障类型,可以根据需求为不同的故障建立相应的故障树.通过这种方式,在线诊断时对于所发生的故障,可以在其对应的故障树中寻找重要度最高的若干个底层故障事件作为根因候选,从而有针对性地开展故障处理措施.

目前,故障树模型已在工业过程的故障RCD中得到了充分的应用,涉及化工^[113,123]、石油^[124]、核电^[125-126]、新能源^[127]等行业.除了经典故障树模型方法,一些改进的故障树模型也得到了领域内学者们的关注.在传统故障树模型中,故障事件的取值往往是非零即一的确定值,而专家知识往往具有不确定性,

例如对于温度物理量偏高的描述,究竟超过多少摄氏度是偏高的,这在专家知识中往往是不存在绝对标准的.考虑到这一问题,模糊故障树^[128]被提出并应用于IRCD^[129-130],该方法利用隶属度函数描述事件处于不同状态的模糊概率,从而充分刻画了专家知识的不确定性.此外,传统的故障树模型描述的是一种静态的故障传播规律,然而,在实际过程中故障往往是随时间演化的,从而形成了一个时间维度上的故障序列,难以利用静态的树结构来描述.为了处理该问题,动态故障树被提出^[131],该方法在原有故障树的基础上引入了新型的逻辑门结构以描述时间上的变化关系,目前这种方法已在IRCD中得到了一定应用^[132].

尽管上述研究表明故障树在IRCD中有着一定的适用空间,但这种方法对知识的充分性和可靠性要求较高,可能不适合于大规模且结构复杂的现代工业过程,近年来的相关研究较少.此外,故障树方法结构较为固定,往往局限于历史已有的故障类型,难以应对多样的未知故障.已有学者对故障树模型和前述的贝叶斯网络方法进行了充分对比,发现贝叶斯网络方法在通用性、可靠性以及灵活性方面相比于故障树具有优势^[133].因此,一些学者也主张将故障树转化为贝叶斯网络以实现更加灵活可靠的故障根因识别^[134].

3.2.3 基于SDG的根因识别模型

SDG是另一种典型的基于图的知识表征方式,用于数据与知识融合驱动的IRCD任务,由Kramer等^[14]于1987年提出.图15给出了SDG的典型结构,其中各个节点代表变量,而有向的连边则表示变量间的依赖关系.与一般的图结构不同,SDG中的节点和边均有正负属性.节点的正负分别表示对应变量取值超出正常范围的上限或低于正常范围的下限,当变量取值在正常范围内时以零表示;边的正负则表示变量间是按相同还是相反的方向变化.在建模阶段,可以根据专家知识和过程机理建立起变量间的连边并确定正负.在线诊断时,根据不同变量数据的取值情况,在SDG中寻找一致边进行根因追溯与识别.具体地,如果起始节点的符号乘以边的符号再乘以结果节点的符号,其结果是正的,则称该边为一致边,能够反映故障的传播路径.根据在线获取的故障数据能

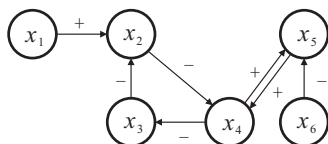


图15 符号有向图模型结构示意图

够找到所有的一致边以给出故障传播的足迹并确定根因变量,这一过程体现了知识与数据的融合.

目前,SDG方法已在不同种类的工业过程中得到了应用,包括化工过程^[135]、核能发电过程^[136]等.此外,针对过程的复杂特性,一些相关的改进SDG模型也被提出,如引入时序信息的SDG方法^[137]、引入模糊语义描述的SDG^[138-139]、针对大规模过程的层次化SDG方法^[140],以及引入条件概率精细化描述变量间关系的概率化SDG^[141]等.然而,随着现代工业过程的发展,测点数目不断增加也导致SDG的节点众多,构建成本高昂.此外,由于过程变量间的复杂耦合关系,一些故障实际上无法通过单变量是否超出高低阈值的规则来识别,而仅能通过提取多变量的协同关联才能被反映出来,因此,SDG方法用于复杂工业过程的RCD还存在着一定制限制.

3.2.4 基于MFM的根因识别模型

MFM将一个工业过程中各个设备、环节之间的相互作用表示为“流”的形式.即:在过程中,物料被输入到不同的设备中进行物理处理或化学反应,表现为物质流的形式;在不同流程中,能量在电能、化学能、热能或机械能等形式间相互转化,呈现为能量流;在过程中,诸多控制系统的执行、观测等作用构成了信息流.MFM正是从过程中的物质流、能量流、信息流的原理入手,结合不同的流所期望实现的目标或功能,对过程机理进行直观的描述.在进行RCD时,能够直接分析故障中各个环节相互影响所导致的连锁反应,从而定位故障的根因环节.

为便于说明MFM的结构和思想,图16给出一个相对简单的供水过程案例.图16(a)为该供水过程的示意图^[142],在该过程中,水泵 C_2 从存水箱 C_1 中抽水并供给左水箱 C_3 ;同时,左右水箱 C_3 和 C_5 通过水管 C_4 连接以维持水位平衡;另外, C_5 中的水通过水管 C_6 流回存水箱 C_1 .整个过程工作条件为电源 C_8 能够正常供电且开关 C_7 处于闭合状态.图16(b)展示了该过程的MFM.其中:物质流的功能节点 $F_1 \sim F_7$ 分别表示存水箱水源、由 C_2 实现的传送水、 C_3 中的存储水、由 C_4 实现的传送水、 C_5 中的存储水、由 C_6 实现的传送水、由 C_1 实现的排水;能量流中的 $F_8 \sim F_{10}$ 分别表示由电源 C_8 提供的供电电能、由开关 C_7 实现的电能传输、由水泵 C_2 电机实现的电能消耗.需要说明的是,由于篇幅限制,这里的示例过程较为简单,仅为说明MFM的思想理念,并不包含信息流.此外,该MFM中还包含两个目标节点:目标 G_1 与物质流中的功能节点 $F_1 \sim F_7$ 对应, G_2 与能量流中的 $F_8 \sim F_{10}$

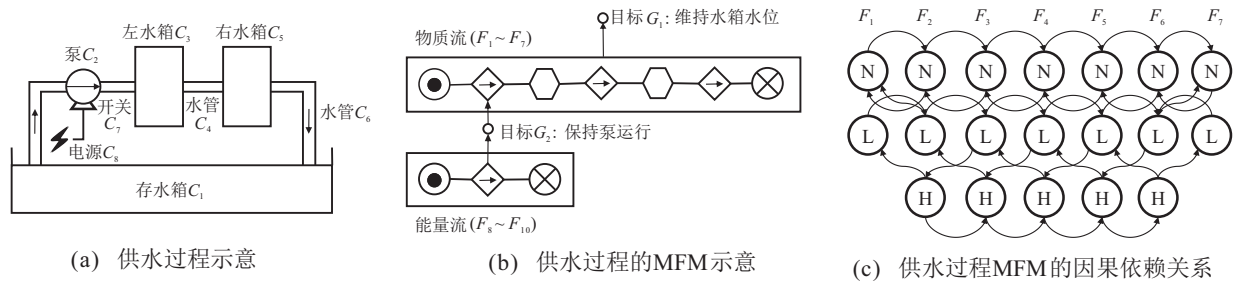


图 16 供水过程的MFM结构以及因果依赖关系

对应。为实现故障根因的追溯,MFM 还需包含图 16(c)所示的因果依赖关系,这里仅列出 $F_1 \sim F_7$ 的依赖关系。其中:每列表示节点的状态,N表示正常,L表示偏低,H表示偏高,不同方向的箭头表示节点间的正向及反向信息传播关系。MFM的构建均依赖于过程机理或专家知识。当故障发生时,根据在线数据获取各个节点的状态,再利用MFM中因果依赖关系的正向及反向传播关系即可确定根因节点。在这一过程中体现了数据与知识的融合。

MFM的思想理念于20世纪80年代就已被提出,并成为当时一种专家系统中的描述语言^[115]。从20世纪90年代开始被用于工业过程的RCD^[143]。迄今为止,MFM已在火电过程^[142]、化工过程^[144-145]、核电过程^[146]等场景中得到应用。然而,与故障树和SDG类似,MFM由于对过程机理知识的可靠性和精细度要求较高,通常难以适用于复杂工业过程,在近5年内相关研究较为少见。

3.3 根因识别方法的性能对比

上文中介绍了不同根因识别方法的原理与研究现状,表2中给出了这些方法的优势与不足以及所适

用的工业对象。可以发现,在处理具有复杂因果结构的工业对象方面,基于分数的方法具有显著优势,而DBN等包含了过程知识的方法则在可解释性和机理吻合性上更胜一筹。然而,在近几年的RCD研究领域,相比于基于结构推断结果直接对故障数据进行根因识别的纯数据驱动方法,融合了知识的方法在RCD上的研究相对较少,这是因为过程规模日益庞大,结构日趋复杂,很难利用充分的过程知识确定这些模型的结构。以DBN为例,其整个概率图的结构几乎都要基于过程变量的连接关系进行设计,尽管部分研究方法^[97]主张采用TE等结构推断模型生成DBN的结构,但这样基于DBN的方法又转变成了纯数据驱动模型,无法体现其在可解释性等方面的优势。相比于DBN,故障树、SDG、MFM对知识的精细化和可靠性要求更高,不仅要求明确变量间的连接关系,还需要对变量间的作用和影响进行定性描述。因此,知识与数据融合驱动的方法对于机理明确的简单设备或系统较为适用,而在面对复杂的现代工业过程时,数据驱动的根本因识别手段则更具优势。

表 2 根因识别模型的特点及适用范围比较

方法	模型	根因可识别性	可解释性	机理吻合度	应用场景特点	适用工业对象
纯数据驱动的根本因识别方法	基于根节点的根因识别模型	无法保证	较强,能够给出故障传播路径	无法保证	故障传播结构较为简单,没有环形、双向结果的情况	通常为单设备、开环系统
知识与数据融合驱动的根本因识别方法	基于DBN的根因识别模型	部分能够处理有环结构的改进方法能够保证	强,能够融合过程机理知识	取决于建模使用的过程知识的准确性	故障数据较少,但可以提供知识描述的情况	有明确过程机理模型的简单设备或系统
	故障树、SDG、MFM	能够保证	极强,结构完全由机理知识构建,仅在在线诊断阶段使用数据	取决于建模使用的过程知识的准确性	故障数据较少,但可以提供极为明确的知识描述的情况	过程机理完全清楚的简单设备或系统,需要具体明确变量间的影响关系

4 工业过程故障RCD中的问题与挑战

由于过程复杂性和数据层面的种种限制,IRCD领域内仍存在着一些问题与挑战。这些问题已被研究人员意识到,但由于理论层面的难题与制约,尚无

明确的解决方法。基于上述回顾与分析,本节针对工业过程故障RCD研究中存在的问题,分别从根因可识别性不定、性能评估指标匮乏、结构信息单一限制这3个方面进行总结与分析。

4.1 根因可识别性不定

根因可识别性的概念已在3.1节中有所提及,由于过程中反馈控制以及复杂的耦合关系,双向以及环形的信息传递结构是不可避免的,这会使得许多结构推断的结果中无法清晰地呈现出根因,造成根因可识别性无法保证,面临不确定性的挑战. 尽管现有研究已从信息传递结构简化^[16,18]、根因评分^[17,80,83,91]等方面对该问题进行了处理,但这些研究尚不充分,仍存在一些应用限制.

一方面,过程中一些反馈控制本质上会导致变量间的相互影响,因此,过程中部分双向、环形结构是本来就存在且符合机理的,硬性地将这些结构去除或简化未必符合过程的实际运行模式,甚至可能带来错误的简化. 例如,Chen等^[77]对TEP基准数据集的故障传播机理进行了分析,发现在一些典型故障案例中,根因变量可能有不止一个,而这些变量恰处于反馈控制回路中,如果强行去除它们之间的双向或环形结构,则势必会造成根因变量的漏报. 如何在简化过程中设置合理的标准,确定哪些是的确应该存在的结构,而哪些是需要进行简化的冗余结构,是目前研究中尚未解决的难题.

另一方面,根因评分方法能够在不改变信息传递结构的前提下识别根因变量,但此类方法仍存在缺陷,即无法合理确定根因数目. 尽管可以根据根因分数将变量由高到低进行排列,但并不清楚应取分数最高的前多少个变量作为故障的根因变量. 例如上文所介绍的RootRank模型^[17],其对于双根因变量的RCD任务,能够正确地将两个根因变量赋予最高的评分,但两个根因变量对于整体故障传播所提供的信息量难以避免地存在差异,因此二者的根因分数仍旧有高低之分. 在实际应用场景中,根因的数目往往是未知的,在这种情况下难以确定应取多少个变量作为故障根因.

值得注意的是,上述的信息传递结构简化以及根因评分两种途径都是在根因识别阶段对变量间故障传播所产生的图结构进行后处理,本质上与结构推断相互割裂,难以确定这种独立的后处理操作是否与变量间的信息传递结构相吻合,从而可能导致错误的处理结果. 反之,如果能在结构推断步骤中直接去除冗余的环形、双向结构,则可在符合信息传递机制的前提下从根本上保证根因的可识别性,避免割裂处理所引发的不一致性. 近年来,在因果分析的理论研究中已有学者关注到了该问题,并设计了结构推断中的无环约束项来保证根因的可识别性(此处的无环约束

对双向因果也具有抑制效果,因为双向因果可以视为最简单的环形结构). Zheng等^[147]为因果图结构的邻接矩阵提出了一种指数型无环约束,即最小化邻接矩阵的指数变换的迹. 该研究指出,由于邻接矩阵的指数变换本质上等价于各阶可达性矩阵的加权求和,指数变换后矩阵的对角线元素表示了各变量节点的自可达性,即节点通过因果图中的环形结构回到自身的可能性. 由于该约束是可微分的,可以直接嵌入到因果分析模型的优化目标中,在推断显著因果关系的同时去除冗余环状结构,而对于显著的环形因果结构则给予保留. 在此基础上,Yu等^[148]对这一指数约束进行了简化并将其拓展至神经网络框架中实现非线性无环因果图的提取. Pamfil等^[149]将该约束方法进一步拓展到时间序列的结构推断中. 目前,这种基于无环约束的思想理念已在因果分析的理论研究领域被广泛采用^[150-151].

尽管人们从理论上提出了上述无环约束方法,但该类方法所基于的因果分析框架与目前工业领域中广为接受的GC、TE等并不相同,且未在工业过程数据上得到应用和验证. 因此,这种无环约束该如何用于IRCD仍有待探究. 此外,这些研究中所提出的无环约束需要作为一个可微分项纳入优化目标中,可能与GC模型的适配度较高,而对于TE、CCM等不具有优化目标的方法则不能适用. 综上,如何保证IRCD中的根因可识别性仍然是一个亟待解决的问题.

4.2 性能评估指标匮乏

目前领域中缺乏对RCD方法合理的定量评估指标. 在故障分类和软测量等典型工业任务中,均可采用准确率和预测误差等定量指标直观、量化地评估模型性能. 然而,RCD的目标是识别故障的根因变量,难以用合理的量化指标描述识别的性能. 例如,对于仅有一个根因变量的故障案例,识别结果仅有成功识别和未识别两种可能,诊断成功率非1即0,因此常规的评判准则很难适用. 此外,一些RCD模型会受到随机性影响,且具有较多的超参数. 例如,基于深度学习的RCD方法在不同的参数设定、不同的训练轮次后可能会得到截然不同的诊断结果,即便在某次训练后恰好正确地识别了根因,也无法确定该方法一定具有准确的性能. 如果缺乏评估指标,则无法衡量方法的优劣,从而无法有效地在实际应用场景中进行模型选取.

迄今为止,多数的RCD研究中并不会直接给出RCD的性能指标,例如,绝大多数基于根节点和DBN的研究,通常采用可视化的方法说明根因变量确实

落在了所识别的故障传播路径的源头上.少数方法利用结构推断的准确性从侧面反映RCD的性能.例如,Zhou等^[83]利用因果推断的 F_1 指标来评估RCD的性能.然而,这种评估方式并不能科学地反映诊断的准确度,在极端情况下,即便大多数因果推断结果均是正确的,只要根因变量被错误识别成其他变量的下游变量,则RCD的最终结果也是错误的.Song等^[17]指出,针对基于深度学习等受随机性影响显著的RCD方法,不应以一次诊断结果的正确与否作为性能评判的依据,而应进行多次随机实验,取其诊断准确率的均值.这种方法能够有效比较具有随机性的RCD模型之间的性能,而对于传统GC等不受随机性影响的方法则无法进行处理.因此,如何为RCD任务设定一个通用、可靠的评估指标仍是一个尚未解决的问题.

4.3 结构信息单一限制

目前领域内大多RCD方法,尤其是基于因果关系的方法,通常仅考虑变量间的直接信息传递关系,通过多变量之间若干直接信息传递关系组合成整体的信息传递结构.而实际上,根因变量对其他下游变量的影响不仅局限于直接的信息传递,间接的关系也应纳入考虑^[17,80,152-153].图17展示了RCD中考虑间接关系的必要性.其中:红色节点表示根因变量,蓝色节点表示受故障影响的下游变量,箭头代表故障信息的传播路径.图17中上半部分描述的是绝大多数现有方法所提供的信息传递结构,可以发现该结构中,对于根因变量的识别真正起作用的仅仅是与红色根因变量节点直接相连的信息传播路径,而在更下游变量间的故障信息传播并没有对根因的识别做出贡献.因此,仅凭借单一的直接关系无法对根因变量的故障传播贡献进行完整地刻画.更合理的RCD方法如图17中下半部分所示,不仅需要关注直接与根因变量相连的故障传播路径,还应将根因变量对下游变量的间接影响纳入考虑,从而充分表征根因对故障传播的贡献.

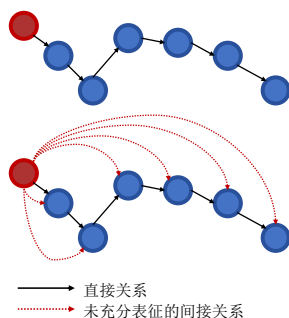


图17 RCD任务中的直接与间接信息传递关系

目前,已经有一些RCD方法意识到了间接信息传递结构的重要性.Jiang等^[152-153]将直接因果关系构成的邻接矩阵扩展到考虑多步影响的可达性矩阵.Chen等^[80]在此基础上设计了用于根因识别的可达性根因分数,已在3.1.2节中进行了介绍.然而,上述方法仅是基于直接信息传递关系的扩展,事实上,多阶可达性并不等同于真正的间接信息传递关系.Song等^[17]提出了MPGE方法,该方法通过在预测任务中引入多层信息融合机制来提取间接的非线性GC关系.该方法证实了同时考虑直接和间接因果能够显著提高模型对根因变量的表征能力.然而,这种方法仅限于在GC框架中使用,对于其他的典型RCD方法,尚不存在能够分析间接信息传递关系的改进方法.因此,如何突破单一的直接关系限制,使模型考虑间接影响,是未来RCD研究中需要突破的难题.

5 工业过程故障RCD研究展望

工业过程故障的RCD在近年来逐步受到学者的深入研究与应用,为工业过程的安全维护做出了显著贡献.如何提取可解释的、可靠的信息传递结构并高效率识别根因变量一直是IRCD领域所追求的目标.本节针对IRCD的可解释性和可靠性两方面需求,分别从知识与数据融合驱动的结构推断、开源数据集与统一评估标准两方面做出研究展望.这两方面研究在IRCD领域内具有重要意义且具备一定的可行思路,然而尚未引起广泛的关注和重视.

5.1 知识与数据融合驱动的结构推断

在过程的构建与运行期间,工业现场会积累一定数目的工业知识,如故障排查检修记录、故障特点、过程机理与测点分布、测点间的连接关系等.对于这些知识信息的应用能够显著提高模型的可解释性,使数据分析结果与过程知识相吻合.尽管如此,工业知识在对可解释性要求较高的RCD领域却仅有少量应用,主要是利用测点间连接关系和分布信息作为知识来对结构推断结果进行约束.事实上,这些方法对知识的利用还不够充分,前述的检修记录、故障特点等知识尚未利用.近年来,将知识嵌入为数据驱动模型可理解的向量、矩阵等形式是过程控制领域研究的新兴方向之一,通过这种方式可以直接将知识融入模型,指导结构推断过程.在故障分类领域中,已有一定数目的研究证实了知识信息的引入能够提高故障判别的准确性并拓展模型的可应用场景.例如,Feng等^[154]提出了工业零样本学习范式,将过程中对于故障特点的描述嵌入为语义属性向量,从而将自然语言信息转换成计算机可理解的向量以供模型学习.Li

等^[155]为火力发电过程构建了故障描述知识库,并通过知识嵌入将其应用于工业联邦学习的故障分类任务中. 这些方法均为知识与数据融合驱动的IRCD带来启发与参考.

5.2 开源基准数据集与统一评估标准

本文在4.2节中阐述了IRCD缺乏合理的性能评估指标,事实上,导致该问题的原因之一是目前可供进行性能验证的开源基准数据集过少,且缺少统一的数据标签. 如前文所述,对于一个故障案例,只要一个RCD模型正确诊断出它的故障根因变量,则该模型

的诊断就是成功的,准确率为100%,然而这种评估显然是不可靠的. 为了获取足够可靠的性能评估标准,应在足够多的故障案例中进行诊断,以获得平均的诊断准确率. 然而,由于缺乏足够的具备根因标签的故障案例,这种评估方式往往难以实现. 表3中统计了本文所涉及的近5年来在IRCD领域发表的部分研究中使用的数据、开源情况以及模型性能的评估标准. 从表3中可以总结出如下几点结论:

1) 目前,能够开源获取的基准数据集较少,且主要为仿真数据. 实际工业数据由于隐私性、保密要求

表3 本文所涉及的近5年(2018~2023年4月)部分IRCD研究中使用的实验数据

文献编号	发表年份	实验数据	数据开源	诊断结果准确性的评估依据
[16]	2019	1) 数值仿真 2) 青霉素发酵过程仿真数据	1) 非开源 2) 开源	1) 数据生成机制 2) 故障生成机制
[17]	2023	1) 数值仿真 2) TEP 仿真数据集, 故障案例 1 和 8 3) 卷烟切割过程实测数据	1) 开源 2) 开源 3) 非开源	1) 数据生成机制 2) 过程故障机理 3) 过程故障机理
[18]	2020	TEP 仿真数据集, 故障案例 1、4、5、6、7	开源	过程故障机理
[46]	2018	TEP 仿真数据集, 故障案例 1 和 7	开源	过程故障机理
[48]	2020	实测的热轧带钢生产过程数据 (故障类型: 热轧带钢头部拉窄)	非开源	过程故障机理
[49]	2020	液化天然气分馏仿真过程	非开源	数据生成机制
[56]	2018	1) TEP 仿真数据集, 故障案例 1、8~15 2) 实测离心机数据	1) 开源 2) 非开源	过程故障机理
[58]	2020	选矿厂浮选回路中的实测振荡数据	开源	过程故障机理
[63]	2018	实测热轧钢过程数据	非开源	过程故障机理
[70]	2020	火力发电过程中实测磨煤机数据	非开源	过程故障机理
[71]	2021	1) TEP 仿真数据集, 故障案例 1 和 7 2) 火力发电过程中实测磨煤机数据	1) 开源 2) 非开源	过程故障机理
[75]	2021	1) 数值仿真 2) TEP 仿真数据集, 故障案例 1	1) 非开源 2) 开源	1) 数据生成机制 2) 过程故障机理
[77]	2018	1) 数值仿真 2) TEP 仿真数据集, 故障案例 1 和 7	1) 非开源 2) 开源	1) 数据生成机制 2) 过程故障机理
[79]	2021	TEP 仿真数据集, 故障案例 1	开源	过程故障机理
[80]	2022	1) 数值仿真 2) TEP 仿真数据集, 故障案例 1 3) 卷烟切割过程实测数据	1) 非开源 2) 开源 3) 非开源	1) 数据生成机制 2) 过程故障机理 3) 过程故障机理
[83]	2022	1) 数值仿真 2) TEP 仿真数据集, 故障案例 3、8、14	1) 非开源 2) 开源	1) 因果分析结果的 F_1 指标 2) 过程故障机理
[86]	2023	1) 数值仿真 2) TEP 仿真数据集, 故障案例 1 和 4	1) 非开源 2) 开源	1) 根因定位与传播路径识别的召回率和精确率 2) 过程故障机理
[90]	2023	实测热轧钢过程数据	非开源	专家知识
[91]	2022	1) TEP 仿真数据集, 故障案例 1、4、7 2) 火力发电过程冷凝系统实测数据	1) 开源 2) 非开源	过程故障机理
[95]	2023	火力发电过程中实测磨煤机数据	非开源	过程故障机理
[97]	2019	TEP 仿真数据集, 故障案例 1	开源	过程故障机理
[119]	2022	TEP 仿真数据集	开源	过程故障机理
[120]	2019	TEP 仿真数据集, 故障案例 4	开源	过程故障机理

等原因,往往无法开放获取。

2) 在已开源的基准数据集中,同样缺乏通用的、广为认可的故障案例。即便是应用最为广泛的TEP数据集,现有的研究工作通常只使用部分故障案例,且不同的研究中使用的案例也往往不一致。

3) 对于工业数据,其根因标签较难获取。从表3中可以发现,由于缺乏带有根因标签的数据集,大多数IRCD研究往往需要借助过程机理和专家知识来确定真实标签,再评估模型的诊断结果是否正确。这无疑会耗费大量的人力和时间成本。

综上,在未来IRCD领域的研究中,亟需构建标签信息丰富、故障案例充足的开源数据集,以便在足够的案例上对模型的性能进行评估,进而保证模型的可靠性和有效性。

6 结论

工业过程的RCD技术对于保障生产安全、提高运维效率具有重要价值。随着工业信息化、智能化的不断进展,IRCD也必将在未来获得更加广泛的关注与研究。本文构建了IRCD框架中的两个层级:结构推断以及根因识别,以新颖的视角回顾了IRCD相关的理论发展与实际应用。在结构推断方面,本文梳理并分析了不同推断模型的特点与适用场景,建立起理论模型与实际过程特性间的桥梁。在根因识别方面,本文从根因可识别性、可解释性以及机理吻合性等方面综合对比了不同识别范式的性能与适用范围。根据对现有研究的总结与讨论,本文从根因可识别性、评估指标、结构信息3方面揭示了领域内亟待解决的问题与挑战。鉴于此,本文从可解释性和可靠性两方面出发,对IRCD领域未来的发展进行了研究展望,为后续研究提供了有益的参考。

参考文献(References)

- [1] 顾幸生, 周冰倩. 基于LNS-DEWKECA算法的多模态工业过程故障检测[J]. 控制与决策, 2020, 35(8): 1879-1886.
(Gu X S, Zhou B Q. Multimodal industrial process fault detection based on LNS-DEWKECA[J]. Control and Decision, 2020, 35(8): 1879-1886.)
- [2] 赵春晖, 余万科, 柴铮. 运行工况监测与故障溯源推理: 机器学习方法[M]. 北京: 化学工业出版社, 2022: 1-9.
(Zhao C H, Yu W K, Chai Z. Operating condition monitoring and fault tracing reasoning[M]. Beijing: Chemical Industry Press, 2022: 1-9.)
- [3] 姚远, 佟佳蓉, 高军, 等. 基于分层分块DLPPCA-SVM的复杂工业过程监测与故障诊断方法[J]. 控制与决策, 2022, 37(5): 1402-1408.
(Yao Y, Tong J R, Gao J, et al. Monitoring and fault diagnosis method of complex industrial process based on DLPPCA-SVM[J]. Control and Decision, 2022, 37(5): 1402-1408.)
- [4] Zhao C H, Sun H. Dynamic distributed monitoring strategy for large-scale nonstationary processes subject to frequently varying conditions under closed-loop control[J]. IEEE Transactions on Industrial Electronics, 2019, 66(6): 4749-4758.
- [5] Yu W K, Zhao C H, Huang B. MoniNet with concurrent analytics of temporal and spatial information for fault detection in industrial processes[J]. IEEE Transactions on Cybernetics, 2022, 52(8): 8340-8351.
- [6] Cheon S W, Chang S H, Chung H Y. Development strategies of an expert system for multiple alarm processing and diagnosis in nuclear power plants[J]. IEEE Transactions on Nuclear Science, 1993, 40(1): 21-30.
- [7] Li G, Qin S J, Yuan T. Data-driven root cause diagnosis of faults in process industries[J]. Chemometrics and Intelligent Laboratory Systems, 2016, 159: 1-11.
- [8] Yuan T, Qin S J. Root cause diagnosis of plant-wide oscillations using Granger causality[J]. Journal of Process Control, 2014, 24(2): 450-459.
- [9] Bauer M, Thornhill N F, Meaburn A. Specifying the directionality of fault propagation paths using transfer entropy[J]. IFAC Proceedings Volumes, 2004, 37(9): 203-208.
- [10] Hellesteth T. Some results about the cross-correlation function between two maximal linear sequences[J]. Discrete Mathematics, 1976, 16(3): 209-232.
- [11] Li H H, Liu J X, Yang Z L, et al. Adaptively constrained dynamic time warping for time series classification and clustering[J]. Information Sciences, 2020, 534: 97-116.
- [12] Fukunaga K, Narendra P M. A branch and bound algorithm for computing k -nearest neighbors[J]. IEEE Transactions on Computers, 1975, C-24(7): 750-753.
- [13] Granger C W J. Investigating causal relations by econometric models and cross-spectral methods[J]. Econometrica, 1969, 37(3): 424.
- [14] Schreiber T. Measuring information transfer[J]. Physical Review Letters, 2000, 85(2): 461-464.
- [15] Sugihara G, May R, Ye H, et al. Detecting causality in complex ecosystems[J]. Science, 2012, 338(6106): 496-500.
- [16] He F, Wang C J, Shu-Kai S F. Fault detection and root cause analysis of a batch process via novel nonlinear dissimilarity and comparative granger causality analysis[J]. Industrial & Engineering Chemistry Research, 2019, 58(47): 21842-21854.
- [17] Song P Y, Zhao C H, Huang B. MPGE and RootRank: A sufficient root cause characterization and quantification framework for industrial process faults[J]. Neural Networks, 2023, 161: 397-417.
- [18] Liu Y, Chen H S, Wu H B, et al. Simplified Granger

- causality map for data-driven root cause diagnosis of process disturbances[J]. *Journal of Process Control*, 2020, 95: 45-54.
- [19] Yu J, Rashid M M. A novel dynamic Bayesian network-based networked process monitoring approach for fault detection, propagation identification, and root cause diagnosis[J]. *AIChE Journal*, 2013, 59(7): 2348-2365.
- [20] Duan P, Chen T W, Shah S L, et al. Methods for root cause diagnosis of plant-wide oscillations[J]. *AIChE Journal*, 2014, 60(6): 2019-2034.
- [21] Lucke M, Chioua M, Thornhill N F. From oscillatory to non-oscillatory disturbances: A comparative review of root cause analysis methods[J]. *Journal of Process Control*, 2022, 113: 42-67.
- [22] 马亮, 彭开香, 董洁. 工业过程故障根源诊断与传播路径识别技术综述[J]. *自动化学报*, 2022, 48(7): 1650-1663.
(Ma L, Peng K X, Dong J. Review of root cause diagnosis and propagation path identification techniques for faults in industrial processes[J]. *Acta Automatica Sinica*, 2022, 48(7): 1650-1663.)
- [23] Bauer M, Thornhill N F. A practical method for identifying the propagation path of plant-wide disturbances[J]. *Journal of Process Control*, 2008, 18(7/8): 707-719.
- [24] Landman R, Jämsä-Jounela S L. Data-driven causal analysis and its application on a large-scale board machine[D]. Espoo: Aalto University, 2013.
- [25] Stockmann M, Haber R, Schmitz U. Source identification of plant-wide faults based on k nearest neighbor time delay estimation[J]. *Journal of Process Control*, 2012, 22(3): 583-598.
- [26] De Maesschalck R, Jouan-Rimbaud D, Massart D L. The mahalanobis distance[J]. *Chemometrics and Intelligent Laboratory Systems*, 2000, 50(1): 1-18.
- [27] Perez-Cruz F, Bousquet O. Kernel methods and their potential use in signal processing[J]. *IEEE Signal Processing Magazine*, 2004, 21(3): 57-65.
- [28] Camastra F, Verri A. A novel kernel method for clustering[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(5): 801-805.
- [29] Holland P W. Statistics and causal inference[J]. *Journal of the American Statistical Association*, 1986, 81(396): 945-960.
- [30] Spirtes P. Introduction to causal inference[J]. *Journal of Machine Learning Research*, 2010, 11(5): 1643-1662.
- [31] Yao L Y, Chu Z X, Li S, et al. A survey on causal inference[J]. *ACM Transactions on Knowledge Discovery from Data*, 2021, 15(5): 1-46.
- [32] Kuang K, Li L, Geng Z, et al. Causal inference[J]. *Engineering*, 2020, 6(3): 253-263.
- [33] Simon H A. Spurious correlation: A causal interpretation[J]. *Journal of the American Statistical Association*, 1954, 49(267): 467-479.
- [34] Buchanan M. Cause and correlation[J]. *Nature Physics*, 2012, 8(12): 852.
- [35] Sinha S, Vaidya U. Causality preserving information transfer measure for control dynamical system[C]. *IEEE the 55th Conference on Decision and Control*. Las Vegas, 2016: 7329-7334.
- [36] Geweke J F. Measures of conditional linear dependence and feedback between time series[J]. *Journal of the American Statistical Association*, 1984, 79(388): 907-915.
- [37] Liao W, Mantini D, Zhang Z Q, et al. Evaluating the effective connectivity of resting state networks using conditional Granger causality[J]. *Biological Cybernetics*, 2010, 102(1): 57-69.
- [38] Chicharro D. On the spectral formulation of Granger causality[J]. *Biological Cybernetics*, 2011, 105(5): 331-347.
- [39] Toda H Y, Phillips P C B. Vector autoregression and causality: A theoretical overview and simulation study[J]. *Econometric Reviews*, 1994, 13(2): 259-285.
- [40] Siggiridou E, Kugiumtzis D. Granger causality in multivariate time series using a time-ordered restricted vector autoregressive model[J]. *IEEE Transactions on Signal Processing*, 2016, 64(7): 1759-1773.
- [41] Shojaie A, Fox E B. Granger causality: A review and recent advances[J]. *Annual Review of Statistics and Its Application*, 2022, 9: 289-319.
- [42] Indyk P, Motwani R. Approximate nearest neighbors: Towards removing the curse of dimensionality[C]. *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*. New York: ACM, 1998: 604-613.
- [43] Baccalá L A, Sameshima K. Partial directed coherence: A new concept in neural structure determination[J]. *Biological Cybernetics*, 2001, 84(6): 463-474.
- [44] Yuan T, Qin S J. Root cause diagnosis of plant-wide oscillations using Granger causality[J]. *Journal of Process Control*, 2014, 24(2): 450-459.
- [45] Landman R, Kortela J, Sun Q, et al. Fault propagation analysis of oscillations in control loops using data-driven causality and plant connectivity[J]. *Computers & Chemical Engineering*, 2014, 71: 446-456.
- [46] Chen H S, Yan Z B, Zhang X L, et al. Root cause diagnosis of process faults using conditional granger causality analysis and maximum spanning tree[J]. *IFAC-PapersOnLine*, 2018, 51(18): 381-386.
- [47] Downs J J, Vogel E F. A plant-wide industrial process control problem[J]. *Computers & Chemical Engineering*, 1993, 17(3): 245-255.
- [48] 何飞, 杜学飞, 王朝俊. 基于对比格兰杰因果关系的热轧带钢头部拉窄根因诊断[J]. *中国机械工程*, 2020, 31(19): 2340-2346.

- (He F, Du X F, Wang C J. Root cause diagnosis of head narrowing in hot rolled strip based on comparative Granger causality[J]. *China Mechanical Engineering*, 2020, 31(19): 2340-2346.)
- [49] Pyun H, Kim K, Ha D, et al. Root causality analysis at early abnormal stage using principal component analysis and multivariate Granger causality[J]. *Process Safety and Environmental Protection*, 2020, 135: 113-125.
- [50] Ahmed U, Ha D, Shin S, et al. Estimation of disturbance propagation path using principal component analysis (PCA) and multivariate granger causality (MVGCC) techniques[J]. *Industrial & Engineering Chemistry Research*, 2017, 56(25): 7260-7272.
- [51] Gharahbagheri H, Imtiaz S, Khan F. Combination of KPCA and causality analysis for root cause diagnosis of industrial process fault[J]. *The Canadian Journal of Chemical Engineering*, 2017, 95(8): 1497-1509.
- [52] Tao Y, Li G, Zhang Z H, et al. Deep causal mining for plant-wide oscillations with multilevel Granger causality analysis[C]. 2016 American Control Conference. Piscataway: IEEE, 2016: 5056-5061.
- [53] Barnett L, Barrett A B, Seth A K. Granger causality and transfer entropy are equivalent for Gaussian variables[J]. *Physical Review Letters*, 2009, 103(23): 238701.
- [54] Bauer M, Cox J W, Caveness M H, et al. Finding the direction of disturbance propagation in a chemical process using transfer entropy[J]. *IEEE Transactions on Control Systems Technology*, 2007, 15(1): 12-21.
- [55] Duan P, Yang F, Shah S L, et al. Transfer zero-entropy and its application for capturing cause and effect relationship between variables[J]. *IEEE Transactions on Control Systems Technology*, 2015, 23(3): 855-867.
- [56] Rashidi B, Singh D S, Zhao Q. Data-driven root-cause fault diagnosis for multivariate non-linear processes[J]. *Control Engineering Practice*, 2018, 70: 134-147.
- [57] Naghoosi E, Huang B, Domlan E, et al. Information transfer methods in causality analysis of process variables with an industrial application[J]. *Journal of Process Control*, 2013, 23(9): 1296-1305.
- [58] Lindner B, Auret L, Bauer M. A systematic workflow for oscillation diagnosis using transfer entropy[J]. *IEEE Transactions on Control Systems Technology*, 2020, 28(3): 908-919.
- [59] Zope K, Singhal T, Nistala S H, et al. Transfer entropy-based automated fault traversal and root cause identification in complex nonlinear industrial processes[J]. *Industrial & Engineering Chemistry Research*, 2023, 62(9): 4002-4018.
- [60] Vakorin V A, Krakovska O A, McIntosh A R. Confounding effects of indirect connections on causality estimation[J]. *Journal of Neuroscience Methods*, 2009, 184(1): 152-160.
- [61] Duan P, Yang F, Chen T W, et al. Direct causality detection via the transfer entropy approach[J]. *IEEE Transactions on Control Systems Technology*, 2013, 21(6): 2052-2066.
- [62] Landman R, Jämsä-Jounela S L. Hybrid approach to casual analysis on a complex industrial system based on transfer entropy in conjunction with process connectivity information[J]. *Control Engineering Practice*, 2016, 53: 14-23.
- [63] Ma L, Dong J, Peng K X. Root cause diagnosis of quality-related faults in industrial multimode processes using robust Gaussian mixture model and transfer entropy[J]. *Neurocomputing*, 2018, 285: 60-73.
- [64] Ma L, Dong J, Peng K X, et al. A novel data-based quality-related fault diagnosis scheme for fault detection and root cause diagnosis with application to hot strip mill process[J]. *Control Engineering Practice*, 2017, 67: 43-51.
- [65] Jiao J F, Zhen W T, Zhu W X, et al. Quality-related root cause diagnosis based on orthogonal kernel principal component regression and transfer entropy[J]. *IEEE Transactions on Industrial Informatics*, 2021, 17(9): 6347-6356.
- [66] Sharma P, Bucci D J, Brahma S K, et al. Communication network topology inference via transfer entropy[J]. *IEEE Transactions on Network Science and Engineering*, 2020, 7(1): 562-575.
- [67] Takens F. Detecting strange attractors in turbulence[C]. *Lecture Notes in Mathematics*. Berlin, Heidelberg: Springer, 1981: 366-381.
- [68] Luo L, Cheng F F, Qiu T, et al. Refined convergent cross-mapping for disturbance propagation analysis of chemical processes[J]. *Computers & Chemical Engineering*, 2017, 106: 1-16.
- [69] Wang Y L, Hu F X, Cao Y, et al. Improved CCM for variable causality detection in complex systems[J]. *Control Engineering Practice*, 2019, 83: 67-82.
- [70] Tian C, Zhao C H, Fan H D, et al. Causal network construction based on convergent cross mapping (CCM) for alarm system root cause tracing of nonlinear industrial process[J]. *IFAC-PapersOnLine*, 2020, 53(2): 13619-13624.
- [71] Tian C, Zhao C H. Single model-based analysis of relative causal changes for root-cause diagnosis in complex industrial processes[J]. *Industrial & Engineering Chemistry Research*, 2021, 60(34): 12602-12613.
- [72] Song P Y, Zhao C H, Huang B, et al. Explicit representation and customized fault isolation framework for learning temporal and spatial dependencies in industrial processes[J]. *IEEE Transactions on Neural Networks and Learning Systems*, DOI: 10.1109/TNNLS.
- [73] Chen X, Zheng J L, Zhao C H, et al. Full decoupling high-order dynamic mode decomposition for advanced static and dynamic synergetic fault detection and

- isolation[J]. *IEEE Transactions on Automation Science and Engineering*, 2022, 99: 1-15.
- [74] Arnold A, Liu Y, Abe N. Temporal causal modeling with graphical granger methods[C]. *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM, 2007: 66-75.
- [75] Song P Y, Zhao C H. Sparse adjacency forecasting and its application to efficient root cause diagnosis of process faults[J]. *IFAC-PapersOnLine*, 2021, 54(3): 439-444.
- [76] Tibshirani R. Regression shrinkage and selection via the lasso[J]. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 1996, 58(1): 267-288.
- [77] Chen H S, Yan Z B, Yao Y, et al. Systematic procedure for granger-causality-based root cause diagnosis of chemical process faults[J]. *Industrial & Engineering Chemistry Research*, 2018, 57(29): 9500-9512.
- [78] Schulz E, Speekenbrink M, Krause A. A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions[J]. *Journal of Mathematical Psychology*, 2018, 85: 1-16.
- [79] Chen J W, Zhao C H, Sun Y X. Sparse causal residual neural network for linear and nonlinear concurrent causal inference and root cause diagnosis[C]. *The 16th International Conference on Control, Automation, Robotics and Vision*. Shenzhen, 2021: 1182-1187.
- [80] Chen J W, Zhao C H. Multi-lag and multi-type temporal causality inference and analysis for industrial process fault diagnosis[J]. *Control Engineering Practice*, 2022, 124: 105174.
- [81] Nicholson William B, Ines W, Jacob B, et al. High dimensional forecasting via interpretable vector autoregression[J]. *Journal of Machine Learning Research*, 2020, 21: 6690-6741.
- [82] Wang S, Zhao Q, Han Y H, et al. Root cause diagnosis for complex industrial process faults via spatiotemporal coalescent based time series prediction and optimized Granger causality[J]. *Chemometrics and Intelligent Laboratory Systems*, 2023, 233: 104728.
- [83] Zhou Y J, Xu K, He F. Root cause diagnosis in multivariate time series based on modified temporal convolution and multi-head self-attention[J]. *Journal of Process Control*, 2022, 117: 14-25.
- [84] Ma L, Dong J, Peng K X. A practical root cause diagnosis framework for quality-related faults in manufacturing processes with irregular sampling measurements[J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 71: 1-9.
- [85] Qin K, Chen L, Shi J T, et al. Root cause analysis of industrial faults based on binary extreme gradient boosting and temporal causal discovery network[J]. *Chemometrics and Intelligent Laboratory Systems*, 2022, 225: 104559.
- [86] Wang S, Zhao Q, Han Y H, et al. Root cause diagnosis for process faults based on multisensor time-series causality discovery[J]. *Journal of Process Control*, 2023, 122: 27-40.
- [87] Zhao C H, Huang B. A full-condition monitoring method for nonstationary dynamic chemical processes with cointegration and slow feature analysis[J]. *AICHe Journal*, 2018, 64(5): 1662-1681.
- [88] Zhao C H, Chen J H, Jing H. Condition-driven data analytics and monitoring for wide-range nonstationary and transient continuous processes[J]. *IEEE Transactions on Automation Science and Engineering*, 2021, 18(4): 1563-1574.
- [89] Mukherjee T K, Naka A. Dynamic relations between macroeconomic variables and the Japanese stock market: An application of a vector error correction model[J]. *Journal of Financial Research*, 1995, 18(2): 223-237.
- [90] Zhang H J, Peng K X, Ma L. A systematic nonstationary causality analysis framework for root cause diagnosis of faults in manufacturing processes[J]. *Control Engineering Practice*, 2023, 131: 105404.
- [91] Song P Y, Zhao C H, Huang B, et al. Sparse and time-varying predictive relation extraction for root cause quantification of nonstationary process faults[J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 71: 1-13.
- [92] Staniek M, Lehnertz K. Symbolic transfer entropy[J]. *Physical Review Letters*, 2008, 100(15): 158101.
- [93] Kugiumtzis D. Partial transfer entropy on rank vectors[J]. *The European Physical Journal Special Topics*, 2013, 222(2): 401-420.
- [94] Wen C M, Yan Z B, Liang Y C, et al. A control chart-based symbolic conditional transfer entropy method for root cause analysis of process disturbances[J]. *Computers & Chemical Engineering*, 2022, 164: 107902.
- [95] Duan S Y, Zhao C H, Wu M. Multiscale partial symbolic transfer entropy for time-delay root cause diagnosis in nonstationary industrial processes[J]. *IEEE Transactions on Industrial Electronics*, 2023, 70(2): 2015-2025.
- [96] Landman R, Jämsä-Jounela S L. Fault propagation analysis by implementing nearest neighbors method using process connectivity[J]. *IEEE Transactions on Control Systems Technology*, 2019, 27(5): 2058-2067.

- [97] Zhu Q X, Luo Y, He Y L. Novel multiblock transfer entropy based Bayesian network and its application to root cause analysis[J]. *Industrial & Engineering Chemistry Research*, 2019, 58(12): 4936-4945.
- [98] Luo Y, Gopaluni B, Xu Y, et al. A novel approach to alarm causality analysis using active dynamic transfer entropy[J]. *Industrial & Engineering Chemistry Research*, 2020, 59(18): 8661-8673.
- [99] Dong J, Cao K R, Peng K X. Hierarchical causal graph-based fault root cause diagnosis and propagation path identification for complex industrial process monitoring[J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72: 1-11.
- [100] Ma L, Dong J, Peng K X. A novel key performance indicator oriented hierarchical monitoring and propagation path identification framework for complex industrial processes[J]. *ISA Transactions*, 2020, 96: 1-13.
- [101] Yu W K, Zhao C H. Sparse exponential discriminant analysis and its application to fault diagnosis[J]. *IEEE Transactions on Industrial Electronics*, 2018, 65(7): 5931-5940.
- [102] Van den Kerkhof P, Vanlaer J, Gins G, et al. Analysis of smearing-out in contribution plot based fault isolation for Statistical Process Control[J]. *Chemical Engineering Science*, 2013, 104: 285-293.
- [103] 赵春晖, 胡赟昀, 郑嘉乐, 等. 数据驱动的燃煤发电装备运行工况监控——现状与展望[J]. *自动化学报*, 2022, 48(11): 2611-2633.
(Zhao C H, Hu Y Y, Zheng J L, et al. Data-driven operating monitoring for coal-fired power generation equipment: The state of the art and challenge[J]. *Acta Automatica Sinica*, 2022, 48(11): 2611-2633.)
- [104] Zhao C H. Perspectives on nonstationary process monitoring in the era of industrial artificial intelligence[J]. *Journal of Process Control*, 2022, 116: 255-272.
- [105] Ge Y, Jiang G F, Ding M, et al. Ranking metric anomaly in invariant networks[J]. *ACM Transactions on Knowledge Discovery from Data*, 2014, 8(2): 1-30.
- [106] Cheng W, Zhang K, Chen H F, et al. Ranking causal anomalies via temporal and dynamical analysis on vanishing correlations[C]. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM, 2016: 805-814.
- [107] Ni J C, Cheng W, Zhang K, et al. Ranking causal anomalies by modeling local propagations on networked systems[C]. *2017 IEEE International Conference on Data Mining*. New Orleans, 2017: 1003-1008.
- [108] Chai Z, Zhao C H, Huang B. Multisource-refined transfer network for industrial fault diagnosis under domain and category inconsistencies[J]. *IEEE Transactions on Cybernetics*, 2022, 52(9): 9784-9796.
- [109] Chai Z, Zhao C H. Enhanced random forest with concurrent analysis of static and dynamic nodes for industrial fault classification[J]. *IEEE Transactions on Industrial Informatics*, 2020, 16(1): 54-66.
- [110] Yu W K, Zhao C H. Broad convolutional neural network based industrial process fault diagnosis with incremental learning capability[J]. *IEEE Transactions on Industrial Electronics*, 2020, 67(6): 5081-5091.
- [111] Chai Z, Zhao C H. A fine-grained adversarial network method for cross-domain industrial fault diagnosis[J]. *IEEE Transactions on Automation Science and Engineering*, 2020, 17(3): 1432-1442.
- [112] Weber P, Jouffe L. Reliability modelling with dynamic Bayesian networks[J]. *IFAC Proceedings Volumes*, 2003, 36(5): 57-62.
- [113] Powers G J, Tompkins F C Jr. Fault tree synthesis for chemical processes[J]. *AIChE Journal*, 1974, 20(2): 376-387.
- [114] Kramer M A, Palowitch B L. A rule-based approach to fault diagnosis using the signed directed graph[J]. *AIChE Journal*, 1987, 33(7): 1067-1078.
- [115] Sussman G J, Steele G L. Constraints — A language for expressing almost-hierarchical descriptions[J]. *Artificial Intelligence*, 1980, 14(1): 1-39.
- [116] Zhang Z D, Dong F L. Fault detection and diagnosis for missing data systems with a three time-slice dynamic Bayesian network approach[J]. *Chemometrics and Intelligent Laboratory Systems*, 2014, 138: 30-40.
- [117] Yue W C, Chen X F, Gui W H, et al. A knowledge reasoning fuzzy-Bayesian network for root cause analysis of abnormal aluminum electrolysis cell condition[J]. *Frontiers of Chemical Science and Engineering*, 2017, 11(3): 414-428.
- [118] Gharahbagheri H, Imtiaz S A, Khan F. Root cause diagnosis of process fault using KPCA and Bayesian network[J]. *Industrial & Engineering Chemistry Research*, 2017, 56(8): 2054-2070.
- [119] Kumari P, Bhadriraju B, Wang Q S, et al. A modified Bayesian network to handle cyclic loops in root cause diagnosis of process faults in the chemical process industry[J]. *Journal of Process Control*, 2022, 110: 84-98.
- [120] Gao J, Zhao C H. A multi-level Bayesian network based on causality analysis for fault diagnosis of nonstationary

- processes[C]. 2018 Chinese Automation Congress. Xi'an, 2019: 1626-1631.
- [121] Zhang Q. Dynamic uncertain causality graph for knowledge representation and reasoning: Continuous variable, uncertain evidence, and failure forecast[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2015, 45(7): 990-1003.
- [122] Ruijters E, Stoelinga M. Fault tree analysis: A survey of the state-of-the-art in modeling, analysis and tools[J]. *Computer Science Review*, 2015, 15/16: 29-62.
- [123] Ulerich N H, Powers G J. On-line hazard aversion and fault diagnosis in chemical processes: The digraph fault-tree method[J]. *IEEE Transactions on Reliability*, 1988, 37(2): 171-177.
- [124] Lavasani S M, Zendegani A, Celik M. An extension to fuzzy fault tree analysis (FFTA) application in petrochemical process industry[J]. *Process Safety and Environmental Protection*, 2015, 93: 75-88.
- [125] Purba J H. A fuzzy-based reliability approach to evaluate basic events of fault tree analysis for nuclear power plant probabilistic safety assessment[J]. *Annals of Nuclear Energy*, 2014, 70: 21-29.
- [126] Volkanovski A, Čepin M, Mavko B. Application of the fault tree analysis for assessment of power system reliability[J]. *Reliability Engineering & System Safety*, 2009, 94(6): 1116-1127.
- [127] Chiacchio F, Famoso F, D'Urso D, et al. Dynamic performance evaluation of photovoltaic power plant by stochastic hybrid fault tree automaton model[J]. *Energies*, 2018, 11(2): 306.
- [128] Tanaka H, Fan L T, Lai F S, et al. Fault-tree analysis by fuzzy probability[J]. *IEEE Transactions on Reliability*, 1983, R-32(5): 453-457.
- [129] Senol Y E, Aydogdu Y V, Sahin B, et al. Fault tree analysis of chemical cargo contamination by using fuzzy approach[J]. *Expert Systems with Applications*, 2015, 42(12): 5232-5244.
- [130] Yazdi M, Korhan O, Daneshvar S. Application of fuzzy fault tree analysis based on modified fuzzy AHP and fuzzy TOPSIS for fire and explosion in the process industry[J]. *International Journal of Occupational Safety and Ergonomics*, 2020, 26(2): 319-335.
- [131] Dugan J B, Bavuso S J, Boyd M A. Fault trees and sequence dependencies[C]. *Annual Proceedings on Reliability and Maintainability Symposium*. Los Angeles, 2002: 286-293.
- [132] Baek S, Heo G. Application of dynamic fault tree analysis to prioritize electric power systems in nuclear power plants[J]. *Energies*, 2021, 14(14): 4119.
- [133] Khakzad N, Khan F, Amyotte P. Safety analysis in process facilities: Comparison of fault tree and Bayesian network approaches[J]. *Reliability Engineering & System Safety*, 2011, 96(8): 925-932.
- [134] Hamza Z, Abdallah T. Mapping fault tree into Bayesian network in safety analysis of process system[C]. *The 4th International Conference on Electrical Engineering*. Boumerdes, 2016: 1-5.
- [135] Gao D, Wu C G, Zhang B K, et al. Signed directed graph and qualitative trend analysis based fault diagnosis in chemical industry[J]. *Chinese Journal of Chemical Engineering*, 2010, 18(2): 265-276.
- [136] Liu Y K, Wu G H, Xie C L, et al. A fault diagnosis method based on signed directed graph and matrix for nuclear power plants[J]. *Nuclear Engineering and Design*, 2016, 297: 166-174.
- [137] Shiozaki J, Shibata B, Matsuyama H, et al. Fault diagnosis of chemical processes utilizing signed directed graphs-improvement by using temporal information[J]. *IEEE Transactions on Industrial Electronics*, 1989, 36(4): 469-474.
- [138] Han C C, Shih R F, Lee L S. Quantifying signed directed graphs with the fuzzy set for fault diagnosis resolution improvement[J]. *Industrial & Engineering Chemistry Research*, 1994, 33(8): 1943-1954.
- [139] He B, Chen T, Yang X H. Root cause analysis in multivariate statistical process monitoring: Integrating reconstruction-based multivariate contribution analysis with fuzzy-signed directed graphs[J]. *Computers & Chemical Engineering*, 2014, 64: 167-177.
- [140] Yang F, Xiao D Y, Shah S L. Signed directed graph-based hierarchical modelling and fault propagation analysis for large-scale systems[J]. *IET Control Theory & Applications*, 2013, 7(4): 537-550.
- [141] Peng D, Geng Z Q, Zhu Q X. A multilogic probabilistic signed directed graph fault diagnosis approach based on Bayesian inference[J]. *Industrial & Engineering Chemistry Research*, 2014, 53(23): 9792-9804.
- [142] 刘仲宇. 基于多级流模型的故障诊断技术研究[D]. 杭州: 浙江大学, 2008.
(Liu Z Y. Research on fault diagnosis technology based on multistage flow model[D]. Hangzhou: Zhejiang University, 2008.)
- [143] Moore R L. G2: A software platform for intelligent process control[C]. *Proceedings of the 1991 IEEE International Symposium on Intelligent Control*. Arlington, 1991: 1-5.

- [144] Wu J, Zhang L B, Liang W, et al. A novel failure mode analysis model for gathering system based on multilevel flow modeling and HAZOP[J]. *Process Safety and Environmental Protection*, 2013, 91(1/2): 54-60.
- [145] Nielsen E K, Gofuku A, Zhang X X, et al. Causality validation of multilevel flow modelling[J]. *Computers & Chemical Engineering*, 2020, 140: 106944.
- [146] Lind M, Yoshikawa H, Jørgensen S B, et al. Multilevel flow modeling of monju nuclear power plant[J]. *International Journal of Nuclear Safety and Simulation*, 2011, 2(3): 275-285.
- [147] Zheng X, Aragam B, Ravikumar P K, et al. Dags with no tears: Continuous optimization for structure learning[C]. *Advances in Neural Information Processing Systems*. Montréal: MIT Press, 2018: 9472-9483.
- [148] Yu Y, Chen J, Gao T, et al. DAG-GNN: DAG structure learning with graph neural networks[J/OL]. 2019, arXiv: 1904.10098.
- [149] Pamfil R, Sriwattanaworachai N, Desai, et al. DYNOTEARS: Structure learning from time-series data[J/OL]. 2020, arXiv: 2002.00498.
- [150] Luo Y N, Peng J, Ma J Z. When causal inference meets deep learning[J]. *Nature Machine Intelligence*, 2020, 2(8): 426-427.
- [151] Yang M Y, Liu F R, Chen Z T, et al. CausalVAE: Disentangled representation learning via neural structural causal models[C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, 2021: 9588-9597.
- [152] Jiang H L, Patwardhan R, Shah S L. Root cause diagnosis of plant-wide oscillations using the adjacency matrix[J]. *IFAC Proceedings Volumes*, 2008, 41(2): 13893-13900.
- [153] Jiang H L, Patwardhan R, Shah S L. Root cause diagnosis of plant-wide oscillations using the concept of adjacency matrix[J]. *Journal of Process Control*, 2009, 19(8): 1347-1354.
- [154] Feng L J, Zhao C H. Fault description based attribute transfer for zero-sample industrial fault diagnosis[J]. *IEEE Transactions on Industrial Informatics*, 2021, 17(3): 1852-1862.
- [155] Li B, Zhao C. Federated zero-shot industrial fault diagnosis with cloud-shared semantic knowledge base[J]. *IEEE Internet of Things Journal*, DOI: 10.1109/JIOT.2023.3243401.

作者简介

赵春晖(1979—),女,教授,博士生导师,从事面向工业应用的人工智能方法、大数据分析与应用等研究, E-mail: chhzhao@zju.edu.cn;

宋鹏宇(1998—),男,博士生,从事工业大数据分析、工业过程故障检测、故障根因追溯等研究, E-mail: pysong@zju.edu.cn.



特邀专家 赵春晖,分别于2003年、2009年获东北大学学士学位和博士学位。浙江大学求是特聘教授,国家杰出青年科学基金获得者。主要研究方向为面向工业应用的人工智能方法。已在国际权威期刊上发表(第1作者/通讯作者)高水平SCI研究论文190多篇,出版3部专著,1部大数据专业课教材,授权发明专利60余项。担任国家自然科学基金委杰青项目负责人。主持国家自然科学基金、国家重点研发计划课题、省级项目和企业合作项目等科研项目20余项。曾获教育部自然科学一等奖、浙江省科技进步二等奖、浙江省首届青年科技奖、中国自动化学会自然科学一等奖、中国自动化学会首届青年女科学家奖等。担任三种国际期刊编委和国内多种期刊的编委。

专家寄语 曾经在东北大学生活了9年多,母校给了我一个专心求学的良好环境,让我打好科研基础,开心快乐地成长。今天,是母校百年华诞,在这个历史性的时刻,向母校致以最诚挚的祝福。母校,这个光彩夺目的名称,将像一颗灿烂的明星,永远引领东大学子不断前进!