

基于反事实推理的柴油机故障诊断

李 恒¹, 黄守金¹, 丁 勇¹, 陆宁云^{1,2}

(1. 南京航空航天大学 自动化学院, 南京 211106;

2. 机械结构力学及控制国家重点实验室, 南京 211106)

摘要: 针对柴油机故障因果关系的隐蔽性、网络模型转换为诊断模型过程中的不确定性, 以及传统反事实推理框架在处理复杂系统时推理效率低的问题, 提出了一种基于 Leaky Noisy-OR 假设和孪生网络的因果推理模型, 以明确定量因果关系; 首先, 构建故障因果网络, 并融入 Leaky Noisy-OR 假设和隐变量节点, 以表示不确定性因素, 构建了一个适用于反事实推理的因果推理模型; 其次, 将该模型转换成孪生网络, 并采取反事实干预措施进行简化, 提高了推理效率; 最后, 定义了故障充分因和故障必要因这两个指标, 以定量衡量故障原因和故障征兆之间的因果关系强度, 并且在推理的准确性上构成了双重保障; 实验结果表明, 该方法克服了传统相关性分析的局限性, 为复杂系统故障诊断提供了可解释的因果推理范式。

关键词: 柴油机; 因果学习; 反事实推理; 故障诊断; Leaky Noisy-OR 假设; 孪生网络

Fault Diagnosis in Diesel Engines Based on Counterfactual Reasoning

LI Heng¹, HUANG Shoujin¹, DING Yong¹, LUN Ningyun^{1,2}

(1. College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China;

2. State Key Laboratory of Mechanics and Control of Mechanical Structures, Nanjing 211106, China)

Abstract: A causal inference model based on Leaky Noisy-OR assumption and twin network is proposed to address the hidden causality of diesel engine faults, uncertainties in the transition from network model to diagnostic model, and low inference efficiency in traditional counterfactual inference frameworks when dealing with complex systems. Firstly, a fault causal network is constructed, and a causal inference model suitable for counterfactual reasoning is built by incorporating the Leaky Noisy-OR assumption and hidden variable nodes into the model to represent uncertainty factors. Secondly, the model is converted into a twin network and simplified with counterfactual interventions, which improves the reasoning efficiency. Finally, fault sufficient cause and fault necessary cause are defined to quantitatively measure the strength of causal relationship between the cause and sign of faults, and constitute a double guarantee in terms of reasoning accuracy. Experimental results demonstrate that this method overcomes the limitations of conventional correlation analysis and provides an interpretable causal reasoning paradigm for complex system fault diagnosis.

Keywords: diesel engines; causal learning; counterfactual reasoning; Leaky Noisy-OR assumption; twin network

0 引言

大功率船用柴油机作为船舶动力系统中的核心组成部分, 其性能、效率以及排放的优化需求日益迫切。高压共轨系统已成为满足现代船用柴油机高性能、高效率和环保需求的关键技术^[1], 然而该系统结构复杂, 内部部件耦合紧密且故障传递规律复杂, 单个部件性能的微小变化可能对整个系统性能产生显著影响。其中, 高压

油泵和喷油器作为共轨系统的两个核心部件, 其部件间故障原因与故障征兆之间的因果关系隐蔽性强、影响因素关联复杂、诊断难度较高。因此, 研究故障因果关系对复杂系统的故障诊断研究具有重要意义。

典型的柴油机故障诊断方法包括振动分析法^[2]、声发射分析^[3-4]、压力分析^[5]、油和磨损颗粒分析^[6]、热力参数分析^[7]等方法; 其中典型部件故障诊断方案大都是依次进行信号采集、特征提取、特征降维、故

收稿日期: 2025-03-14; 修回日期: 2025-04-08。

基金项目: 国家自然科学基金(62273176)。

作者简介: 李 恒(1999-), 女, 硕士研究生。

通讯作者: 陆宁云(1978-), 女, 博士, 教授。

引用格式: 李 恒, 黄守金, 丁 勇, 等. 基于反事实推理的柴油机故障诊断[J]. 计算机测量与控制, 2025, 33(5): 117-124.

障诊断等步骤。目前本文研究的柴油机装备故障诊断大多数研究仍停留在相关性分析层面, 缺乏对因果机制的解释。而因果学习的核心思想正是突破了这种传统统计的相关性思维, 建立在因果逻辑与干预主义基础上。目前, 因果学习的应用已逐步深入医学、农业、智能制造以及其他领域。其对因果关系的挖掘方法已显示出显著的研究价值和实践效果。在医学诊断中, 文献 [8] 利用因果学习识别引发患病症状的潜在疾病因素, 将诊断过程转化为反事实推理任务, 并构建了一种反事实推理的算法框架, 从而更精准地解释病人症状的成因; 在农业领域中, 文献 [9] 基于因果学习的方法剖析农业环境损害的因果关系, 结合知识图谱和数据库的研究, 解决因果关系“鉴定难”的问题; 在计算机视觉的应用中, 文献 [10] 利用因果学习来增强深度学习模型的鲁棒性, 避免数据偏差导致虚假关联, 并建立基于因果结构的深度学习框架; 在电力系统故障诊断中, 文献 [11] 通过因果学习将配电网中各类设备关联关系划分为 3 种基本因果类型, 提供因果网络实现逻辑逆向推理, 准确定位故障位置; 在通信系统中, 文献 [12] 利用逻辑树组法并结合因果学习构建候选诊断模型, 并且运用概率推理方法优化系统故障的诊断策略。特别在医学诊断领域, 因果学习的应用提供了许多典型范例。装备故障诊断与医学诊断具有显著的同构性。医生建立症状与疾病之间的病症因果链对患者进行诊断。故障诊断同样需要通过故障征兆追溯产生这些征兆的根本故障模式及故障原因。

然而, 目前国内外的研究多集中于单一部件的故障分析, 并且仍停留在相关性分析层面, 缺乏对深层因果机制的解释, 该问题的突破需要依赖于因果发现与学习理论的进一步发展。特别是在处理柴油机故障诊断中, 将故障因果网络转换成诊断模型的过程中, 考虑到高压油泵和喷油器的故障诊断存在的复杂因果关系以及知识和证据的不确定性, 本文将尝试将因果学习引入装备故障诊断领域, 通过构建部件故障原因与征兆之间的故障因果链, 进行复杂系统的故障诊断研究, 有利于现场工程师诊断维修。

1 因果推理模型构建

1.1 结构性因果模型基本原理

结构性因果模型 (SCM, structural causal model) 由 Pearl 教授提出^[13], 为解决复杂系统中的多因素概率因果推理问题提供了数学化的框架, 也被称为功能因果模型或结构方程模型^[14-15]。因其能够在建模过程中引入外部干预的概念^[16], 为研究因果关系提供了一种新方法。SCM 在贝叶斯网络的基础上进一步扩展了

外部干预的概念, 通过引入潜在变量或者噪声变量, 能够更全面地理解和分析复杂系统中变量之间的因果关系^[17-18]。这一模型通过将变量组织成一个有向无环图 (DAG, directed acyclic graph): G 。其中, 每个节点都代表着一个变量; 节点与节点之间的有向边表示由原因到结果的因果关系; 未观察到的外生变量作为外部干预, 模拟模型外部的所有影响因素。SCM 的结构示意图如图 1 所示。其中 X, Y, Z 为观测变量, $\{X, Y, Z\} \in V$; U_X, U_Y, U_Z 为潜在变量, $\{U_X, U_Y, U_Z\} \in U$ 。

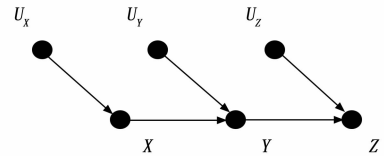


图 1 SCM 结构示意图

1.2 Leaky Noisy-OR 假设

在柴油发电机故障诊断领域, 故障征兆的影响通常被看作是独立的, 这意味着不同故障原因间的相互作用不会削弱它们对征兆的影响, 只要存在一个能引发征兆发生的故障原因发生, 征兆便会表现出来。这种原因与征兆间的关系可通过 Noisy-OR 假设^[19]来描述。但在高压共轨系统实际的诊断过程中, 由于诊断系统的复杂性、故障知识表达的模糊性以及征兆发生的不确定性等因素, 直接应用 Noisy-OR 假设建模可能面临以下挑战: 1) 故障原因与征兆之间普遍存在多对多的复杂对应关系, 这就容易导致诊断时无法准确区分不同类型的故障; 2) 逐步排查存在不准确性, 故障原因的排查可能无法一步到位。

针对这样的问题, Noisy-OR 假设的原有框架不具备处理这些未知的例外性因素的能力, 因而需将其进一步扩展为 Leaky Noisy-OR 假设, 后者通过引入失效概率的概念表征这些未知的例外性因素。

1) Noisy-OR 假设

Noisy-OR 假设可以用来表示 n 个变量 $F_1, F_2, \dots, F_n$ 与它们产生的影响变量 S 之间的一种内在关系, 其核心逻辑在于描述父节点对结果变量的影响。使用该假设时, 必须要满足以下 3 个条件: (1) 父节点之间相互独立, 子节点之间也相互独立; (2) 所有节点变量为二值节点; (3) 当每个父节点变量在除其本身外的所有父节点变量都为假时, 都能够引起结果变量 S 的发生。

Noisy-OR 假设通过上述条件明确描述了变量间的因果关系, 并以简单的逻辑图进行表示, 如图 2 所示。

2) Leaky Noisy-OR 假设

在高压共轨系统实际故障诊断中, 故障征兆不一定

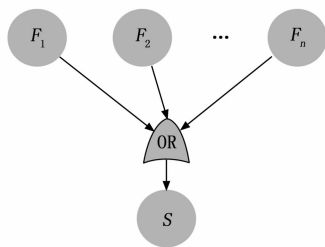


图2 Noisy-OR 假设逻辑图

完全由其对应的故障原因导致, 即子节点的发生不一定是由其父节点触发, 还可能受到一些其他未观测到的未知因素的影响, 例如环节干扰、测量误差或未考虑的故障模式, 也能够导致结果事件的发生。这种复杂情景需要 Leaky Noisy-OR 假设来进行建模。

Leaky Noisy-OR 假设在原有 Noisy-OR 假设的基础上, 增加了一个表示所有未观测到的未知因素的一个父节点 F_l , 用于综合表征系统中的例外性因素。Leaky Noisy-OR 假设声明了当所有前提条件都为假的情况下, 其结果事件仍可能以一定的泄露概率发生^[20]。其逻辑结构图见图3所示。

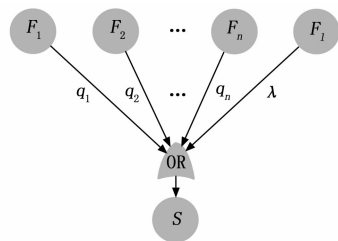


图3 Leaky Noisy-OR 假设逻辑图

基于 Leaky Noisy-OR 假设的变量 S 的概率计算表达式为:

$$P(S | F_1, F_2, \dots, F_i, \dots, F_n) = 1 - (1 - \lambda) \prod_{i=1}^n (1 - q_i) \quad (1)$$

式中, 变量 S 和 $F_1, F_2, \dots, F_n$ 发生用 1 表示, 未发生用 0 表示; q_i 表示为每个父节点与子节点之间的连接概率, 则 $1 - q_i$ 表示变量 F_i 对变量 S 的失效概率, 即当变量 F_i 为真时, 变量 S 有 $1 - q_i$ 的概率为假; λ 表示变量 S 的泄露概率, 表示其他所有未考虑到的因素导致变量 S 发生的概率, 即当变量 $F_1, F_2, \dots, F_n$ 均为假时, 变量 S 有 λ 的概率为真。

1.3 服从 Leaky Noisy-OR 假设的因果推理模型

在复杂工业系统的故障诊断领域, 准确建立因果推理模型是获得故障原因与故障征兆之间因果关系的核心环节。为了克服传统因果模型在处理不确定性因素方面的局限, 本研究提出将 Leaky Noisy-OR 假设与 SCM 框架有机融合, 分别单独评估每个父节点对其子节点的失效概率和每个子节点的泄露概率, 减少所需

赋值数量。

此外, 该方法还通过独立量化假设中两类关键概率参数实现对复杂系统的精准建模^[21]。这两类概率参数分别为失效概率和泄露概率。其中, 失效概率用于量化模型在概率权重赋值过程中对每个父节点对子节点因果强度的独立评估; 泄露概率用于描述随机性因素, 以及环境扰动可能对节点状态产生的潜在影响。通过将 Leaky Noisy-OR 假设融入因果推理模型中, 能够对故障机理中的因果关系与现象之间的随机性和不确定性影响进行有效抽象化建模。

为了在 SCM 中融入 Leaky Noisy-OR 假设, 本文对故障原因和征兆的故障因果网络结构进行了调整, 如图4示例所示。

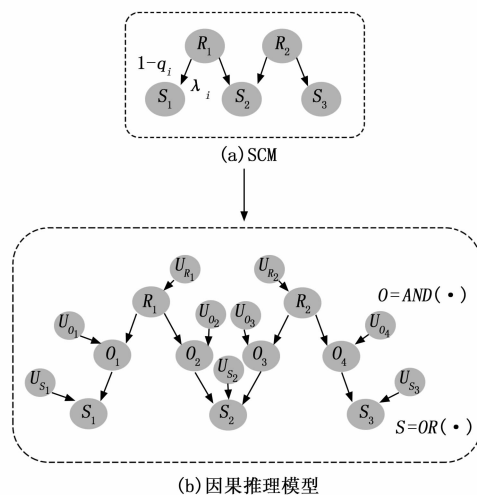


图4 因果推理模型构建流程

图中, R_i 表示故障原因, S_i 表示故障征兆, O_i 表示故障原因 R_i 对故障征兆 S_i 产生影响的激励状态将与或门逻辑应用在 SCM 的结构性等式中, U_i 表示外部隐变量。

1) 引入隐变量

外部隐变量节点 U 和激励节点 O 作为隐变量引入原有的 SCM 中。在图3(1)中引入了3种外部隐变量节点, 分别为 U_R 、 U_O 和 U_S , 并分别对应先验概率、失效概率和泄露概率所引发的特殊例外事件。根据这些隐变量节点在推理过程中发挥的不同作用和产生的不同影响, 将其划分为两类节点: 充分性例外节点和必要性例外节点。

充分性例外节点 U_O 的含义在于, 当故障原因已经发生的情况下, U_O 发生即 ($U_O = 1$) 必然会导致故障征兆的出现。因此, 其发生的概率 $P(U_O = 1)$ 可以视为 Leaky Noisy-OR 假设中的失效概率, 即 $1 - q_i$ 。

另一方面, 必要性例外节点 U_R 和 U_S 的作用在于, 当故障原因或者故障征兆未发生时, 相对应的必要性例

外节点必定不会被激活。其中,根节点的必要性例外节点 U_R 可以被视为故障先验概率;而故障征兆的必要性例外节点 U_S 发生的概率 $P(U_S=1)$ 则等同于 Leaky Noisy-OR 假设中的泄露概率 λ_i 。

2) 构建结构性等式

在构建故障诊断的因果推理模型时,通过引入激励节点并采用与或门逻辑机制^[22]。

激励节点的状态由故障原因和充分性例外节点的联合作用所决定,只有当这两者同时满足条件时,激励才会被触发。数学表达式为:

$$O = \text{AND}(R, U_o) = \begin{cases} 1, R = 1, U_o = 1 \\ 0, \quad \text{其他} \end{cases} \quad (2)$$

在征兆节点的建模过程中,综合考虑激励节点与必要性例外节点的双重影响,当激励节点或必要性例外节点中的任意一个发生时,征兆即会出现。数学表达式为:

$$S = \text{OR}(O_i, U_s) = \begin{cases} 0, \sum O_i = 0, U_s = 0 \\ 1, \quad \text{其他} \end{cases} \quad (3)$$

此外,对于根节点的处理,其状态仅与必要性例外节点 U_R 存在直接关联,数学表达式为:

$$R = \begin{cases} 1, U_R = 1 \\ 0, U_R = 1 \end{cases} \quad (4)$$

3) 确定隐变量概率权重值

为了确保隐变量概率权重赋值的科学性和可操作性,本文提出了一个基于专家经验的隐变量初始概率权重赋值框架。本文严格遵循统计独立性假设,认为各故障模式、故障原因以及故障征兆之间相互独立,不存在直接的影响。该框架采用以下方法:将概率权重等级划分为 1 至 10 级,其中 10 级代表事件发生的可能性最大,1 级则表示事件发生的可能性最小。考虑到概率的取值范围为 $[0, 1]$,本文又将概率权重等级缩小为原来的十分之一,使概率权重赋值区间为 $[0, 1]$ 。

2 基于反事实推理的故障诊断

2.1 反事实推理基本理论

反事实推理^[23]是因果推断领域中的一个核心概念,可以有效利用因果模型预测现象或者推断变量间的因果关系。这种推理方法不同于传统的推理分析,它是在已有的现实观测的基础上,通过构建假设性的场景,即建立一个与现实空间相似但在某些方面存在特定差异的假设空间,对于这样的空间称其为反事实空间,然后在反事实空间中进行分析进行干预后,以此衡量事件之间的因果关系。

在高压共轨系统实际推理的过程中,通过主动干预变量的状态并推断其对其他变量状态的影响。这种干预操作在 SCM 框架中,用数学语言 $do(X)$ 进行描述。

do 算子干预逻辑图如图 5 所示。

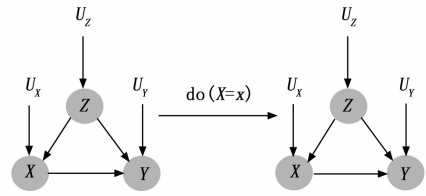


图 5 do 算子干预逻辑图

文献 [24] 中提出了一中用于计算任意反事实推理的通用算法框架,该框架基于给定 SCM 展开,其核心流程包括:1) 溯因:更新潜在隐变量分布 $P(U)$;2) 干预:对观测变量子集 X 中指定变量进行 $do(X)$ 算子干预操作;3) 预测:用干预后的模型计算 $Y=y$ 的概率。

这种经典的反事实推理框架在处理大型复杂模型时存在显著的计算挑战。主要的问题来源于第一步,即溯因步骤中隐变量后验概率计算的复杂度。当模型规模扩大,潜在的隐变量节点数目随之增加,因而需要计算的后验概率的数量也呈增长趋势。尤其在多层次、多变量的因果结构中,这种对应关系的复杂性进一步加剧了后验概率计算的困难,显著降低后验概率推理的及时性和准确性。因此,如何优化隐变量后验概率的计算量以减轻计算负担、提升反事实推理效率成为因果推断的重要问题之一。

2.2 基于孪生网络的反事实推理定量研究

针对经典的三步法反事实推理框架中存在的计算复杂度高、推理效率较低的问题,提出了一个基于孪生网络^[25-26]的反事实推理框架。

孪生网络作为一种创新性结构设计的网络模型,具有独特的架构设计。与传统三步法框架不同,孪生网络的核心在于其采用了一种更为高效的推理机制,通过对单个“孪生”因果推理模型直接执行反事实查询,巧妙地避免复杂度高的溯因步骤下后验概率的计算,显著降低计算负担,从而极大地提升了推理效率。

本文提出的基于孪生网络的反事实推理新框架,通过 3 个关键步骤实现对复杂系统因果关系的定量研究:1) 构建孪生网络;2) 简化孪生网络;3) 计算反事实推理指标。为了有效量化模型中变量之间的因果解释强度,以及双重保障推理的准确性,本研究设置了两个衡量因果关系程度的指标:故障充分因和故障必要因。

2.1.1 孪生网络构建

孪生网络核心设计在于将传统的单一因果网络扩展为一个包含事实空间和反事实空间的复合网络,并且原始网络中的每个节点均与其孪生的副本变量共享相同的外部隐变量。其中,原始网络代表真实的事实空间,即

表征实际观测到的事实状态; 其孪生副本网络则代表假设的反事实空间, 即表征在特定干预下可能出现的反事实状态。双网络结构严格对齐事实空间和反事实空间的拓扑, 确保了因果关系的基本框架在不同空间中保持一致。同时, 这种双网络的映射关系也保证了因果路径的完整性。

本文构建的因果推理模型采用了基于三层 DAG 结构的 SCM, 以图 6 中的故障因果网络为基础, 介绍孪生网络的构建流程。为了确保模型的可解释性和计算可行性, 本研究假设了严格的层内无连接约束, 即属于同一层级的节点之间不存在直接的因果关系, 没有有向边连接。

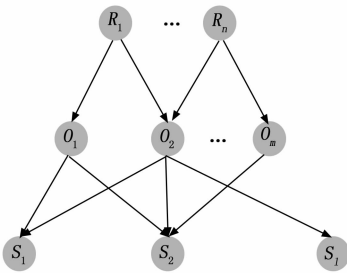


图 6 故障因果网络

在图 6 的三层故障因果网络的基础上, 构建孪生网络, 其模型构建示意图如图 7 所示。

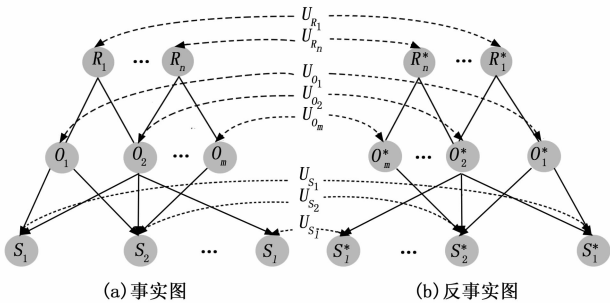


图 7 孪生网络模型构建示意图

在构建孪生网络的过程中, 复制图 6 中的故障因果网络 (即图 7a), 生成孪生副本 (即图 7b), 构成孪生网络的双网络结构。图中, 左侧的网络称其为事实图; 右侧网络将在反事实空间称其为反事实图。新增的反事实图中的孪生节点是由事实图中的相应节点映射过来的, 这些节点通过节点表示符号右上角加星号 (*) 与原始节点进行区分, 互为对偶变量。此外, 事实图与反事实图之间的关系通过共享的外部变量进行连接。在示意图中, 这种联系以虚线形式表现。

2.1.2 孪生网络化简

孪生网络中, 一旦应用了反事实干预措施, 就可以通过节点合并消除冗余节点和关系, 大大简化孪生网络

的结构^[27], 从而提升模型的可操作性和计算效率。简化孪生网络的具体步骤如下。

1) 合并顶层

在因果推理模型中, 某一节点的取值是由其父节点以及其外部隐变量节点的取值共同决定的。当一个变量的所有父节点以及其外部隐变量的瞬时值已知时, 该节点的取值则为一个固定值。因此, 如果两个节点的外部隐变量和父节点变量完全相同, 那么这两个节点就可以合并为一个节点。对于图 7 中的节点, 事实图和反事实图中的顶层层无父节点, 并且每一对对偶节点共享同一个外部隐变量, 因此可将所有节点对 合并为单一节点。图 8 为合并顶层层示意图。

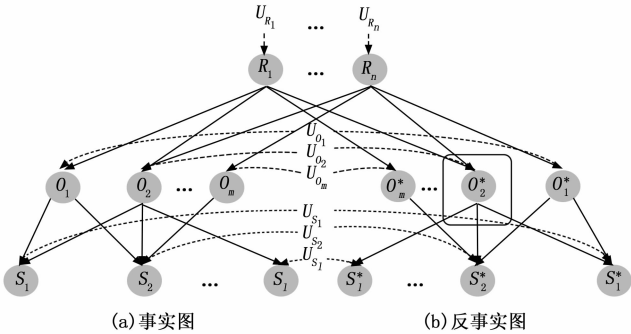


图 8 孪生网络合并顶层层

2) 合并第二层

在图 8 完成顶层层节点的合并后, 孪生网络可以再进一步简化, 尤其是考虑对单一节点进行反事实干预的情况下。图 8 中, 假设在反事实图中选择某一个节点进行干预 (本例中为节点)。根据反事实推理中的 do 算子运算, 切断原有因果网络中所有指向该节点的有向边。这一化简过程减少了网络中不必要的冗余关系。合并第二层节点后的孪生网络如图 9 所示 (为了简化图示, 图中不再显示外部隐变量节点)。

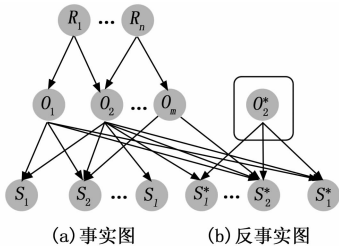


图 9 孪生网络合并第二层

3) 合并第三层

继续对第三层节点进行合并。观察网络中第三层的节点, 如果某些节点的变化不会影响反事实推理, 或者与反事实查询无关, 可将其删除。图 10 为合并第三层示意图。假设图中观测到的变量仅为 S_1 时, 除节点 S_1 , 其余节点 O_2^* 的子节点观测为不发生, 被视作为冗

余节点, 可将其删去。

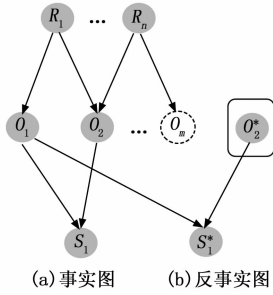


图 10 孪生网络合并第三层

通过这种逐层合并节点的过程, 最终可以得到一个更加紧凑的孪生网络。

2.1.3 反事实推理指标计算

因果性解释包括充分性和必要性, 将充分因和必要因定义为量化一般模型中因果解释的两个指标。利用 do 算子来表示这两个指标的目标函数, 从而量化原因对结果的解释程度。充分因和必要因两个指标从充分性和必要性两方面衡量故障原因和故障征兆之间因果关系的差异。充分因适用于事件尚未发生, 如果采取某些干预, 当前状态会发生改变的的概率; 必要因适用于事件已经发生, 通过反事实推理得到如果原因不发生, 征兆本不发生的概率。在推理的准确性上, 充分因和必要因共同构成双重保障, 两个指标共同作用, 避免单一指标可能带来的不准确性。

1) 故障充分因:

一个观察到的征兆往往可能由多个原因所导致, 这些原因相互竞争并共同构成对观察结果的不同解释。这一现象可以通过“因果充分性假设”来进行概括。该假设认为, 任何单一的故障都能作为其相关征兆的充分原因。根据这一假设, 我们可以通过确定一种故障原因引起征兆的可能性来简化推理过程。

在因果推理模型中, Leaky Noisy-OR 模型本身常常遵循因果充分性假设, 因为在该模型中, 每个原因都足以单独引发其相关的征兆。Leaky Noisy-OR 假设的核心思想是, 某一原因通过其影响路径, 独立地解释该征兆的发生。在此模型中, 征兆节点 S 的产生可以由其父节点中任一原因通过激活其对应的 OR 函数所引起。

基于反事实推理的故障充分因定义: 在已知当前征兆发生的条件下, 如果能够引起征兆发生的其他可能的故障原因都不出现, 原有发生的征兆仍然发生的期望是故障的充分因。表达式为:

$$E_{\text{suf}}(O_k, \epsilon) =$$

$$\sum_{S^*} |S^*| P\{S^* | \epsilon, do[Pa(S) \setminus O_k = F]\} \quad (5)$$

式中, $E_{\text{suf}}(O_k, \epsilon)$ 表示针对目标节点 O_k 干预得到的

期望充分性。公式中的求和是对所有可能的反事实证据中征兆 S^* 状态的求和。 $|S^*|$ 表示反事实证据中的一组当前已发生的征兆的基数。 $Pa(S)$ 表示能够引起该征兆发生的所有父节点集合; $Pa(S) \setminus O_k$ 表示除了指定的干预节点 O_k 外, 能够引起该征兆发生的所有父节点集合; $do[Pa(S) \setminus O_k = F]$ 则表示在反事实推理的过程中, 采用 do 算子进行干预, 将所有属于 $Pa(S) \setminus O_k$ 的父节点的值固定为 F , 以此排除除节点 O_k 以外的其他因素对征兆 ϵ 的影响。 ϵ 则表示所有事实证据的集合。

2) 故障必要因:

基于反事实推理的故障必要因作为故障充分因的补充, 其定义为: 在已知当前征兆发生的条件下, 如果故障原因没有发生, 征兆本可以不发生的期望是故障的必要因。表达式为:

$$E_{\text{dis}}(O_k, \epsilon) = \sum_{S^*} |S \setminus S^*| P[S^* | \epsilon, do(O_k = F)] \quad (6)$$

式中, $E_{\text{dis}}(O_k, \epsilon)$ 表示针对目标节点 O_k 干预得到的期望必要性。 $|S \setminus S^*|$ 表示在事实证据中存在, 但在反事实证据中不存在的征兆的基数。 $do(O_k = F)$ 表示在反事实推理干预中设置所有节点 O_k 的值为 F 。

3 实验结果与分析

3.1 实验步骤和方法

本实验以上海某研究所某 16 缸 V 型船用大功率柴油机共轨系统中的高压油泵和喷油器为研究对象, 采用结合故障树 (FTA, fault tree analysis) 和失效模式与影响分析 (FMEA, failure mode and effects analysis) 的集成方法获取领域知识, 得到这两个部件的故障原因和故障征兆。对发生的故障征兆 S_7 和未发生的故障征兆 $S_2, S_3, S_4, S_5, S_6, S_8, S_9, S_{10}, S_{11}$ 应用反事实推理。

3.2 实验数据

故障原因、征兆以及其之间的概率权重值由该研究所共轨系统运行线路的现场工程师基于日常故障数据积累、多年工作经验以及参考该型号柴油机《EMERGENCY GENERATING SET DIESEL ENGINE》赋值, 如表 1 所示。

构建故障原因与征兆的因果推理模型如图 11 所示 (为使故障原因与征兆的因果关系表示更加紧凑与清晰, 图中未绘制激励节点)。并对该模型进行基于孪生网络的反事实推理定量研究, 完成孪生网络构建、孪生网络化简和反事实推理指标计算三步骤后, 得到推理计算结果如表 2 所示。并对推理计算结果进行对比, 对比结果如图 12 所示。

表1 故障因果网络中节点的概率权重值

必要性例外节点		充分性例外节点	
$P(U_{R_0}=1)=0.4$	$P(U_{S_5}=1)=0.4$	$P(O(R_0,S_2)=0)=0.2$	$P(O(R_3,S_3)=0)=0.3$
$P(U_{R_0}=1)=0.4$	$P(U_{S_5}=1)=0.4$	$P(O(R_0,S_6)=0)=0.4$	$P(O(R_3,S_1)=0)=0.2$
$P(U_{R_0}=1)=0.4$	$P(U_{S_5}=1)=0.4$	$P(O(R_0,S_7)=0)=0.6$	$P(O(R_3,S_1)=0)=0.2$
$P(U_{R_0}=1)=0.4$	$P(U_{S_5}=1)=0.2$	$P(O(R_1,S_8)=0)=0.2$	$P(O(R_4,S_0)=0)=0.2$
$P(U_{R_0}=1)=0.6$	$P(U_{S_5}=1)=0.4$	$P(O(R_1,S_7)=0)=0.2$	$P(O(R_4,S_7)=0)=0.2$
$P(U_{R_0}=1)=0.6$	$P(U_{S_5}=1)=0.4$	$P(O(R_2,S_0)=0)=0.3$	$P(O(R_5,S_1)=0)=0.2$
	$P(U_{S_5}=1)=0.2$	$P(O(R_2,S_1)=0)=0.2$	$P(O(R_5,S_7)=0)=0.2$
	$P(U_{S_5}=1)=0.6$	$P(O(R_2,S_1)=0)=0.2$	
	$P(U_{S_5}=1)=0.3$	$P(O(R_2,S_5)=0)=0.3$	
	$P(U_{S_5}=1)=0.3$	$P(O(R_2,S_7)=0)=0.2$	
	$P(U_{S_{10}}=1)=0.4$	$P(O(R_3,S_7)=0)=0.3$	
	$P(U_{S_{11}}=1)=0.4$	$P(O(R_3,S_8)=0)=0.3$	

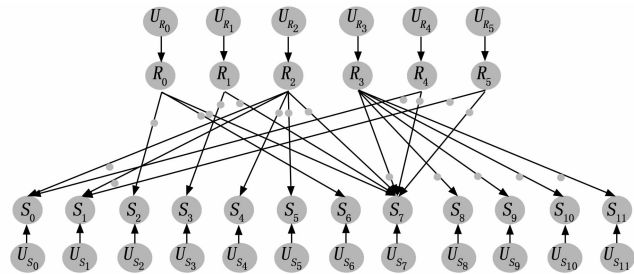


图 11 因果推理模型

表2 反事实推理结果

证据	充分因	必要因
$S_7=1, S_2=0,$	$E_{\text{suf}}(R_0)=0.041\ 02$	$E_{\text{dis}}(R_0)=0.035\ 77$
$S_3=0, S_4=0,$	$E_{\text{suf}}(R_1)=0.145\ 64$	$E_{\text{dis}}(R_1)=0.042\ 33$
$S_5=0, S_6=0,$	$E_{\text{suf}}(R_2)=0.000\ 50$	$E_{\text{dis}}(R_2)=0.000\ 02$
$S_8=0, S_9=0,$	$E_{\text{suf}}(R_3)=0.011\ 99$	$E_{\text{dis}}(R_3)=0.000\ 36$
$S_{10}=0, S_{11}=0$	$E_{\text{suf}}(R_4)=0.191\ 63$	$E_{\text{dis}}(R_4)=0.179\ 20$
	$E_{\text{suf}}(R_5)=0.191\ 63$	$E_{\text{dis}}(R_5)=0.104\ 43$

3.3 实验结果分析

表2推理结果显示,导致“无法喷射”情况发生的原因的可能性排序为 $R_4>R_5>R_1>R_0>R_3>R_2$,现场工程师可以依次对故障原因进行分析并制定维修策略。其中 R_4 为“出油阀无法正常启闭”的可能性最大。该推理结果与现场工程师提供的实际情况一致:如果出油阀无法正常启闭,那么高压燃油就无法流出高压油泵,而在高压油泵内压缩产生的高压燃油也就无法送入喷油器中,因此,喷油器内无燃油供给,从而导致无法喷射的现象出现。

图11(1)为基于故障因果网络的定性研究结果,

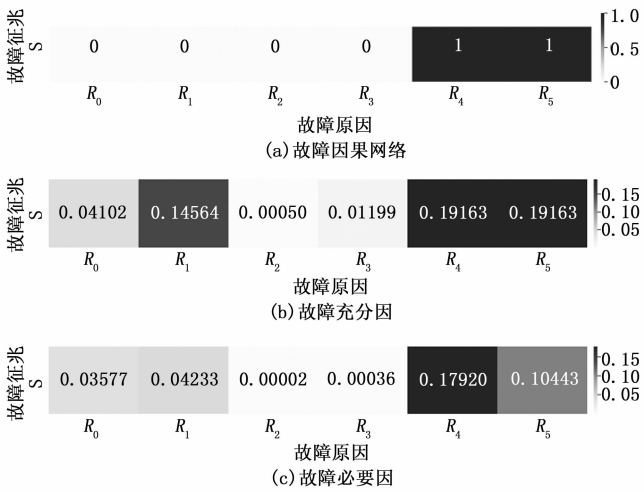


图 12 结果对比

图11(2)和(3)为反事实推理定量研究结果。前者能够快速判断出单一征兆是否由特定的故障原因引起,但在细致区分多重原因时存在不足;而后者弥补了这一缺陷,在前者生成的故障因果网络的基础上构建因果推理模型,对征兆的故障原因的可能性进行排序。

图中出现了 R_4 和 R_5 故障充分因数数值一样的情况,对推理过程加以分析。在反事实推理过程中构建孪生网络后,分别对 R_4 和 R_5 应用 do 算子进行干预并化简网络。化简后的两个网络结构脉络一致,化简后的节点虽有部分不同,但这些不同的节点所对应的概率权重数值完全一致。因此,根据故障充分因公式计算得到的结果也是相同的。在此情形下,故障必要因作为充分因的补充就显得格外重要。与充分因的充分性角度不同,必要因从必要性角度对因果强度进行重新计算,从而实现 R_4 和 R_5 的精确区分。这一案例不仅验证了本文方法的可用性与有效性,也进一步表明了故障充分因和必要因指标从两个不同的角度保障了推理的准确性,具有双重保障的作用。

4 结束语

针对柴油机故障原因与故障征兆之间的因果关系隐蔽性强、影响因素关联复杂、诊断难度较高的问题,提出了一种基于 Leaky Noisy-OR 假设和孪生网络的因果推理模型,实现了不确定性因素的量化、推理效率的提高、因果关系的度量,以及推理准确性上的双重保障。实验结果表明,通过对高压油泵和喷油器进行研究,证明了本文方法能够清晰地呈现共轨系统中故障原因与故障征兆之间的故障因果链,以及充分因和必要因对推理的准确性具有双重保障作用,提升了该方法在高压共轨系统实际应用中的可解释性和实用性,并为现场工程师提供了更加直观的故障因果链参考,有助于提高检修效率和降低检修成本。然而,受限于时间、知识储备以及资源条件,研究仍存在不足之处,尚需进一步深入探

索, 如 Leaky Noisy-OR 假设主要适用于静态因果推理, 未考虑到因果关系随时间变化的场景, 或者原因与结果之间的时间延迟效应, 无法建模动态系统中的因果关系的演化, 未来可以将时间作为变量引入模型中, 为每个原因变量赋予时间依赖的激活概率函数, 使模型能够适应不断变化的因果关系。

参考文献:

- [1] VERMA P, ZARE A, JAFARI M. Diesel engine performance and emissions with fuels derived from waste tyres [J]. *Scientific Reports*, 2018, 8 (1): 1–13.
- [2] LU N, JIANG B, MENG X. Diagnosis, diagnosticability analysis, and test point design for multiple faults based on multisignal modeling and blind source separation [J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2017, 50 (1): 137–148.
- [3] JIANG B, SHEN Q, SHI P. Neural-networked adaptive tracking control for switched nonlinear pure-feedback systems under arbitrary switching [J]. *Automatica*, 2015, 61 (1): 119–125.
- [4] BEJGER A, DRZEWIENIECKI J B. The use of acoustic emission to diagnosis of fuel injection pumps of marine diesel engines [J]. *Energies*, 2019, 12 (24): 4661–4672.
- [5] SONG E, KE Y, YAO C. Fault diagnosis method for high-pressure common rail injector based on IFOA-VMD and hierarchical dispersion entropy [J]. *Entropy*, 2019, 21 (10): 923–943.
- [6] GEORGE S, BALLA S, GAUTAM M. Effect of diesel soot contaminated oil on engine wear [J]. *Wear*, 2007, 262 (9/10): 1113–1122.
- [7] COLAO M J, TEIXEIRA C V, DUTRA L M. Thermal analysis of a diesel engine operating with diesel-biodiesel blends [J]. *Fuel*, 2010, 89 (12): 3742–3752.
- [8] HERNÁNDEZ M A. Methods of public health research—Strengthening causal inference from observational data [J]. *New England Journal of Medicine*, 2021, 385 (15): 1345–1348.
- [9] 王伟, 赵晋宇, 强沥文, 等. 农业环境损害因果关系鉴定研究 [J]. *农业环境科学学报*, 2022, 41 (9): 1855–1863.
- [10] ZHANG K, SUN Q, ZHAO C. Causal reasoning in typical computer vision tasks [J]. *Science China Technological Sciences*, 2024, 67 (1): 105–120.
- [11] 文向南, 撤奥洋, 于立涛, 等. 基于因果网络的配电网故障诊断模型研究 [J]. *青岛大学学报 (工程技术版)*, 2016, 31 (4): 68–73.
- [12] 马存宝, 马 婷, 周方旺. 基于因果网络模型的飞机通信系统故障诊断 [J]. *测控技术*, 2013, 32 (10): 52–54+58.
- [13] YAO L, CHU Z, LI S. A survey on causal inference [J]. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 2021, 15 (5): 1–46.
- [14] BONGERS S, FORRÉ P, PETERS J. Foundations of structural causal models with cycles and latent variables [J]. *The Annals of Statistics*, 2021, 49 (5): 2885–2915.
- [15] PETERS J, JANZING D, SCHLKOPF B. *Elements of Causal Inference: Foundations and Learning Algorithms* [M]. Cambridge: The MIT Press, 2017.
- [16] PEARL J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference* [M]. Amsterdam: Elsevier, 2014.
- [17] YAO L, CHU Z, LI S. A survey on causal inference [J]. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 2021, 15 (5): 1–46.
- [18] LAURITZEN S L. *Graphical Models* [M]. Oxford: Clarendon Press, 1996.
- [19] NOGUCHI T, FENTON N, NEIL M. Addressing the practical limitations of Noisy-OR using conditional inter-causal anti-correlation with ranked nodes [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2018, 31 (4): 813–817.
- [20] LIAO F, LIANG M, LI Z. Evaluate the malignancy of pulmonary nodules using the 3-D deep Leaky Noisy-OR network [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2019, 30 (11): 3484–3495.
- [21] ZHOU K, MARTIN A, Pan Q. The belief Noisy-OR model applied to network reliability analysis [J]. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2016, 24 (6): 937–960.
- [22] 王 仲. 燃气—蒸汽联合循环机组智能诊断与健康维护技术研究 [D]. 北京: 华北电力大学 (北京), 2021.
- [23] OH K, YOON J S, SUK H I. Learn-explain-reinforce: Counterfactual reasoning and its guidance to reinforce an Alzheimer's Disease diagnosis model [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 45 (4): 4843–4857.
- [24] PEARL J. *Causality* [M]. Cambridge: Cambridge University Press, 2009.
- [25] VLONTZOS A, KAINZ B, GILLIGAN-LEE C M. Estimating categorical counterfactuals via deep twin networks [J]. *Nature Machine Intelligence*, 2023, 5 (2): 159–168.
- [26] RICHENS J G, LEE C M, JOHRI S. Improving the accuracy of medical diagnosis with causal machine learning [J]. *Nature Communications*, 2020, 11 (1): 3923–3932.
- [27] PROSPERI M, GUO Y, SPERRIN M. Causal inference and counterfactual prediction in machine learning for actionable healthcare [J]. *Nature Machine Intelligence*, 2020, 2 (7): 369–375.