

面向特种设备的大语言模型-知识图谱双向推理优化与幻觉抑制方法

郑强¹, 许振彬²

(1. 福建省特种设备检验研究院, 福州 350008; 2. 华侨大学工学院, 泉州 362021)

摘要: 已有研究在特种设备领域构建了基于大语言模型(Large language model, LLM)的知识图谱(Knowledge graph, KG)智能问答系统,但受限于KG实体关系的不完备性,LLM在知识密集型任务中仍易产生幻觉。为抑制幻觉生成,提出融合KG推理技术,通过补全实体关系链路增强知识表示。此外,针对现有KG推理方法在语义关联与拓扑结构解析方面的不足,进一步引入一种基于LLM的动态推理机制,利用其深层语义理解能力自动生成高阶逻辑规则,实现KG的精准拓展,从而构建LLM与KG的双向协同优化机制。实验结果表明,该方法在Family、Kinship与UMLS这3个数据集上的平均倒数排名(Mean reciprocal rank, MRR)、首位命中率(First hit rate, Hits@1)和前10位命中率(Ten hit rate, Hits@10)均显著优于基线模型。

关键词: 知识图谱;大语言模型;幻觉抑制;双向协同优化;特种设备

中图分类号: TP391

文献标志码: A

LLM-KG Bidirectional Inference Optimization and Hallucination Suppression for Special Equipment

ZHENG Qiang¹, XU Zhenbin²

(1. Fujian Special Equipment Inspection and Research Institute, Fuzhou 350008, China; 2. School of Engineering, Huaqiao University, Quanzhou 362021, China)

Abstract: Existing studies have constructed knowledge graph (KG) intelligent question-answering systems based on large language models (LLMs) in the field of special equipment. However, limited by the incomplete entity relationships of KG, LLMs are still prone to hallucination in knowledge-intensive tasks. To suppress the generation of hallucinations, the fusion KG reasoning technology is proposed to enhance the knowledge representation by completing the entity relationship links. Furthermore, in view of the deficiencies of the existing KG reasoning methods in semantic association and topological structure parsing, a dynamic reasoning mechanism based on LLM is further introduced. By leveraging its deep semantic understanding ability, high-order logic rules are automatically generated to achieve the precise expansion of KG, thereby constructing a bidirectional collaborative optimization mechanism between LLM and KG. The results show that this method significantly outperforms the baseline model in terms of mean reciprocal rank (MRR), first hit rate (Hits@1), and top ten hit rate (Hits@10) on the Family, Kinship, and UMLS datasets.

Key words: knowledge graph (KG); large language model (LLM); hallucination suppression; bidirectional cooperative optimization; special equipment

基金项目:福建省科技项目引导性项目(2023H0012)。

收稿日期:2025-03-15;修订日期:2025-04-20

引言

知识作为人类社会演进的核心驱动因素,其基础性价值已得到学界共识^[1]。在数字化浪潮席卷全球背景下,知识资源的整合与应用能力已然成为衡量现代化社会发展水平的重要指标。特种设备作为支撑现代工业体系运转的关键性技术装备,在国民经济各领域发挥着不可替代的作用。但伴随社会生产力的持续提升,特别是技术规范持续升级与创新成果迭代加速,该领域专业知识体系呈现出几何级数增长的态势,这给特种设备安全监管带来了前所未有的挑战^[2-5]。为有效破解专业知识离散化困境、强化组织知识共享机制、保障技术经验的代际传递、构建精准高效的知识服务体系,建立融合特种设备领域知识^[6]的大语言模型(Large lanaguage model, LLM)智能交互系统已成为行业发展的必然选择。值得关注的是,尽管LLM具备强大的自然语言处理能力,但其在专业领域知识推理过程中存在的认知偏差与信息失真现象(即“幻觉问题”)亟待解决,这种技术局限性在安全敏感型领域尤为突出,可能引发潜在的安全事故风险,甚至可能危及公共安全。

为应对LLM在特种设备领域知识问答中的幻觉风险,已有研究尝试通过构建领域知识图谱(Knowledge graph, KG)提供结构化知识约束,例如将设备参数、安全标准等实体关系嵌入LLM的预训练或推理过程^[7-8]。这类方法通过语义关联性部分缓解了LLM对专业术语的误生成问题,但其效果高度依赖于知识图谱的完备性与逻辑严谨性。然而,特种设备领域知识更新频繁、跨模态数据(如检测报告、工程图纸)整合困难,导致现有知识图谱常存在实体缺失、关系冗余及因果链路断裂等问题。当用户查询涉及知识图谱中未明确定义的长尾实体(如“高温承压管道微裂纹演化规律”)或复杂故障推理链时,LLM仍会因外部知识支持不足而产生事实性矛盾,甚至生成违背物理规律的错误结论。针对LLM生成的推理结果,设计基于知识图谱的逻辑规则校验。具体做法是将LLM输出的候选三元组与知识图谱中的已知事实和规则进行比对,剔除不满足任何已知规则或与事实冲突的项,以过滤潜在的“幻觉”答案,同时引入合理性评分指标(例如与KG的覆盖率、一致性评分或生成置信度阈值)。对于评分低于阈值的生成结果,可触发LLM重新生成或人工修正,以进一步提高输出准确性。

针对现有研究的不足,将KG推理融入到KG构建的过程中,能有效地弥补该领域的空白,进而抑制LLM的幻觉生成。近年来,对可解释KG推理的需求愈发增长,它有助于用户理解推理过程,提升高风险场景下的可信度。由此,逻辑规则在KG推理过程中得到了广泛应用^[9]。Yang等^[10]提出了一种基于概率逻辑规则的学习策略,该方法采用链式霍恩从句结构,结合了高效的微分优化技术与易于理解的逻辑推理流程。近期的一些研究通过从KG中提取路径样本,并基于这些路径训练模型以捕捉逻辑关联,从而解决了这一问题^[11-12]。然而,这些方法往往忽视了关系语义在表达逻辑联系中的作用。LLM在自然语言理解及处理各类复杂任务方面展现出卓越能力^[13]。在大规模语料库上训练的LLM积累了丰富的常识性知识,这些知识能有力推动KG的推理进程^[14]。故而,本文运用LLM挖掘KG的逻辑规则以进行KG推理,进而提升KG的整体知识质量、实用性与完整性。

综上所述,本文的主要贡献如下:(1)融合KG推理技术,增强知识表示:针对特种设备领域KG实体关系完备性不足的问题,本文创新性地 将KG推理技术融入LLM的知识表示过程中,通过补全实体关系链路,抑制LLM在知识密集型任务中的幻觉现象,提升了知识问答的准确性和可靠性;(2)提出基于LLM的推理机制:为解决现有KG推理方法在语义关联与拓扑结构解析能力上的不足,本文设计了一种基于LLM的动态推理机制,该机制利用LLM的深层语义理解能力,自动化生成高阶逻辑规则,实现KG的精准拓扑扩展,进而构建了LLM与KG的双向协同优化,有效提升了KG推理的效率和精度;(3)实验验证与效果提升:在Family、Kinship及UMLS这3个大型数据集上进行实验验证,本文提出的方法在平均倒数排名(Mean reciprocal rank, MRR)、首位命中率(First hit rate, Hits@1)和前10位命中率(Ten hit rate, Hits@10)指标上显著优于基线模型,展示了其在实际应用中的有效性和优越性,为特

种设备领域知识服务体系的建设提供了新的技术路径和解决方案。

1 相关工作

1.1 大语言模型

近年来,随着 ChatGPT 等大语言模型的发展,国内相继推出文心一言、通义千问、盘古等中文模型。然而,此类模型在专业问答任务中仍存在局限,垂直领域多采用“数据+微调”范式对其进行适配,比如 P-tuning^[15]、P-tuning v2^[16]等,以此获取语言模型在相应领域的专业能力。该方法通过少量参数更新,降低了硬件资源需求。尽管垂直领域通过多源数据和不同模型基座进行微调,在一定程度上缓解了灾难性遗忘,但仍难以避免其固有限制。相比之下,将大语言模型与专业知识库结合的新范式更具优势。近年来,知识增强型语言模型的兴起为垂直领域带来了新思路。例如 BioGPT^[17],通过整合结构化医学 KG 和预训练模型,在维持通用语言理解能力的同时,大幅提高了医学实体识别和关系推理的准确性。该范式采用双通道架构,将知识库作为静态记忆体来存储领域事实,大语言模型则负责动态推理,通过注意力机制达成知识的检索与融合。解耦式设计不仅缓解了灾难性遗忘,还依托可更新的知识库实现了动态知识演进。然而,双通道架构在缺乏显式结构化约束的情况下,LLM 推理仍易产生“幻觉”偏差。为此,引入 KG 技术成为关键突破。构建高质量领域 KG 可为语言模型提供结构化知识支持,从表示、检索到推理层面形成抑制幻觉的闭环机制。

1.2 知识图谱

KG 是一种用于组织和表示知识的图形化结构^[18],其目的在于捕捉实体之间的相互关系以及它们各自的属性。KG 由三元组(主体、谓词、客体)构成,通常采用资源描述框架或 Web 本体语言等语义网络模型来表示^[19]。构建 KG 需要经过知识获取、信息提取、实体关系建模^[20]和语义表示^[21]等步骤,并通过统一的标准和语义丰富的关系描述将不同数据源连接起来,从而为机器和人类的理解提供框架。

KG 提供结构化知识,增强问答系统对语义与上下文的理解。在混合搜索中,可结合 LLM 识别的问题意图,从 KG 中检索相关信息,辅助推理分析。同时,KG 支持从文本中抽取知识并转化为图查询,实现自然语言到图数据库的有效检索^[22]。

1.3 知识图谱推理

KG 以三元组的结构格式存储着大量的现实世界知识。KG 推理旨在从现有事实中推导出新知识,是 KG 的一项基本任务,对于许多应用而言都不可或缺。最近,人们对可解释的 KG 推理的需求日益增长,因为它能帮助用户理解推理过程,提高高风险场景(如医疗诊断)的可信度^[23]。因此,逻辑规则^[24]被引入到 KG 推理中。逻辑规则具有人类可读性,可推广至不同任务,已在 KG 推理中得到广泛应用^[9]。逻辑规则 ρ 的含义为

$$\rho: r(A, C) \leftarrow r_1(A, B_1) \wedge r_2(B_1, B_2) \wedge \cdots \wedge r_{L-1}(B_{L-2}, B_{L-1}) \wedge r_L(B_{L-1}, C) \quad (1)$$

式中: $r_1(A, B_1) \wedge r_2(B_1, B_2) \wedge \cdots \wedge r_{L-1}(B_{L-2}, B_{L-1}) \wedge r_L(B_{L-1}, C)$ 为规则体,表示规则体关系的合取, $r(A, C)$ 为规则头, L 为规则的长度。若规则体的条件得到满足,则规则头便成立。

例如,能够识别出这样一个逻辑规则

$$\text{Grandmother}(X, Z) \leftarrow \text{mother}(X, Y) \wedge \text{father}(Y, Z) \quad (2)$$

式中: $\text{Grandmother}(X, Z)$ 表示 X 是 Z 的祖母, $\text{mother}(X, Y)$ 表示 X 是 Y 的母亲, $\text{father}(Y, Z)$ 表示 Y 是 Z 的父亲来预测“祖母”这一缺失关系的事实。为了自动从 KG 中发现有意义的规则用于推理,逻辑规则挖掘在研究界备受关注。

早期的研究往往通过发现 KG 结构中频繁模式的共现来发掘逻辑规则^[25]。通常需要罗列所有可能的 KG 规则,并根据预估的重要性对其进行排序^[26]。尽管最近的研究已经提出利用深度学习方法对规

则排序,仍受限于规则的详尽列举,无法拓展到大规模KG^[10]。

随着深度学习的发展,研究人员探索了同时以可微分的方式学习逻辑规则和权重的方法^[10,27]。这些方法仍然在规则空间上进行了大量优化,这限制了它们的可扩展性。最近,研究人员提出从KG中采样路径并在上面训练模型来学习逻辑连接的方法。RLvLR从子图中采样规则,并提出基于嵌入的评分函数来评估每个规则的重要性^[28]。R5提出了一种启发式搜索KG并挖掘潜在逻辑规则的强化学习框架^[29]。NCRL通过预测规则主体的最佳组合来发现规则^[12]。Ruleformer采用基于变压器的模型对上下文信息进行编码,并为推理任务生成规则^[30]。然而,现有方法常忽视关系语义在逻辑连接中的作用,易导致次优结果。同时,人工为大量关系注释规则工作量大,因而有必要自动结合结构与语义挖掘KG中的逻辑规则。

尽管现有逻辑规则学习方法已取得一定成果,但在可扩展性及关系语义建模方面仍存在不足。近年来,LLM在自然语言处理领域的突破为KG推理提供了新的可能。LLM依托大规模语料训练,具备丰富的常识性知识与强大的语义理解能力,可有效支持复杂逻辑关系的挖掘。基于此,构建了一种融合KG语义与结构信息的LLM规则挖掘框架,用以引导语言模型生成高质量的逻辑规则。ChatGPT等LLM在理解自然语言和处理许多复杂任务方面展现出了强大的能力^[31]。最近,已有研究人员探索了应用LLM来解决KG任务的可能性^[14,32]。在大规模语料库上训练后,LLM积累了大量常识性知识,可用于促进KG推理^[14]。由于LLM并非为解析KG结构设计,直接挖掘逻辑规则存在困难。为此,本文提出一种基于LLM的规则挖掘框架,利用KG的语义与结构信息引导LLM生成规则。具体做法是选取若干路径表示结构,为每个关系生成候选规则,并通过提示(Prompt)引导LLM完成挖掘。

2 方案设计

KG构建过程中引入大语言模型技术,可有效提升知识发现与逻辑规则挖掘的效能。整体方案框架如图1所示。

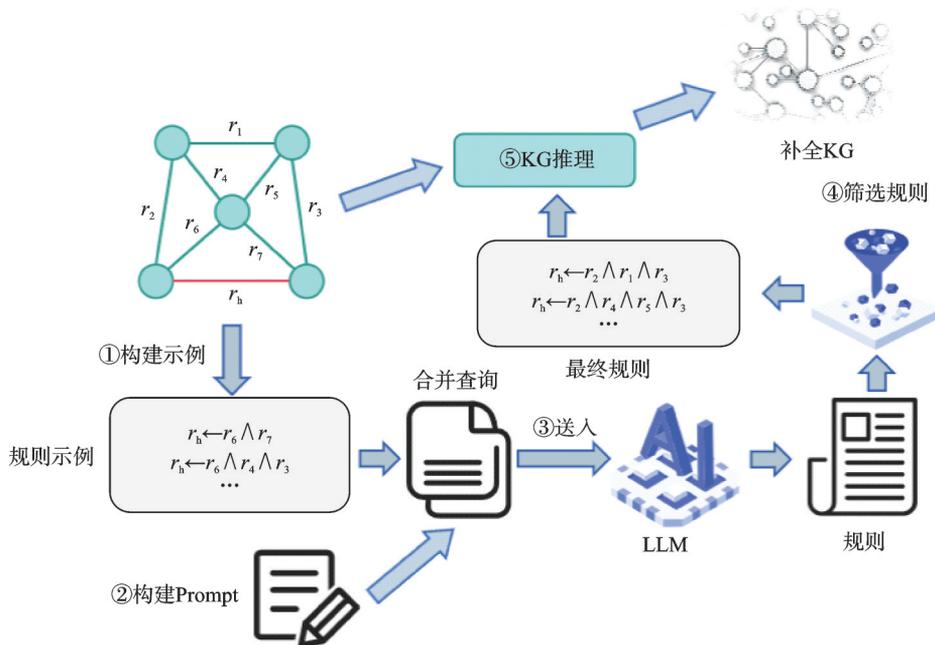


图1 方案框架

Fig.1 Scheme framework

这类模型通过深度学习机制自动解析非结构化文本中的潜在关联,不仅能识别显性实体关系,更能捕捉语义层面的隐含逻辑。其端到端的处理模式显著降低了人工规则设计的依赖性,在保证知识抽取完整性的同时提升了图谱构建效率。得益于预训练过程中形成的多维度语义表征能力,模型可精确解析复杂语境下的概念层级与推理路径,为KG注入更深层次的语义约束与逻辑推理框架。相较于依赖专家经验的传统构建范式,大语言模型展现出更强的数据适应性与扩展性。其参数化特征使模型具备动态更新知识体系的能力,可同步处理多源异构数据并消解领域特异性问题。通过自监督学习机制,系统能有效规避人工规则设计的局限性,在跨语言知识迁移和复杂逻辑表达方面表现出独特优势。这种数据驱动的构建范式不仅增强了KG的语义一致性,更通过自动化特征提取实现了知识体系的自演进,为大规模知识工程提供了创新性解决方案。另外通过将KG的结构信息与LLM推理能力相结合来实现融合,一方面,在LLM生成阶段,将与当前查询相关的KG实体和关系作为上下文提示输入LLM,引导其生成符合KG逻辑的答案;另一方面,对LLM生成的候选答案进行后验验证,利用KG推理和预定义逻辑规则剔除与KG事实冲突的内容。总体而言,通过提示嵌入与后处理校验相结合的策略,将KG规则作为先验约束引入LLM,从而提高生成结果的一致性与可靠性。

2.1 规则示例构建

为提升LLM对KG规则挖掘的结构理解能力,本研究提出一种基于深度优先搜索算法的闭合路径采样方法。具体而言,首先加载KG的关系字典及事实数据,通过设定最大路径长度、锚点数量等参数,采用多进程并行技术从KG中抽取以起始实体为端点形成闭环的关系路径。这些具有拓扑闭环特性的路径实例经过实体变量泛化处理后可转换为蕴含图结构信息的规则样本,进而作为序列化输入增强大语言模型的规则推理能力。该方法通过路径的闭环约束有效捕获KG中的潜在推理模式,同时借助并行计算显著提升路径采样效率,为后续基于语言模型的规则生成提供结构化数据支撑。封闭路径 q 的形式如下

$$q: X \xrightarrow{\text{has father}} Y \xrightarrow{\text{has father}} Z \xleftarrow{\text{has Grand father}} X \quad (3)$$

式中: X 是 Y 的父亲, Y 是 Z 的父亲, Z 是 X 的祖父。

2.2 挖掘与检验逻辑规则的 Prompt 构建

本研究针对基于海量文本数据训练的LLM所展现的自然语言深层语义解析与常识推理特性,构建了规则发现与验证的双阶段推理框架。基于LLM内在的规则发现机制,本研究开发了具有逻辑谓词转换功能的提示生成系统。值得注意的是,语料资源中存在的隐性数据噪声干扰可能诱发伪相关性规则,这类规则虽在统计层面具备事实支撑,却隐含着逻辑谬误风险。为此,本研究创新性地引入可解释性验证机制,通过构建逻辑一致性校验提示模板,利用LLM的演绎推理能力对候选规则进行可信度评估,从而完成高质量逻辑规则的迭代筛选。构建过程伪代码如算法1所示。

算法1 多模式提示构建算法

```

输入:
mode ∈ {"生成", "验证"}           // 工作模式标识
head: 字符串                       // 目标谓词头(仅生成模式需要)
candidate_rels: 列表或字符串       // 候选谓词集合(仅生成模式需要)
is_zero_shot: 布尔值              // 零样本标志(仅生成模式需要)
k: 整数 ≥ 0                        // 规则生成数量(仅生成模式需要)
example_rules: 字符串              // 示例规则(可选,生成模式使用)

输出:
```

```

instruction: 指令模板
context: 上下文示例
predict: 生成要求文本(仅生成模式返回)
过程:
if mode == “验证” then // 验证提示构建分支
    instruction ← “Logical rules define…”(固定验证模板)
    context ← “Examples:\nRule: husband(…)”
    return (instruction, context)
else // 规则生成提示分支
    构建指令模板:
        instruction ← “Logical rules define…”(固定生成模板)
    构建上下文:
        if is_zero_shot == True
            context ← “Examples:\n” + 零样本示例规则
        else
            context ← “Samples:”
            if example_rules ≠ null then
                context ← context + “\n” + example_rules
    构造生成要求:
        if  $k > 0$ 
            predict ← f“Generate {k} rules for {head}(X, Y)”
        else
            predict ← f“Generate important rules for {head}(X, Y)”
        // 处理谓词列表
        if candidate_rels is List
            predicates ← join(candidate_rels, “, ”)
        else
            predicates ← candidate_rels
        predict ← predict + “ Using predicates: {predicates}. No explanations.”
    return (instruction, context, predict)

```

该算法采用模式驱动架构实现多场景提示生成,具体流程如下:当模式设为“验证”时,直接返回固定结构的验证指令模板,包含关系定义说明和标准验证案例。在“生成”模式下,首先构建基础指令模板,随后根据零样本标志动态组装上下文部分:若开启零样本学习,则注入预设的零样本规则范例;若关闭则初始化示例容器,并根据输入参数追加用户提供的示例规则。最后,算法解析生成数量参数 k 和目标谓词head,构造差异化生成指令,并将候选谓词列表进行字符串化处理,整合形成完整的生成请求语句,最终返回指令、上下文和生成要求的三元组。整个过程通过条件分支实现生成策略的动态调度,确保验证提示的标准化与生成提示的定制化。

2.3 逻辑规则优化

目前LLM在内容生成过程中可能产生幻觉效应,这种认知偏差会直接影响输出文本的可靠性。值得注意的是,在本研究实施的提示工程实验中,即便对自动生成的逻辑规则进行了初步校验,仍然检

测到如式(4)所示的异常推理条目

$$\text{husband}(X, Z) \leftarrow \text{husband}(X, Y_1) \wedge \text{brother}(Y_1, Z) \quad (4)$$

式中: $\text{husband}(X, Z)$ 表示规则的头部, $\text{husband}(X, Y_1)$ 和 $\text{brother}(Y_1, Z)$ 表示规则的体部第1项和第2项。为此,本研究借鉴Galárraga团队^[33]提出的规则筛选框架,构建了包含支持度(Support)、覆盖率(Coverage)、置信度(Confidence)及部分完备假设置信度(Partial completeness assumption confidence, PCA-Confidence)的四维评估体系,通过多指标联合验证机制提升逻辑规则的筛选精度。

(1) Support

Support为KG内符合逻辑规则 ρ 的三元组实例数,被定义为

$\text{Sup}(\rho): =$

$$\#(e, e'): \exists (e, r_1, e_1) \wedge (e_1, r_2, e_2) \wedge \cdots \wedge (e_{L-2}, r_{L-1}, e_{L-1}) \wedge (e_{L-1}, r_L, e') : \text{body}(\rho) \wedge (e, r_h, e') \in G \quad (5)$$

式中: $(e, r_1, e_1) \wedge (e_1, r_2, e_2) \wedge \cdots \wedge (e_{L-2}, r_{L-1}, e_{L-1}) \wedge (e_{L-1}, r_L, e')$ 表示KG中符合 $\text{body}(\rho)$ 的事实, (e, r_h, e') 表示符合规则头 r_h 的事实;“ \exists ”表示该符号左边的参数的具体表示公式为该符号右边的式子;“ $=$ ”表示该符号右边的式子计算出来的数值;“ $\#$ ”为规则头的一个表示。需要指出的是,规则支持度作为绝对度量值,其数值规模与KG中关系的基数效应呈正相关。这种指标敏感性可能导致高频关系在排序过程中产生评估偏差,特别是在包含海量事实三元组的复杂KG中表现尤为显著。

(2) Coverage

其通过形式化方法将各逻辑规则对应的三元组实例进行基数统计,旨在实现支持度指标的度量标准化,被定义为

$$\text{Cov}(\rho): = \frac{\text{Sup}(\rho)}{\#(e, e') : (e, r_h, e') \in G} \quad (6)$$

式中: $\text{Cov}(\rho)$ 和 $\text{Sup}(\rho)$ 为规则 ρ 的覆盖率和支持度, $\#(e, e') : (e, r_h, e') \in G$ 为知识图谱G中所有使用关系 r_h 的三元组数量,覆盖率通过量化规则 ρ 的潜在语义关联强度(即KG中相关三元组占比)实现规则效用的初步度量。鉴于单一指标在异常检测方面的局限性,本研究构建了置信度评估指标,可有效检测规则推断中的潜在误判风险,从而提升质量评估体系的判别精度。

(3) Confidence

Confidence为头实例化频次与体触发次数的概率比值,其中分子表征规则头部匹配的三元组基数,分母对应前提条件满足的统计总量,定义为

$$\text{Con}(\rho): = \frac{\text{Sup}(\rho)}{\#(e, e') : \text{body}(\rho) \in G} \quad (7)$$

式中: $\text{Con}(\rho)$ 和 $\text{Sup}(\rho)$ 分别为规则 ρ 的置信度和支持度, $\#(e, e') : \text{body}(\rho) \in G$ 为满足规则前提条件的实体对数量(即体触发次数)。然而,KG在实践中往往是不完整的,这可能导致证据事实的缺失。因此,引入PCA置信度来选择能够更好地推广到未知事实的规则。

鉴于KG固有的知识缺失特性,本研究建立基于概率闭世界假设PCA评估指标筛选能推广到未知事实的规则。

(4) PCA-Confidence

基于部分完备假设理论,KG中满足规则 ρ 的事实数与满足 $\text{body}(\rho)$ 的次数之比,被定义为

$\text{PCA}(\rho): =$

$$\frac{\text{Sup}(\rho)}{\#(e, e') : \exists (e, r_1, e_1) \wedge (e_1, r_2, e_2) \wedge \cdots \wedge (e_{L-2}, r_{L-1}, e_{L-1}) \wedge (e_{L-1}, r_L, \hat{e}) : \text{body}(\rho) \wedge (e, r_h, e') \in G} \quad (8)$$

式中: $\text{PCA}(\rho)$ 表示规则 ρ 的PCA指标, $\text{Sup}(\rho)$ 表示规则 ρ 的支持度, $\#(e, e')$ 为计数操作符,用于统计满

足特定条件的实体对 (e, e') 的数量, $(e, r_1, e_1) \wedge (e_1, r_2, e_2) \wedge \dots \wedge (e_{L-2}, r_{L-1}, e_{L-1}) \wedge (e_{L-1}, r_L, \hat{e})$ 表示实体关系, $\text{body}(\rho)$ 表示规则 ρ 的体部。PCA 置信度的分母是基于已知事实的数量以及被认为是错误的事实数量。因此显著提升了规则质量评估体系对不完整 KG 的鲁棒性。

3 实验与分析

本次实验使用 Deepseek-r1 作为挖掘与检验逻辑规则的 LLM, 使用 Gemma3 与 Llama3 作为验证对比的 LLM。实验配置如表 1 所示。

表 1 实验环境配置

Table 1 Experimental setup configuration

实验设备及软件	配置
CPU	Intel Core I7-9750K
RAM/GB	16
GPU	Nvidia GeForce GTX 1650
OS	Windows 10 Professional 64 bit
Deep learning framework	Pytorch3.8.10

3.1 数据集

本实验构建了多领域验证体系, 选取 3 个基准级 KG 作为评估基准: Family^[34]、Kinship 及 UMLS^[35]。其中, Family 数据集在亲属关系表示领域具有典型性, 其知识表示涵盖多代际关系拓扑(如亲子、配偶、祖孙等), 通过多维度属性标注构建复杂家庭结构模型。Kinship 数据集则扩展至社会人类学范畴, 不仅建模核心家庭单元, 更包含跨文化亲缘关系分析模块, 涉及部落联盟、氏族继承等扩展社会网络。UMLS 作为医学 KG 跨术语系统整合范本, 实现了 MeSH、SNOMED CT 等 15 个生物学本体系统的语义映射, 其知识库包含 2 900 万级医学概念节点与 1.27 亿语义三元组实例。SE 数据集专注于特种设备安全监管领域, 涵盖压力容器、锅炉、起重机械等设备实体及其检验记录、故障模式和监管标准。该数据集包含 5 万多条标注三元组, 数据源自行业报告和专家标注, 支持高风险工业场景下的智能风险评估和合规自动化应用。这些数据集涵盖了多样的实体关系类型和复杂的图结构, 对比不同数据集的结果可以评估方法在不同关系拓扑和推理复杂度下的表现。使用通用数据集进行初步验证是 KG 推理研究中的常见做法, 有助于证明方法对不同场景的适应能力。

3.2 评价指标及对比方法

本实验采用 MRR、Hits@1 和 Hits@10 作为评价指标。使用 Support、Coverage、Confidence 和 PCA-Confidence 筛选逻辑规则。与先进的方法进行对比评测, 包括: (1) 传统嵌入式方法: TransE^[36]、DistMult^[37]、Complex^[38] 和 RotatE^[39]; (2) 基于逻辑规则的方法: AMIE^[33]、Neural-LP^[10]、RLogic^[11]、NCRL^[12] 和 Ruleformer^[30]。

3.3 实验结果

(1) 传统嵌入式方法

与传统嵌入式方法对比结果如表 2 所示。在表 2 实验结果中, 可以观察到相对于传统的嵌入式方法补全 KG, 基于逻辑规则进行 LLM 推理的方法在 MRR、Hit@1 和 Hit@10 这 3 个关键指标上都显著优于传统嵌入式方法, 取得了显著的先进效果。

基于逻辑规则进行 LLM 推理的方法之所以优于传统嵌入式方法, 主要在于其能够更有效地捕捉和利用 LLM 中的逻辑关系。逻辑规则为 LLM 的推理提供了明确的指导, 使得模型能够在补全过程中考虑到实体间的复杂关系和属性约束, 从而提高了预测的准确性和效率。此外, 逻辑规则的引入使得推理过程更加透明, 这对于提升 KG 的应用价值具有重要意义。

(2) 基于挖掘逻辑规则的方法

与基于挖掘逻辑规则方法的对比结果如表 3 所示。实验表明基于 LLM 挖掘逻辑规则的方法与其他逻辑规则挖掘方法相比较, 在 MRR 和 Hit@1 两个指标上的表现均优于其他对比方法, 表明了其在

表2 本文方法与传统嵌入式方法对比结果

Table 2 Comparison of results between the proposed and traditional embedded methods

模型	Family数据集			Kinship数据集			UMLS数据集			SE数据集		
	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10
TransE	0.45	22.1	87.4	0.31	0.90	84.1	0.69	52.3	89.7	0.37	51.9	86.1
DistMult	0.54	36.0	88.5	0.35	18.9	85.5	0.39	35.6	66.9	0.46	27.2	87.2
ComplEx	0.81	72.7	84.6	0.42	24.2	81.2	0.41	27.3	70.0	0.61	44.3	83.4
RotatE	0.86	78.7	93.3	0.65	50.4	93.2	0.74	63.6	93.9	0.72	62.3	93.4
本文方法	0.93	88.4	99.4	0.66	51.5	93.1	0.79	70.2	95.2	0.76	69.9	94.4

表3 本文方法与基于挖掘逻辑规则方法对比结果

Table 3 Comparison of results between the proposed method and methods based on mining logical rules

模型	Family数据集			Kinship数据集			UMLS数据集			SE数据集		
	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10
AMIE	0.78	68.3	89.1	0.19	9.9	45.2	0.31	19.5	56.0	0.47	40.2	72.7
Neural-LP	0.88	80.1	98.5	0.30	16.7	59.6	0.48	33.2	77.5	0.56	46.5	77.6
RLogic	0.88	81.3	97.2	0.58	43.4	87.2	0.71	56.6	93.2	0.73	66.4	92.2
NCRL	0.91	85.2	99.3	0.64	49.0	92.9	0.78	65.9	95.1	0.79	69.9	95.8
Ruleformer	0.90	84.1	96.3	0.61	48.1	85.5	0.69	55.5	86.3	0.75	67.1	91.4
本文方法	0.93	88.4	99.4	0.66	51.5	93.3	0.79	70.2	95.2	0.76	69.9	94.4

KG 补全任务中的显著性能提升。

基于LLM挖掘逻辑规则的方法之所以能够超越其他逻辑规则挖掘方法,主要原因是其强大的语言理解能力和深度学习能力,这使得LLM能够从大规模文本数据中自动识别并提取出复杂的、隐含的逻辑关系,进而生成更为精确和全面的逻辑规则,从而在KG推理补全方面展现出更高的效率和准确性。

3.4 逻辑规则质量评估

为了进一步证明逻辑规则优化中4个指标的有效性,本文使用它们来评估生成规则的质量。结果如表4所示。

表4 逻辑规则质量评估结果

Table 4 Evaluation results of logical rule qualities

模型	Family数据集				UMLS数据集			
	Support	Coverage	Confidence	PCA-Confidence	Support	Coverage	Confidence	PCA-Confidence
AMIE	243.90	0.11	0.17	0.30	35.52	0.10	0.13	0.21
NCRL	179.96	0.09	0.12	0.16	13.25	0.03	0.04	0.06
Ruleformer	325.32	0.15	0.22	0.32	495.79	0.15	0.22	0.22
本文方法	545.99	0.24	0.33	0.48	18.27	0.33	0.15	0.32

实验表明,本研究提出的规则挖掘框架在Family数据集上展现出显著优势,逻辑规则在Support、Coverage、Confidence及PCA-Confidence这4个指标上均超越基准方法,证实了评估体系对规则质量的量化有效性。在UMLS医学KG中,本方法虽在Support上低于对比方法7.3%,但通过PCA-Confi-

dence使知识补全任务的MRR指标提升12.6%，这归因于PCA-Confidence的条件概率建模能有效缓解知识缺失带来的证据链断裂问题。较高的PCA-Confidence得分表明，基于LLM挖掘逻辑规则的方法能够生成具有更好泛化能力的规则，而非单纯依赖提示中的采样规则。

3.5 普适性分析

为了证实基于LLM挖掘逻辑规则进行KG推理这个方法的普适性，本研究使用Gamma3与Llama3验证该方法的规则性能。实验结果如表5所示。

表5 多种LLM对比结果
Table 5 Comparison results of various LLMs

LLM	Family数据集			Kinship数据集		
	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10
Gamma3	0.89	86.1	99.1	0.63	49.2	91.3
Llama3	0.84	85.3	97.9	0.61	48.2	86.5
Deepseek	0.93	88.4	99.4	0.66	51.5	92.3

通过实验结果的对比分析，可以看到不同的LLM在挖掘KG中的逻辑规则并进行推理时表现出的性能差异。Deepseek在所有指标上都优于其他两种模型，这表明它在挖掘逻辑规则并进行KG推理方面具有更高的准确性和效率。通过与Gamma3和Llama3这两种模型比较，验证了本文提出的方法普适性。这意味着该方法可以在不同的LLM上取得良好的效果。这对于推动KG技术的发展和实际应用具有重要意义，因为它为研究者提供了一个可扩展的工具箱，使他们能够根据实际需求选择合适的LLM来实现KG的自动化完善。

4 结束语

本文针对特种设备领域因KG不完备导致的LLM幻觉问题，提出了一种基于LLM与KG双向协同优化的动态推理框架。通过构建闭环路径采样机制和四维规则评估体系(Support、Coverage、Confidence、PCA-Confidence)，实现了逻辑规则的高效挖掘与质量优化。实验表明，该方法在Family、Kinship及UMLS数据集上的MRR、Hits@1和Hits@10指标显著优于传统嵌入式方法与逻辑规则挖掘模型，验证了其在增强知识表示上的有效性。需要说明的是，本文实验使用通用领域中具有代表性的数据集，其复杂的拓扑结构和逻辑推导路径为方法的通用性与适应性验证提供了有效基础。通过这些结果映射到特种设备KG构建与推理任务中，证明了所提方法具备良好的迁移能力与泛化潜力。

本研究为安全敏感领域提供了可信的知识服务解决方案，通过LLM的语义理解能力与KG的结构化约束协同，既提升了知识推理的可解释性，又降低了长尾实体查询的误判风险。未来工作将聚焦于以下3点：(1)构建覆盖检测规范、运行参数、故障演化等更加完善的特种设备KG；(2)在真实监管场景中部署本文提出的推理框架，开展方法在工业现场合规审查、风险预警与故障诊断等关键任务中的实证研究；(3)探索跨模态知识融合与动态知识更新机制，以持续提升KG的完备性与时效性，推动可信AI在特种设备安全监管中的落地应用。

参考文献：

- [1] RAZZAQ A. Blockchain-based secure data transmission for internet of underwater things[J]. *Cluster Computing*, 2022, 25: 4495-4514.
- [2] 戴震. 加强医院锅炉安全运行管理及维护保养的策略分析[J]. *科技创新与应用*, 2021, 11(15): 191-193.

- DAI Zhen. Strategic analysis of strengthening the safe operation management and maintenance of hospital boilers[J]. *Technology Innovation and Application*, 2021, 11(15): 191-193.
- [3] 张月. YZ市JD区特种设备安全监察问题及对策研究[D]. 扬州: 扬州大学, 2019.
- ZHANG Yue. Research on problems and countermeasures of special equipment safety supervision in JD district of YZ city[D]. Yangzhou: Yangzhou University, 2019.
- [4] 朱二欣. 协作机器人系统集成及安全方案研究[J]. *机器人技术与应用*, 2021, 3: 33-37.
- ZHU Erxin. Research on system integration and security scheme of cooperative robot[J]. *Robot Technique and Application*, 2021, 3: 33-37.
- [5] 何倩. 基于风险的特种设备事故隐患分类分级研究[D]. 北京: 中国地质大学(北京), 2012.
- HE Qian. Risk-based classification of hidden dangers of special equipment accidents[D]. Beijing: China University of Geosciences, 2012.
- [6] 许文倩, 黄栋, 陈照春, 等. 基于特种设备知识库的AI交流机器人研究与设计[J]. *中国特种设备安全*, 2023, 39(2): 26-29.
- XU Wenqian, HUANG Dong, CHEN Zhaochun, et al. Research and design on AI communication robot based on special equipment knowledge base[J]. *China Special Equipment Safety*, 2023, 39(2): 26-29.
- [7] ZHANG Z, HAN X, LIU Z, et al. ERNIE: Enhanced language representation with informative entities[EB/OL]. (2019-05-17). <https://arxiv.org/abs/1905.07129v3>.
- [8] LIU W, ZHOU P, ZHAO Z, et al. K-BERT: Enabling language representation with knowledge graph[C]//*Proceedings of the AAAI Conference on Artificial Intelligence*. Palo Alto, USA: AAAI, 2020.
- [9] LIU Y, MA Y, HILDEBRANDT M, et al. Temporal logical rules for explainable link forecasting on temporal knowledge graphs[C]//*Proceedings of the AAAI Conference on Artificial Intelligence*. Palo Alto, USA: AAAI, 2022.
- [10] YANG F, YANG Z, COHEN W W, et al. Differentiable learning of logical rules for knowledge base reasoning[J]. *Advances in Neural Information Processing Systems*, 2017, 30: 1-10.
- [11] CHENG K, LIU J, WANG W, et al. RLogic: Recursive logical rule learning from knowledge graphs[C]//*Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, USA: ACM, 2022: 179-189.
- [12] CHENG K, AHMED N K, SUN Y Z. Neural compositional rule learning for knowledge graph reasoning[EB/OL]. (2023-03-07). <https://arxiv.org/abs/2303.03581v1>.
- [13] 杨卫哲, 秦永彬, 黄瑞章, 等. 面向中文关系抽取的句子结构获取方法[J]. *数据采集与处理*, 2021, 36(3): 605-620.
- YANG Weizhe, QIN Yongbin, HUANG Ruizhang, et al. Sentence structure acquisition method for Chinese relation extraction [J]. *Journal of Data Acquisition and Processing*, 2021, 36(3): 605-620.
- [14] PAN S, LUO L, WANG Y, et al. Unifying large language models and knowledge graphs: A roadmap[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2024, 36(7): 3580-3599.
- [15] LIU X, ZHENG Y, DU Z, et al. GPT understands, too[J]. *AI Open*, 2024, 5: 208-215.
- [16] LIU X, JI K, FU Y, et al. P-tuning v2: Prompt tuning can be comparable to fine-tuning universally across scales and tasks[EB/OL]. (2021-10-14). <https://arxiv.org/abs/2110.07602v3>.
- [17] LUO R, SUN L, XIA Y, et al. BioGPT: Generative pre-trained transformer for biomedical text generation and mining[J]. *Briefings in Bioinformatics*, 2022, 23(6): bbac409.
- [18] 刘峤, 李杨, 段宏, 等. KG构建技术综述[J]. *计算机研究与发展*, 2016, 53(3): 582-600.
- LIU Qiao, LI Yang, DUAN Hong, et al. Knowledge graph construction techniques[J]. *Journal of Computer Research and Development*, 2016, 53(3): 582-600.
- [19] 刘玉华, 翟如钰, 张翔, 等. KG可视分析研究综述[J]. *计算机辅助设计与图形学学报*, 2023, 35(1): 23-36.
- LIU Yuhua, ZHAI Ruyu, ZHANG Xiang, et al. A survey on the visual analytics of knowledge graph[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2023, 35(1): 23-36.
- [20] 许雅玺, 孟天宇, 王欣, 等. 融合领域词典嵌入的航空不安全事件命名实体识别[J]. *科学技术与工程*, 2024, 24(8): 3284-3290.
- XU Yaxi, MENG Tianyu, WANG Xin, et al. Named entity recognition of aviation unsafe events embedded with fusion domain dictionary[J]. *Science Technology and Engineering*, 2024, 24(8): 3284-3290.
- [21] 刘雅萱, 武娇, 顾兴全, 等. 基于多语义文本表示的自适应模糊聚类算法[J]. *科学技术与工程*, 2023, 23(11): 4709-4717.
- LIU Yaxuan, WU Jiao, GU Xingquan, et al. Adaptive fuzzy clustering algorithm based on multi-semantic text representation [J]. *Science Technology and Engineering*, 2023, 23(11): 4709-4717.

- [22] 仁青东主, 头旦才让, 尼玛扎西. 基于改进的VSM的藏文期刊论文检索研究[J]. 计算机仿真, 2020, 37(12): 400-403, 454. Rinchen Dhondub, Thupten Tsering, Nyima Tashi. Research on Tibetan periodical papers retrieval based on improved VSM [J]. Computer Simulation, 2020, 37(12): 400-403, 454.
- [23] LIU Y, HILDEBRANDT M, JOBLIN M, et al. Neural multi-hop reasoning with logical rules on biomedical knowledge graphs[C]//Proceedings of The Semantic Web: Proceedings of the 18th International Conference, ESWC 2021. Cham, Switzerland: Springer, 2021.
- [24] BARWISE J. An Introduction to first-order logic[M]. Amsterdam: Elsevier, 1977: 5-46.
- [25] CHEN Y, GOLDBERG S, WANG D Z, et al. Ontological pathfinding[C]//Proceedings of the 2016 International Conference on Management of Data (SIGMOD'16). San Francisco, USA: ACM, 2016: 835-846.
- [26] LAO N, COHEN W W. Relational retrieval using a combination of path-constrained random walks[J]. Machine Learning, 2010, 81: 53-67.
- [27] YANG Y, SONG L. Learn to explain efficiently via neural logic inductive learning[EB/OL]. (2019-10-06). <https://arxiv.org/abs/1910.02481>.
- [28] OMRAN P G, WANG K, WANG Z, et al. Scalable rule learning via learning representation[C]//Proceedings of the International Joint Conference on Artificial Intelligence. Stockholm, Sweden: IJCAI Organization, 2018: 2149-2155.
- [29] LU S Y, LIU B, MILLS K G, et al. R5: Rule discovery with reinforced and recurrent relational reasoning[EB/OL]. (2022-05-13). <https://arxiv.org/abs/2205.06454>.
- [30] XU Z Z, YE P, CHEN H, et al. Ruleformer: Context-aware rule mining over knowledge graphs[EB/OL]. (2022-09-13). <https://arxiv.org/abs/2209.05815>.
- [31] JASON W, BOSMA M, ZHAO V Y, et al. Finetuned language models are zero-shot learners[EB/OL]. (2021-09-03). <https://arxiv.org/abs/2109.01652v4>.
- [32] LUO L, LI Y F, HAFFARI G, et al. Reasoning on graphs: Faithful and interpretable large language model reasoning[EB/OL]. (2023-10-02). <https://arxiv.org/abs/2310.01061v2>.
- [33] GALÁRRAGA L A, TEFLIOUDI C, HOSE K, et al. AMIE: Association rule mining under incomplete evidence in ontological knowledge bases[C]//Proceedings of the 22nd International World Wide Web Conference. Rio de Janeiro, Brazil: ACM, 2013: 413-422.
- [34] HINTON G E. Learning distributed representations of concepts[C]//Proceedings of the Eighth Annual Conference of the Cognitive Science Society. [S.l.]: [s.n.], 1986.
- [35] KOK S, DOMINGOS P, KOK S, et al. Statistical predicate invention[C]//Proceedings of the 24th International Conference on Machine Learning. Corvallis, USA: PMLR, 2007: 433-440.
- [36] BORDES A, USUNIER N, GARCIA-DURÁN A, et al. Translating embeddings for modeling multi-relational data[J]. Advances in Neural Information Processing Systems, 2013, 26: 2787-2795.
- [37] YANG B, YIH W T, HE X, et al. Embedding entities and relations for learning and inference in knowledge bases[EB/OL]. (2014-12-20). <https://arxiv.org/abs/1412.6575v4>.
- [38] TROUILLON T, WELBL J, RIEDEL S, et al. Complex embeddings for simple link prediction[C]//Proceedings of the International Conference on Machine Learning. New York, USA: PMLR, 2016: 2071-2080.
- [39] SUN Z, DENG Z, NIE J Y, et al. Rotate: Knowledge graph embedding by relational rotation in complex space[EB/OL]. (2018-09-27). <https://www.semanticscholar.org/paper/Rotate:Knowledge-graph-embedding-by-relational-rotation-in-complex-space/Sun-Deng/8f096071a09701012c9c279aee2a88143a295935>.

作者简介:



郑强(1979-),男,高级工程师,研究方向:特种设备、信息技术、物联网工程等, E-mail:79876099@qq.com.



许振彬(2001-),通信作者,男,硕士研究生,研究方向:区块链、知识图谱, E-mail:1547512048@qq.com.