

薛浩, 代艳格, 李美玲. 子图结构信息增强大语言模型的水务运维知识图谱补全[J]. 净水技术, 2025, 44(11): 184–194.

XUE H, DAI Y G, LI M L. Subgraph structural information-enhanced large language model for water operation knowledge graph completion[J]. Water Purification Technology, 2025, 44(11): 184–194.

子图结构信息增强大语言模型的水务运维知识图谱补全

薛浩^{1,2,*}, 代艳格², 李美玲²

(1. 南京航空航天大学经济与管理学院, 江苏南京 211106; 2. 上海威派格智慧水务股份有限公司, 上海 201800)

摘要 【目的】 水务知识图谱为水务管理提供了基于知识推理的决策支持,但其构建与应用因关键实体及关系缺失面临严峻挑战,传统补全方法难以应对水务领域实体异构性高、语义关联复杂等问题。【方法】 文章以水务运维知识图谱为研究对象,提出了一种基于大语言模型的知识图谱补全方法。首先,以缺失实体为中心,从图谱中抽取该实体的子图结构信息;其次,通过提示工程将子图结构信息动态转换为适配大语言模型的源文本序列与目标文本序列;最后通过高效参数微调,深度整合子图结构与大语言模型,提高大语言模型在水务运维知识图谱的推理能力。【结果】 模型试验结果表明,相比于仅利用知识图谱三元组信息的大模型(KG-LLM),文章方法在三元组分类的准确率提升 6.5%、关系预测 hits@1 指标上提升 6.4%,在链接预测的 hits@1、hits@3 及 hits@10 指标分别提升 1.6%、7.3% 及 5.7%。基于文章所提出的水务运维知识图谱补全方法应用于某水司智慧运维现场,泵组运行效率提升 22%,年度维护成本降低 15%,有效避免了因水泵轴承故障导致的突发停水事件。【结论】 子图结构信息能够显著增强大模型对缺失实体的推理准确性,为复杂、多样的运维知识图谱补全任务提供更简单高效的解决方法。

关键词 智慧水务 信息化建设 知识图谱 大语言模型 知识图谱补全

中图分类号: TU991 文献标志码: A 文章编号: 1009-0177(2025)11-0184-11

DOI: 10.15890/j.cnki.jsjs.2025.11.021

Subgraph Structural Information-Enhanced Large Language Model for Water Operation Knowledge Graph Completion

XUE Hao^{1,2,*}, DAI Yange², LI Meiling²

(1. College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China;

2. WPG Smart Water Public Co., Ltd., Shanghai 201800, China)

Abstract [Objective] The water knowledge graph can provide knowledge-driven decision support for water management. However, the construction and application of the water knowledge graph faces the challenge of completeness due to the lack of key entities and relationships. Traditional completion method struggles to address issues such as high entity heterogeneity and complex semantic associations in the water sector. [Methods] This paper proposed a knowledge graph completion method based on large language models for water operation maintenance knowledge graphs. The method consisted of three main steps: First, subgraph structure information of missing entities was extracted from the knowledge graph. Second, prompt engineering was used to convert the subgraph structure information into source and target text sequences suitable for large language models. Finally, efficient parameter fine-tuning was employed to integrate the subgraph structure with the language model, enhancing its reasoning capabilities in water operation and maintenance knowledge graphs. [Results] Compared to a large model that only utilizes knowledge graph triple information (KG-LLM), the proposed method improves the accuracy of triple classification by 6.5% and enhanced the hits@1 metric for relationship prediction by 6.4%. Additionally, in link prediction, the hits@1, hits@3, and hits@10 metrics were improved by 1.6%, 7.3%, and 5.7%, respectively. The knowledge graph completion method proposed in this paper was applied in a smart operation and maintenance site of a water company, resulting in a 22% increase in pump group operating efficiency, a 15% reduction in annual

[收稿日期] 2025-03-03

[基金项目] 国家自然科学基金面上项目(72174086)

[通信作者] 薛浩(1985—),男,博士研究生,研究方向为大数据分析及应用,E-mail: iamxuehao@126.com。

maintenance costs, and effectively preventing incidents of sudden water supply interruptions caused by pump bearing failures. [Conclusion] Subgraph structure information can significantly enhance the inference accuracy of large language models for missing entities, and provide a simpler and more efficient method for complex and diverse operation and maintenance knowledge graph completion tasks.

Keywords smart water informatization construction knowledge graph large language model knowledge graph completion

随着智慧水务的快速发展,水务管理面临着数据孤岛和信息碎片化的问题,如何高效整合各类运维数据并提升决策支持能力,成为亟待解决的挑战^[1]。知识图谱是一种语义网络,能够通过节点和边表示实体及其关系,实现知识的结构化和语义化表达^[2]。因此,水务知识图谱通过构建“水源—水厂—管网—用户”全链条语义网络,不仅可以不同来源、不同格式的数据进行统一建模和语义关联,有效消除数据孤岛,实现数据的互联互通和综合分析,而且能够将水务行业专家的经验 and 知识系统化、结构化地表示,便于知识的管理、共享和应用。同时,知识图谱支持复杂的查询和推理,能够为水务管理提供智能化的决策支持^[3-4]。

然而,构建一个数据丰富且准确的水务全领域知识图谱需要大量的技术和人力成本,水务知识图谱的不完备性仍然是一个普遍存在的问题,即知识图谱中的许多实体及实体间的关系可能存在缺失,进而导致水务知识图谱无法全面准确地表征现实世界的知识结构。这种不完备性会对基于水务知识图谱的应用产生不利影响,如降低水务知识问答的准确性。因此,如何有效地补全水务知识图谱中的缺失信息,是构建智慧水务应用体系的重要研究方向。国内外针对知识图谱补全方法已有大量研究,按发展历程和方法类型可大致分为几个阶段。早期传统表示学习方法,将实体或关系映射到低维向量空间中,通过向量运算表示语义关系,能有效解决数据稀疏性问题,但难以捕捉复杂关系模式,尤其涉及多个实体的复杂关系^[5-8]。随着深度学习技术发展,知识图谱补全方法过渡到基于神经网络的方法^[9-12],通过多层卷积学习高维特征表示,捕捉实体和关系间语义关联,在处理复杂关系时表现更好。近年来,图神经网络(graph neural networks, GNNs)逐渐成为知识图谱补全的主流方法^[13-15],能够通过聚合节点的局部邻域信息,捕捉知识图谱中的全局和局部信息,从而提高补全的效果。但随着知识图谱规模的增大,GNNs需要处理的节点和边的数量也急剧增

加,导致模型的训练和推理过程更加耗时,计算资源需求大幅上升。

随着大语言模型如 GPT-3^[16]等的出现,为知识图谱补全提供了新的范式。部分研究^[17-19]初步探讨了大语言模型在图谱补全及知识推理等任务中的应用,采用通用大语言模型结合 Zero-Shot 或 One-Shot 的方式来完成特定的任务,有一定的效果,但与经过微调的预训练语言模型相比仍存在一定差距。且这些研究大多集中于单一任务,如实体预测,在三元组分类、关系预测等任务上的探索较少。

上述研究表明,大语言模型已在其他领域的图谱补全取得了一定的进展,但依然面临对图谱结构信息利用不全的挑战。本研究利用知识图谱的子图结构信息来增强大语言模型,旨在提高水务知识图谱补全的准确性和效率。本文提取待补全三元组在已知知识图谱中的子图结构信息,并通过提示工程将这些子图结构信息转换为大语言模型能够理解和使用的源文本序列。缺失的知识图谱信息则作为目标文本序列,通过微调的方式对大语言模型进行优化,从而提高其在水务知识图谱补全任务中的表现。

1 数据与方法

1.1 数据采集与处理

本文所使用的水务运维知识图谱(water operation and maintenance knowledge graph, WOMKG)由实验室基于内部模拟数据构建,模拟数据涵盖设备运行状态和运维场景,旨在评估模型的效果,未涉及任何客户的真实数据。在实际应用过程中,该模型仅在客户现场进行结果验证,未使用客户的原始数据用于论文试验,以保障数据的隐私与安全。

WOMKG 共包含了 26 942 条三元组,24 504 个实体,26 种关系。针对 WOMKG 补全中常涉及的三元组分类、关系预测、链接预测任务构建数据集,从 WOMKG 中随机选择 70% 的三元组作为知识图谱中完备的部分,记为 WOMKG-complete,共 18 860 条。剩下的 30% 作为知识图谱中不完备的部分,记为

WOMKG-incomplete,共 8 082 条。对于三元组分类任务,将 WOMKG-incomplete 均分为 2 组,其中一组不作处理,作为正确的三元组标签,另一组随机选择替换实体为其他的实体,或者替换关系为其他的关系类型,作为错误的三元组标签,修改后的数据集记为 WOMKG-incomplete-clis;对于关系预测任务,将 WOMKG-incomplete 中的三元组去掉关系,只保留头尾实体,作为关系预测的数据集,记为 WOMKG-incomplete-rel;对于链接预测任务,将 WOMKG-incomplete 随机去掉头实体或尾实体,记为 WOMKG-incomplete-lin。再分别对 3 个任务的数据集按照 6 : 1 : 3 划分为训练集、验证集、测试集,数量分别为 4 850、808、2 424 条。

1.2 任务定义

将 WOMKG 定义 G 如式(1)^[20]。

$$G = \{(h, r, t)\} \subseteq E \times R \times E \quad (1)$$

其中: h ——头实体;

t ——尾实体;

r ——头实体与尾实体之间的关系;

E ——实体的集合;

R ——关系的集合。

WOMKG 补全任务的目标是通过已有的三元组信息,预测缺失或不完整的三元组。具体来说,WOMKG 补全包括三元组分类、关系预测及链接预测,具体如图 1 所示。

1) 三元组分类

给定三元组 (h, r, t) , 判断该三元组是否为知识图谱中的正确事实,定义如式(2)。

$$f(h, r, t, G) = \begin{cases} 1, & \text{if } (h, r, t) \in G \\ 0, & \text{else} \end{cases} \quad (2)$$

其中: $f(h, r, t, G)$ ——判断该三元组是否正确的模型,正确输出 1,错误输出 0。

2) 关系预测

给定 h 和 t , 预测它们之间的关系 r , 定义如式(3)。

$$r = \operatorname{argmax}_{\hat{r} \in R} f(\hat{r} | h, t, G) \quad (3)$$

其中: $f(\hat{r} | h, t, G)$ ——输出实体间的关系概率

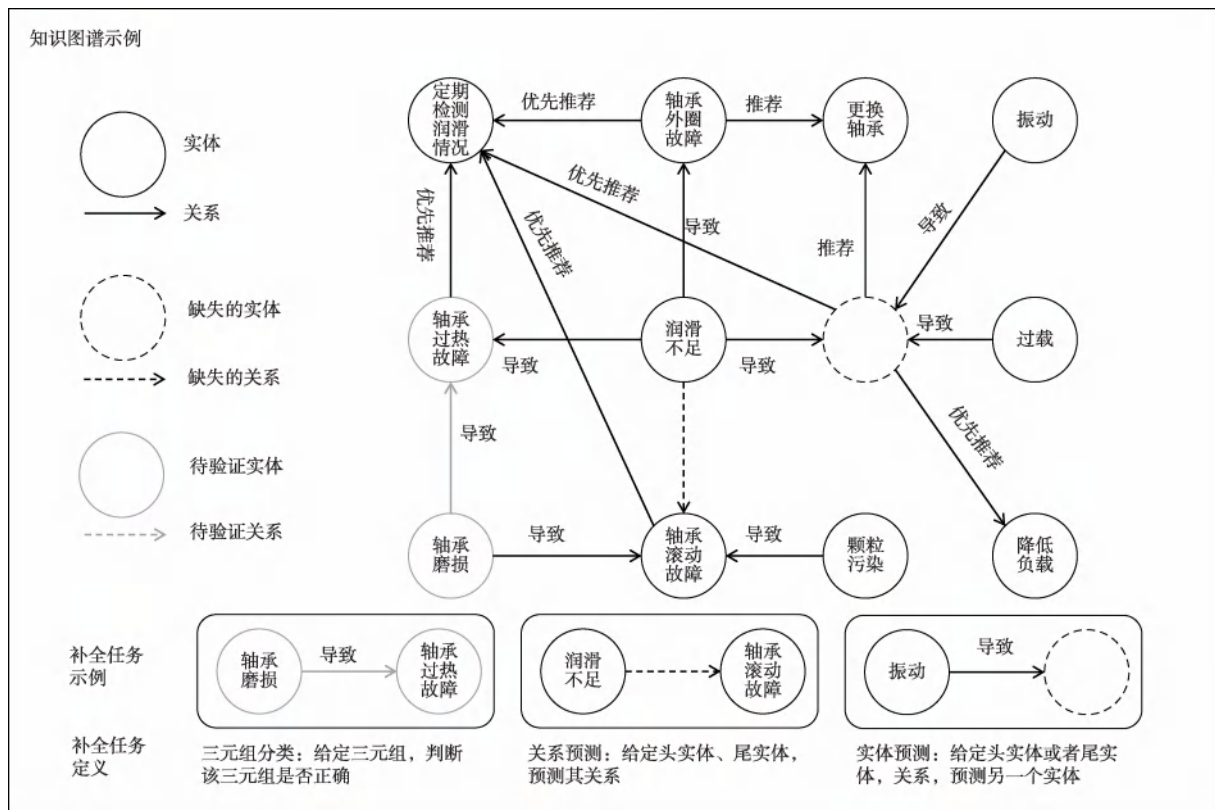


图 1 知识图谱及补全任务示意图(以水泵故障运维为例)

Fig. 1 Illustration of Knowledge Graph and Completion Task (Example of Water Pump Fault Diagnosis)

的模型;

argmax ——选择概率最大的关系。

3) 链接预测

给定头实体 h 和关系 r , 预测尾实体 t , 或给定尾实体 t 和关系 r , 预测头实体 h , 定义如式(4)。

$$t = \underset{\hat{t} \in E}{\text{argmax}} f(\hat{t} | h, r, G), h = \underset{\hat{h} \in E}{\text{argmax}} f(\hat{h} | r, t, G) \quad (4)$$

其中: $f(\hat{t} | h, r, G)$ ——输出尾实体概率的模型;

$f(\hat{h} | r, t, G)$ ——输出头实体概率的模型。

1.3 研究方法

当前大语言模型的知识图谱补全方法主要以三元组为输入,利用其广泛的背景知识和强大的推理能力进行补全。然而,这些方法往往忽略了缺失三元组的子图结构信息,这些子图包含了缺失三元组的上下文知识,与大语言模型的隐式知识相结合,可以进一步增强其推理能力。因此,本文提出了一种子图结构信息增强的大语言模型 WOMKG 补全方法,该方法提出了一种新的基于大语言模型的图谱补全多任务融合策略,将传统的三元组分类、链接预

测及关系预测任务融合为一个生成式任务。主要流程如图 2 所示,分为子图结构信息抽取、提示工程构建、微调与补全 3 个部分。(1)子图结构信息抽取:对于 WOMKG 中缺失的三元组,根据任务类型从已有的知识图谱中抽取相关的子图结构信息。(2)提示工程构建:根据补全任务类型设计相应的描述模板,将采样的子图转化为大语言模型的提示文本序列,将三元组中缺失的信息转化为大语言模型的响应文本序列,从而将知识图谱补全任务转化为序列生成任务。(3)微调与补全:使用高效参数微调算法 P-Tuning v2^[19] 在构建的文本序列数据集上对大语言模型进行微调,得到 WOMKG 补全的大模型,在应用阶段,模型的响应即为知识图谱补全的结果。通过引入子图结构信息,本方法能够更加充分地利用大语言模型的推理能力,提高知识图谱补全的准确性和效果。

2 WOMKG 补全模型构建

2.1 子图信息提取

本文通过动态抽取与任务相关的运维子图结构信息,并将其转化为自然语言形式,以显式的知识补充提供给大语言模型,利用其强大的推理能力进行

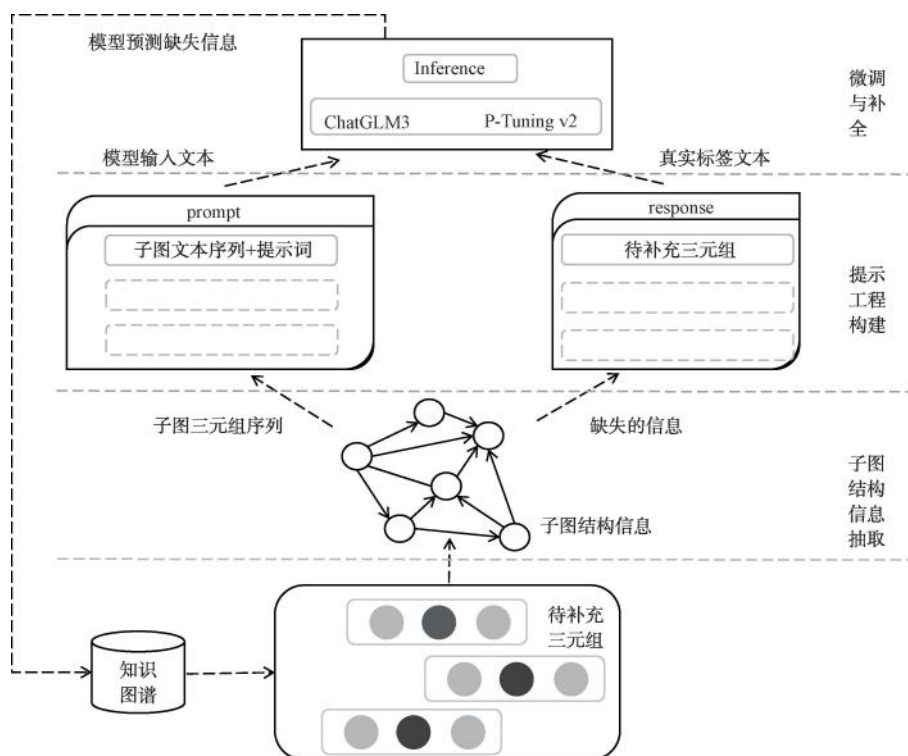


图 2 子图结构信息增强的大语言模型知识图谱补全流程

Fig. 2 Knowledge Graph Completion of Large Language Model Based on Subgraph Structure Information Enhancement

WOMKG 补全。子图结构信息的抽取主要是指在知识图谱中,抽取所有与缺失三元组相关联的节点及其对应的关系,可以用如式(5)。

$$\text{sub}G = \{(x, y, z) \mid (x, y, z) \in G, \\ x \in (h_j, t_j) \text{ or } z \in (h_j, t_j)\} \quad (5)$$

其中: h_j ——待补充三元组头节点;
 t_j ——待补充三元组尾节点;
 (x, y, z) ——知识图谱中的三元组。

子图抽取的伪代码如图 3 所示。

```

Algorithm 1: 子图抽取算法
Input: 知识图谱  $G = \{(x, y, z)\} \subseteq E \times R \times E$ , 待补充三元组  $(h_j, r_j, t_j)$ ,  $T = |G|$ 
Output:  $\text{sub}G \subseteq E \times R \times E$ 
1 for  $(x_i, y_i, z_i), i = 1, 2, \dots, T$  do
2   if  $x_i = h_j$  or  $x_i = t_j$  or  $z_i = h_j$  or  $z_i = t_j$ 
3     if  $(x_i, y_i, z_i) \neq (h_j, r_j, t_j)$  AND  $(x_i, y_i, z_i) \notin \text{sub}G$ 
4        $\text{sub}G.add((x_i, y_i, z_i))$ 
5   end
6 end
7 end

```

图 3 子图抽取伪代码

Fig. 3 Pseudocode for Subgraph Extraction

实际上, WOMKG 以文本文件形式存储, 其中每一行都是一条三元组的文本信息。图 3 的计算量主要集中在判断当前三元组 (x_i, y_i, z_i) 是否和待补充三元组 (h_j, r_j, t_j) 存在相同的实体, 每次遍历影响其时间复杂度的因素主要包含 2 个判断语句及一个列表的添加语句, 时间复杂度为常量阶 $O(1)$, 循环语句一共迭代了 T 轮, 则每个待补充三元组的子图抽取时间复杂度为线性阶 $O(T)$ 。假设待补充三元组的个数为 N 则子图抽取的复杂度为所有单个子图抽取复杂度总和, 估算为 $O(NT)$ 。空间复杂度的影响因素主要取决于待补充三元组 (h_j, r_j, t_j) 与知识图谱 G 中三元组 (x_i, y_i, z_i) 的关联程度。每一次遍历, 如果存在相同的实体, 则需要为三元组 (x_i, y_i, z_i) 开辟常量阶的存储空间 $S(1)$, 遍历一轮最多需要开辟线性阶存储空间 $S(T)$ 。对所有待补充的 N 个三元组, 所需的存储空间为 $S(NT)$, 则子图抽取的空间复杂度估算介于 0 与 $S(NT)$ 之间。

2.2 提示工程构建

为了更好地利用子图结构信息, 需要将其转化为大模型适用的文本形式。采用提示学习的方式来构造知识图谱补全模型的输入文本序列, 其中将子图信息转化为子图文本子序列的计算如式(6)。

$$\text{sub}G_text = \text{concat}(' ', \sum_{k=1}^N \\ (\text{concat}(' ', h_k, r_k, t_k))), (h_k, r_k, t_k) \in \text{sub}G \quad (6)$$

其中: $\text{sub}G$ ——子图三元组集;

(h_k, r_k, t_k) ——子图中的一条三元组;

$\text{concat}(' |')$ ——表示使用符号“|”进行文本拼接操作;

N ——对应子图中三元组个数。

式(6)首先通过空格符号将单个三元组文本拼接, 然后通过符号“|”将子图中的不同三元组文本拼接在一起, 形成子图信息的文本子序列 $\text{sub}G_text = "h_1 r_1 t_1 | h_2 r_2 t_2 | \dots | h_N r_N t_N"$ 。进一步, 将子图信息文本子序列 $\text{sub}G_text$ 构造为知识图谱补全模型的输入文本序列 I_j , 公式如式(7)。

$$I_j = \text{concat}(' ', \text{prefix_text}, \text{sub}G_text, p_j) \quad (7)$$

其中: prefix_text ——前缀子序列, 不同任务类型下统一用文本“已知知识图谱三元组”表示;

$\text{sub}G_text$ ——由子图 $\text{sub}G$ 中的三元组转化而来的子图信息文本子序列;

p_j ——提示文本子序列。

模型的响应则是给出缺失的关系或实体, 或判断三元组是否正确。不同任务下的提示文本及响应文本如表 2 所示。

以图 1 中判断三元组(轴承磨损, 导致, 轴承过热故障)是否正确为例, I_j 中 3 个子序列如下。

1) prefix_text : 前缀文本子序列为“已知知识图谱三元组”。

2) $\text{sub}G_text$: 将子图集合 $\{(轴承磨损, 导致, 轴承滚动故障), (轴承过热故障, 优先推荐, 定期检查润滑情况), (润滑不足, 导致, 轴承过热故障)\}$ 按照式(2)进行拼接, 拼接后的子图信息文本子序列为“轴承磨损 导致 轴承滚动故障 | 轴承过热故障 优先推进 定期检查润滑情况 | 润滑不足 导致 轴承滚动故障”。

3) p_j : 参照表 1, 将提示文本“判断 h, r, t 是否正确?”中的 h, r, t 用真实的三元组进行替换, 替换后的提示文本子序列为“判断轴承磨损 导致 轴承过热故障是否正确?”。

最后,将 3 个文本子序列使用空格符号进行拼接,即得到知识图谱补全模型的输入文本序列 I_j 。

对于图 1 中的 3 个补全子任务提示文本按照式 (2) 及式 (3) 构建,即可得到补全模型的输入文本,输入及输出的示例如表 3 所示。

表 2 知识图谱补全任务提示文本示例

Tab. 2 Examples of Prompt Text for Knowledge Graph Completion Task

任务类型	提示文本	响应文本
链接预测(头实体)	[MASK] r t 中的 [MASK] 是什么?	h
链接预测(尾实体)	h r [MASK] 中的 [MASK] 是什么?	t
三元组分类(正确)	判断 h, r, t 是否正确?	正确
三元组分类(错误)	判断 h, r, t 是否正确?	错误
关系预测	h 与 t 之间的关系是什么,请从 { 关系集合 } 中选择	r

表 3 知识图谱补全模型输入输出示例

Tab. 3 Input and Output Examples for Knowledge Graph Completion Models

任务类型	输入示例	输出示例
链接预测	已知知识图谱三元组 过载 导致 [MASK] [MASK] 优先推荐 降低负载 润滑不足 导致 [MASK] [MASK] 优先推荐 定期检查润滑情况 [MASK] 推荐 更换轴承。振动 导致 [MASK] 中的 [MASK] 是什么?	轴承内圈故障
三元组分类	已知知识图谱三元组 轴承磨损 导致 轴承滚动故障 轴承过热故障 优先推进 定期检查润滑情况 润滑不足 导致 轴承滚动故障。判断 轴承磨损 导致 轴承过热故障 是否正确?	正确
关系预测	已知知识图谱三元组 润滑不足 导致 轴承外圈故障 润滑不足 导致 轴承过热故障 颗粒污染 导致 轴承滚动故障 轴承磨损 导致 轴承滚动故障 轴承滚动故障 优先推荐 定期检索润滑情况。润滑不足与轴承滚动故障之间的关系是什么,请从 导致 推荐 ... 优先推荐 中选择	导致

2.3 模型构建

随着模型参数规模的不断增大,尤其是在大语言模型的应用场景下,在消费级硬件上进行全量微调的可行性逐渐降低。近年来,研究者们提出了各种高效的迁移学习方法,即在固定大语言模型大部分参数的前提下,仅调整小部分参数,以达到与全部参数微调接近的效果。选取高效的参数微调方法 P-Tuning v2^[19],在 WOMKG 训练数据上对大语言模型进行微调。该微调算法在性能、资源效率上等多个方面均具备显著优势,是目前大模型微调的主流算法。与仅在 Transformer 第一层输入中插入 Prompts 的 P-Tuning v1 相比,P-Tuning v2 在每一层都加入了 Prompts 作为输入,从而引入了更多可学习的参数。加入到更深层结构中的 Prompts 能对模型预测产生更直接的影响。

试验采用 ChatGLM3-6B 提供的微调套件,将构造的输入文本序列和响应文本序列按照规定格式进行组织,运行微调脚本。训练完成后,得到微调后的 P-Tuning v2 参数。推理时,加载 ChatGLM3-6B 的模型文件及 P-Tuning v2 的模型参数,并从推理的结果

序列中提取缺失的元素,完成知识图谱补全任务。

大语言模型微调与推理流程如图 4 所示。首先抽取缺失三元组的子图信息,并构建提示工程,得到微调与推理的文本数据;然后通过 P-Tuning v2 的方式微调,输入文本序列转化为向量形式,在 ChatGLM3-6B 模型的每一层,包括输入层,引入长度为 L 的 Prompts 前缀向量;该向量进入到 GLM 架构内部各层与 Key、Value 向量拼接,参与到注意力分数的计算;经过输出层处理后得到预测的响应文本序列,与真实的响应文本序列一起计算损失并反向传播,此时只更新 P-Tuning v2 中每层 Prompts 向量的参数,ChatGLM3-6B 模型的参数保持不变,不断迭代,直至微调完成。推理时,加载微调好的 P-Tuning v2 的参数和 ChatGLM3-6B 的参数,输入缺失三元组对应构建的文本序列,模型输出响应文本序列,抽取三元组中缺失的元素进行补全。ChatGLM3-6B 初始模型及所提方法微调后的 WOMKG 补全大模型的接口调用流程如图 5 所示。首先,对于知识图谱中缺失的三元组,构建相应的子图,并通过提示工程得到模型的输入文本序列。接下来,加载模型文件

及词向量文件;如果使用的是微调后的模型,还需要加载 P-Tuning v2 的参数文件。最后,模型生成对应

的输出文本序列,从中抽取缺失的内容,并补全到知识图谱中。

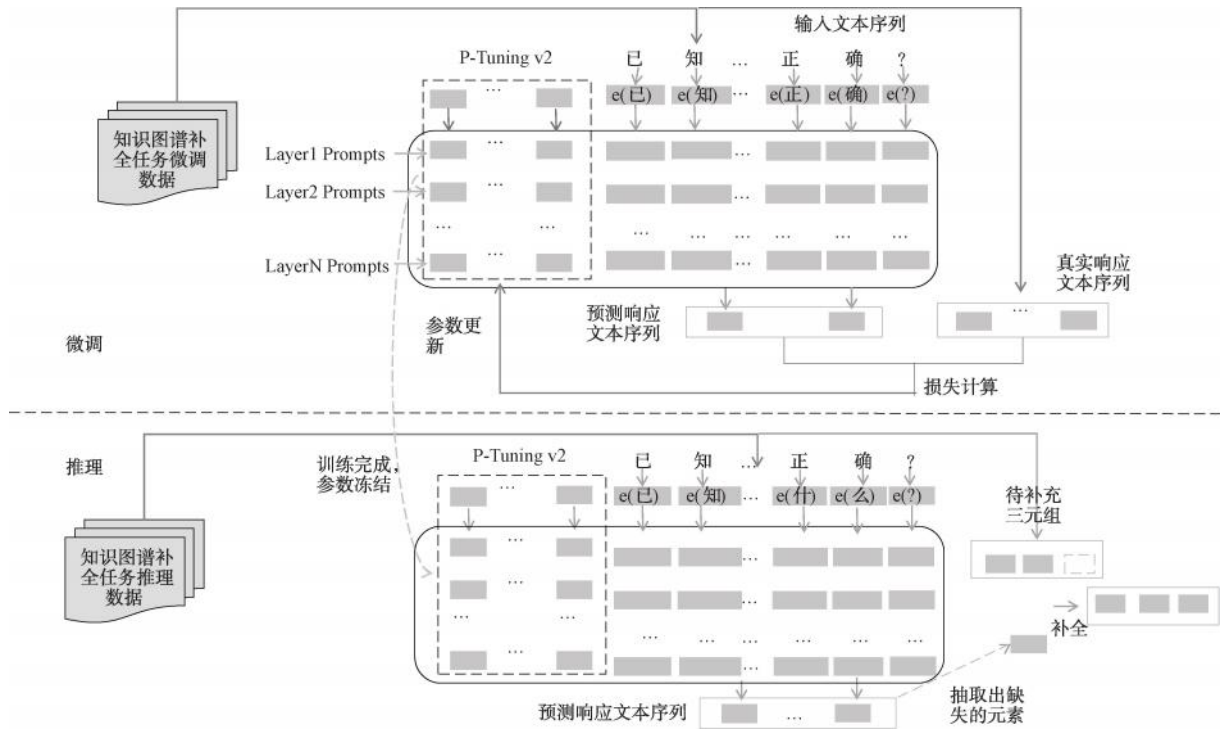


图4 知识图谱补全大模型微调及推理流程

Fig. 4 Knowledge Graph Completion: Large Model Fine-Tuning and Inference Workflow

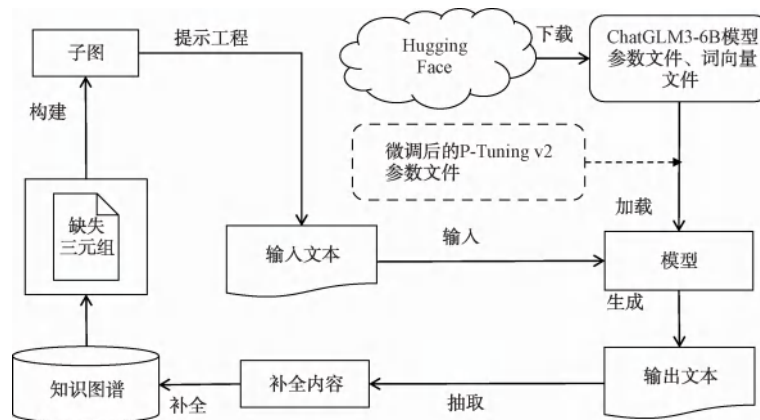


图5 知识图谱补全任务模型流程

Fig. 5 Model Interfaces of Knowledge Graph Completion Task

3 试验过程与结果讨论

3.1 模型参数设置

ChatGLM3 是智谱 AI 和清华大学 KEG 实验室联合发布的对话预训练模型。ChatGLM3-6B 是 ChatGLM3 系列中的开源模型,具有对话流畅、部署门槛低等众多优良特性。综合考虑大语言模型参数大小、模型效果、模型部署与微调门槛等因素选择

ChatGLM3-6B 作为基础模型。本研究在一台配备 512 G 内存和 4 张 Nvidia RTX 3090 显卡的服务器上进行试验。模型的微调与推理均基于 PyTorch 框架,微调过程中采用 Adam 优化器,其中一阶矩估计的指数衰减率设置为 0.9,二阶矩估计的指数衰减率为 0.99,初始学习率设置为 0.02。输入序列包含三元组的子图信息,构建的序列较长,因此,在试

验中将 ChatGLM3-6B 输入的最大长度设置为 1 024 个 token,输出的最大长度设置为 128 个 token。P-Tuning v2 微调方法中 Prompts 的长度 L 设置为 128 个 token。微调过程中,验证集的 batch_size 设置为 4,梯度累计步数设置为 16,微调中模型的 checkpoint 间隔设置为 500。

为了验证提出的模型的效果,文章选择了一系列方法进行对比试验,如嵌入表示的 TransE、TransH、TransD、TransR 等,神经网络及图网络模型 ConvKB、DKRL、DRR-GAT、HRGAT 等,预训练语言模型 KG-BERT、StAR、KGT5 等,以及利用大语言模型进行知识图谱补全任务的 KG-LLM。上述模型基于不同范式,在不同时期推动了知识图谱补全任务的发展,能够帮助本研究验证所提出的方法在知识图谱补全任务中的有效性和适用性。

选择准确率作为 WOMKG 中三元组分类子任务的评价指标,衡量三元组的正确性,准确率指标越高越好,当准确率指标为 1 时表示模型的判断全部正确。选择命中率指标 hits@ 1 作为关系预测的评价指标,hits@ 1、hits@ 3 及 hits@ 10 作为链接预测的评价指标。hits@ k 表示预测的前 k 个结果中存在正确结果的准确率, k 值越小对模型的要求越高,hits@ k 指标越高表示模型的前 k 个结果预测越准确。

3.2 补全效果评估

在构造的 WOMKG-incomplete-cls、WOMKG-incomplete-rel、WOMKG-incomplete-lin 上分别针对三元组分类、关系预测、链接预测补全子任务进行补全试验。选择不同时期的最先进方法作为对比,试验结果如表 4 所示。

表 4 中各评估指标中加粗值为最优结果,下划线表示次优结果。本文所提的融合了子图结构的大语言模型补全方法在 WOMKG 补全中的三元组分类、关系预测及链接预测的 hits@ 3、hits@ 10 指标上都取得了最好的结果。相比于次优方法,各指标上提升超过 6%,表明了所提方法的有效性。传统的嵌入表示补全方法仅通过低维向量建模实体间线性关系,无法捕捉复杂语义,如“泵组效率”与“维护”成本之间的非线性关联;而 GNNs 方法需依赖显式图结构(如邻接矩阵)进行训练,对知识图谱的完整性和数据质量敏感,但在 WOMKG 场景中存在大量实体缺失的情况。因此,GNNs 方法在此场景难以

表 4 水务知识图谱补全效果评估结果
Tab. 4 Evaluation Results of Knowledge Graph Completion for Water Management

模型	三元组 分类选择 准确率	关系预测 hits@ 1	链接预测		
			hits@ 1	hits@ 3	hits@ 10
TransE	76.5%	61.2%	35.4%	48.6%	60.7%
TransH	78.3%	64.5%	37.6%	50.4%	68.8%
TransR	77.5%	64.2%	37.2%	51.6%	70.4%
TransD	84.1%	69.5%	34.5%	47.8%	62.2%
ConvKB	83.8%	73.6%	39.1%	60.7%	79.4%
DKRL	74.8%	70.5%	36.2%	58.9%	76.5%
KG-BERT	85.4%	78.2%	42.1%	67.1%	84.3%
StAR	86.4%	<u>79.2%</u>	46.3%	72.4%	88.5%
KGT5	80.3%	67.1%	41.9%	70.5%	87.2%
KG-LLM	<u>90.2%</u>	78.3%	43.7%	<u>74.3%</u>	<u>89.7%</u>
本文	96.7%	84.7%	<u>45.3%</u>	81.6%	95.4%

达到最优的效果;预训练语言模型依赖通用训练语料,遵循长尾分布,低频实体描述信息不足,且知识迁移成本较高。本文提出的方法利用了参数量巨大的大语言模型,能够有效捕捉非线性关系,且具有上下文学习的能力,通过少量的样本学习,就能将通用领域的知识推理能力适配到水务运维图谱补全领域。此方法也全面超过了利用大语言模型的 KG-LLM,表明了子图结构信息能够为不完备的 WOMKG 提供丰富的上下文信息,再结合大语言模型的推理能力,提高大模型的补全准确率。该方法在链接预测 hits@ 1 上要略低于预训练的语言模型方法,由于 hits@ 1 指标的要求更为严格,大语言模型有着更丰富的背景知识,更高的推理能力,但也存在冗余的信息影响了 hits@ 1 上的效果,对要求稍微放宽的 hits@ 3、hits@ 10 指标则提升更加明显。

3.3 消融试验

为了验证子图结构(Subgraph)与微调(P-Tuning v2)的在 WOMKG 补全任务上的有效性,文章设计了消融试验进行验证,通过从模型中移除掉不同的模块,研究这些模块对于模型性能的影响,试验结果如表 5 所示。

表 5 中“w/o”表示在 baseline 的基础上移除掉该模块,在移除掉微调与子图结构时,模型性能大幅度下降,甚至不如此前的嵌入表示、图网络及预训练语言模型方法。这主要是因为通用的大语言模型难

表 5 消融试验结果
Tab. 5 Results of Ablation Experiment

模型	三元组分类 选择准确率	关系预测 hits@ 1	链接预测		
			hits@ 1	hits@ 3	hits@ 10
baseline(本文)	96. 7%	84. 7%	45. 3%	81. 6%	95. 4%
w/o Subgraph	91. 5%	77. 9%	41. 5%	73. 5%	90. 3%
w/o P-Tuning v2	83. 3%	74. 5%	37. 6%	66. 8%	83. 7%
w/o P-Tuning & Subgraph	50. 8%	43. 7%	30. 5%	45. 6%	65. 1%

以理解水务运维的专业知识,又缺乏足够的上下文使得大模型的推理产生了较为严重的“幻觉”,导致模型的准确率不佳。结合子图结构信息后的模型(w/o P-Tuning v2)评价指标提升较大,尤其是在较为简单的运维图谱三元组分类及关系预测任务上,提升超过 30%。分析不同模型的输出结果发现,“w/o P-Tuning & Subgraph”的输出中回复类似于“对不起,我没有任何的信息回答您提供的问题”占比为 30. 3%。但在采用子图结构信息后,占比下降为 7. 4%,子图结构信息转化成的提示词一方面能够充当样例的角色,加强通用大语言模型的任务意图对齐能力,另一方面能够为大语言模型提供一定的上下文信息,加强通用大语言模型的推理能力。再进一步结合高效参数微调,模型准确率进一步提升,微调能够有效地整合大语言模型与子图结构信息,更好地将通用大模型转化为水务运维的知识图谱补全大模型。

3. 4 鲁棒性分析

在实际的水务运维场景中不完备知识图谱的缺失率各不相同,本文采用子图结构信息增强大语言模型,子图的构造与缺失率存在一定的关联。为验证本文所提方法在不同缺失率下的鲁棒性,将 WOMKG 的缺失率分别设置为 0. 2、0. 3、0. 4、0. 5、0. 6、0. 7、0. 8。按照 1. 1 节所述构造数据集,针对三元组分类、关系预测、链接预测子任务进行补全试验,不同缺失率下的模型效果如图 6 所示。

随着运维知识图谱缺失率的上升,所提方法在不同任务上的指标均呈现下降的趋势,但不同的任务表现出不同的幅度,链接预测中 hits@ 10 指标下降幅度最小,三元组分类、关系预测任务次之,链接预测 hits@ 1 下降幅度最大。hits@ 10 指标对模型准确率的要求较为宽松,即使存在较大的缺失率,所提方法依然有较高的准确率;而三元组分类任务较

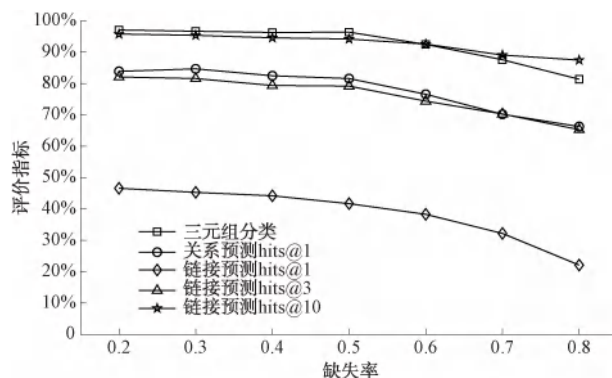


图 6 不同缺失率下评价指标

Fig. 6 Evaluation Metrics under Different Missing Rates

为简单,对子图的依赖较低,因此,下降幅度也较小;链接预测任务中存在大量的相似实体及实体的长尾分布,缺失率上升导致难以构建子图结构信息,因此,链接预测 hits@ 1 指标下降幅度最大。从下降幅度来看,缺失率为 0. 5 是一个较为明显的拐点,当缺失率小于 0. 5 时,知识图谱仍能保持基本拓扑结构,实体间的核心关联路径得以保留,模型可通过局部子图推理弥补信息缺失;但当缺失率超过 0. 5 后,图谱连通性被破坏,关键桥接实体大量丢失(如运维场景中的核心设备节点),导致子图结构信息呈现碎片化。此时,大语言模型难以从离散的上下文片段中重建有效逻辑链,性能出现断崖式下降。这表明所提方法的准确率与缺失率并不是线性相关,当缺失率超过某一个拐点时,子图信息不足,无法为大语言模型提供足够的上下文信息进行推理,导致模型性能下降较快。在实际应用场景中,缺失率超过 0. 5 的情况较为少见,缺失率小于 0. 5 时,模型在三元组分类准确率、关系预测 hits@ 1、链接预测 hits@ 3 及 hits@ 10 指标上均能保持 80% 以上。因此,本文方法在实际应用中具有一定的鲁棒性。

4 结论

本文针对 WOMKG 中关键实体及关系缺失的

不完备问题,提出了一种子图结构信息增强的大语言模型 WOMKG 补全方法。该方法的核心思路在于通过精心设计的提示工程动态,构建与补全任务相关的子图结构描述,通过微调深度融合大模型的知识推理能力与子图结构提供的关键运维上下文信息。本方法的主要优势体现在:补全精准性与鲁棒性,通过引入子图结构信息和对大模型微调,显著提高知识图谱的补全准确率。且在不同知识缺失情况下,模型依然保持良好的鲁棒性;补全高效性与客观性,相比人工校验,本方法实现自动化、规模化知识补全,节省时间和成本,避免主观偏差,保证知识的及时性和一致性。

在实际应用层面,本方法通过提升知识图谱的完备性与准确性,为水务运维带来了切实的效益。如在某地区二次供水泵房智能化改造实践中,针对泵房内各类关键设备(如水泵、电机、传感器等)梳理其历史故障记录、维修流程及运维人员经验知识,构建了“设备—状态—故障类型—成因—解决方案—资源(人员、工具)”的结构化设备运维知识图谱,并采用本文提出的图谱补全方法,自动补全了图谱中缺失的实体及关系,针对老旧设备,系统自动触发预防性维护工单,提供维保建议。系统运行后,当检测到某老旧泵组出现持续高频振动和电流波动时,系统基于图谱中“设备型号—典型症状—高发故障—预防性措施”的结构化知识链条进行检索,自动判定其由轴承磨损导致,并据此生成预防性维保工单,明确更换时间节点和所需备件。通过实施预防性维护保养措施,泵组运行效率实现 22% 的提升,同时有效避免了 3 起因老旧水泵轴承故障引发的突发停水事故;当泵组在运行中出现异常停机时,系统自动调用图谱中“故障现象—可能原因—历史案例—维修方案—人员匹配”路径,进行多轮推理,快速定位问题成因(如联轴器松脱或轴承故障),并匹配过往的典型维修记录,生成包含操作指引、配件需求与推荐责任人的维修建议单,直接推送至相关维保人员,设备故障平均修复时间由原先约 60 min 缩短至 40 min,维修效率提高 33%。

综上所述,本文提出的 WOMKG 补全方法,有效增强大模型的知识补全能力,提升了图谱中潜在的因果关系、应对方案等知识的覆盖度与实用性,并支撑了快速响应与决策辅助,推动水务行业智能化发展,提高安全性和降本增效,为行业数字化转型提

供技术支持。

参考文献

- [1] 赵立. 浅谈智慧水务发展对传统供水企业的启发[J]. 净水技术, 2019, 38(s2): 106-108.
ZHAO L. Inspiration of traditional water supply enterprises in the development of smart water platform [J]. Water Purification Technology, 2019, 38(s2): 106-108.
- [2] 高凯丽. 水务知识图谱构建及问答应用系统的研究与实现[D]. 北京: 北京工业大学, 2020.
GAO K L. Research and implementation of water affairs knowledge graph construction and question answering application system[D]. Beijing: Beijing University of Technology, 2020.
- [3] 宫强. 我国水务企业信息化建设主要问题分析与建议[J]. 净水技术, 2021, 40(s1): 363-366.
GONG Q. Analysis and suggestions on the main problems of water utility informatization construction in China [J]. Water Purification Technology, 2021, 40(s1): 363-366.
- [4] 吴秋明, 陈磊, 刘婉春, 等. 水利设备故障知识图谱建设探讨[J]. 长江技术经济, 2024, 8(1): 126-131, 138.
WU Q M, CHEN L, LIU W C, et al. Construction of knowledge mapping of water equipment failures [J]. Technology and Economy of Changjiang, 2024, 8(1): 126-131, 138.
- [5] 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. DOI: 10.5555/2999792.2999923.
- [6] WANG Z, ZHANG J, FENG J, et al. Knowledge graph embedding by translating on hyperplanes [C]//AAAI Press. Proceedings of the AAAI Conference on Artificial Intelligence, 2014.
- [7] LIN Y K, LIU Z Y, SUN M S, et al. Learning entity and relation embeddings for knowledge graph completion [C]//AAAI Press. Proceedings of the AAAI Conference on Artificial Intelligence, 2015.
- [8] JI G L, HE S Z, XU L H, et al. Knowledge graph embedding via dynamic mapping matrix [C]. Beijing: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, 2015.
- [9] XIE R B, LIU Z Y, JIA J, et al. Representation learning of knowledge graphs with entity descriptions [C]//IEEE. Proceedings of the AAAI Conference on Artificial Intelligence, 2016.
- [10] DETTMERS T, MINERVINI P, STENETORP P, et al. Convolutional 2D knowledge graph embeddings [C]//AAAI Press. Proceedings of the AAAI Conference on Artificial Intelligence, 2018, 32(1).
- [11] JIANG X, WANG Q, WANG B. Adaptive convolution for multi-

- relational learning[C]//BURSTEIN J, DORAN C, SQLORIO T. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics; Human Language Technologies, 2019.
- [12] NGUYEN D Q, NGUYEN T D, NGUYEN D Q, et al. A novel embedding model for knowledge base completion based on convolutional neural network[C]//WAIKER M, JI H, STENT A. Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics; Human Language Technologies, 2018.
- [13] SCHLICHTKRULL M, KIPF T N, BLOEM P, et al. Modeling relational data with graph convolutional networks[C]//BERTINO E, GAO W, STEFFEN B, et al. Lecture notes in computer Science, 2018.
- [14] NATHANI D, CHAUHAN J, SHARMA C, et al. Learning attention-based embeddings for relation prediction in knowledge graphs[C]. Florence; Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019.
- [15] XU H, BAO J, LIU W. Double-branch multi-attention based graph neural network for knowledge graph completion [C]//ROGER A, BOYD-GRABER J, OKAZAKI N. Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, 2023; 15257–15271.
- [16] BROWN T, MANN B, RYDER N, et al. Language models are few-shot learners [C]//LAROCHETTE H, RANZATO M, HADSEIN R. Advances in Neural Information Processing Systems, 2020, 33; 1877–1901.
- [17] PAN S, LUO L, WANG Y, et al. Unifying large language models and knowledge graphs: A roadmap[C]. IEEE. IEEE Transactions on Knowledge and Data Engineering, 2024.
- [18] ZHU Y Q, WANG X H, CHEN J, et al. LLMs for knowledge graph construction and reasoning: Recent capabilities and future opportunities [J]. Computer Science, 2023. DOI: 10.1007/s11280-024-01297-w.
- [19] LIU X, JI K X, FU Y C, et al. P-Tuning v2: Prompt tuning can be comparable to fine-tuning universally across scales and tasks [C]//MURESAN S, NAKOV P, VILLAVI CENCIOA. Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, 2022.
- [20] 李子豪, 冯林, 徐凌泉, 等. 基于少样本学习的知识图谱补全综述[J]. 小型微型计算机系统, 2024, 45(12): 2838–2857.
- LI Z H, FENG L, XU L X, et al. Few-shot knowledge graph completion overview[J]. Journal of Chinese Computer Systems, 2024, 45(12): 2838–2857.

(上接第 173 页)

- [12] WU W H, PAN X Y, KANG Y K, et al. To feel the spatial: Graph neural network-based method for leakage risk assessment in water distribution networks[J]. Water, 2024, 16(14): 2017.
- [13] 王云亮, 乔庆, 韩冰. 数据驱动的供水管网安全评估模型应用实践[J]. 给水排水, 2023, 49(7): 140–147.
- WANG Y L, QIAO Q, HAN B. Application practice of data-driven safety assessment model of water supply pipeline network [J]. Water & Wastewater Engineering, 2023, 49(7): 140–147.
- [14] 李彬彬. 基于机器学习的市政管网运维风险评估[D]. 哈尔滨: 哈尔滨工业大学, 2021.
- LI S S. Risk assessment of municipal pipenetwork operation and maintenance based on machine learning [D]. Harbin: Harbin Institute of Technology, 2021.
- [15] 应路瑶. 基于水力模型的供水管网独立计量分区优化方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2022.
- YING L Y. Optimization study on district metered area by hydraulic model in water distribution system [D]. Harbin: Harbin Institute of Technology, 2022.
- [16] 朱鹏飞. 基于神经网络的供水管网爆管风险评估模型研究[D]. 上海: 上海财经大学, 2023.
- ZHU P F. Risk assessment model for pipe bursts in water supply networks using neural networks [D]. Shanghai: Shanghai University of Finance and Economics, 2023.
- [17] 李振博. 基于改进差分进化的供水管网 DMA 分区研究[D]. 昆明: 昆明理工大学, 2023.
- LI Z B. Research on district metered areas of water distribution network based on improved differential evolution [D]. Kunming: Kunming University of Science and Technology, 2023.
- [18] WANG M, MA Y, WU C, et al. A data-driven online model system for urban water pipeline network based on GIS [C]. Wuhan: International Conference on Electronics Information and Emergency Communication, 2024.
- [19] ABDULSAMAD A A, ABDULRAZZAQ K A. Calibration and analysis of the potable water network in the Al-Yarmouk region employing WaterGEMS and GIS[J]. Journal of the Mechanical Behavior of Materials, 2022, 31(1): 298–305.
- [20] 李正泉, 吴尧祥. 顾及方向遮蔽性的反距离权重插值法[J]. 测绘学报, 2015, 44(1): 91–98.
- LI Z Q, WU Y X. Inverse distance weighted interpolation involving position shading[J]. Acta Geodaetica et Cartographica Sinica, 2015, 44(1): 91–98.