

薛浩, 代艳格. 融合联邦知识图谱隐私计算与大语言模型推理的水质应急处置方法[J]. 净水技术, 2026, 45(3): 52-64.

Xue H, Dai Y G. Water quality emergency disposal integrated with federated knowledge graph privacy computation and large language model inference [J]. Water Purification Technology, 2026, 45(3): 52-64.

## 融合联邦知识图谱隐私计算与大语言模型推理的水质应急处置方法

薛浩<sup>1,2,\*</sup>, 代艳格<sup>2</sup>

(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】水质应急管理是保障饮用水安全的关键环节, 本文旨在解决多业务域协同下的水质应急管理中面临的信息孤岛、隐私泄露风险及智能决策支持不足等问题, 提升应急响应的准确性与安全性。【方法】本文提出联邦知识图谱隐私计算与大模型应急处置方案生成融合方法(Fed-WaterKG-LLM), 该方法通过关系加权的实体聚合实现跨业务域的知识图谱共享, 结合同态加密实现知识图谱嵌入的安全计算与传输, 在保证隐私保护约束的同时, 支持多业务域知识的协同融合。在此基础上, 将全局聚合的知识图谱与本地知识图谱进一步融合, 以提升知识图谱的个性化服务能力, 最后引入联邦知识图谱作为外部知识增强大语言模型的因果推理能力, 为跨域水质应急处置方案生成提供支持。【结果】以水质联邦知识图谱链接预测任务进行验证, Fed-WaterKG-LLM 在平均倒数排名(MRR)、Hits@K 等指标上较独立训练方法提升 2% 以上, 并在隐私保护需求下能够保持模型预测性能; 将所提方法应用在水质应急方案生成任务中, 采用专家评分, 其在水质应急事件溯源准确率和应急方案可行性亦显著优于对比方法。【结论】Fed-WaterKG-LLM 能够兼顾知识共享与隐私保护, 有效提升跨域水质应急事件的智能决策水平。

**关键词** 水质应急处置 大模型 知识图谱 联邦学习 隐私安全

中图分类号: TU991 文献标志码: A 文章编号: 1009-0177(2026)03-0052-13

DOI: 10.15890/j.cnki.jsjs.2026.03.005

## Water Quality Emergency Disposal Integrated with Federated Knowledge Graph Privacy Computation and Large Language Model Inference

Xue Hao<sup>1,2,\*</sup>, Dai Yange<sup>2</sup>

(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] Water quality emergency management is a critical component in ensuring drinking water safety. This paper aims to address challenges in multi-domain collaborative water quality emergency management, including information silos, privacy leakage risks, and insufficient intelligent decision-making support, thereby enhancing the accuracy and security of emergency responses. [Methods] This paper proposed a fused method of federated knowledge graph privacy computing and large language model-based emergency response generation, named Fed-WaterKG-LLM. The method enabled cross-domain knowledge graph sharing through relation-weighted entity aggregation, ensures privacy security via homomorphic encryption. On this basis, the globally aggregated knowledge graph was further integrated with the local knowledge graph to enhance the personalized service capabilities of the knowledge graph. Finally, federated knowledge graphs were introduced as external knowledge to enhance the causal reasoning ability of the large language model, providing support for cross-domain water quality emergency response plan generation. [Results] Evaluated on a water quality federated knowledge graph link prediction task, Fed-WaterKG-LLM achieved an average improvement of over 2% in indices such as mean reciprocal rank (MRR) and Hits@K compared to independent training approaches, while maintaining model

[收稿日期] 2025-10-28

[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, 工程师, 主要从事大数据分析及应用等工作, E-mail: iamxuehao@126.com。

predictive performance under privacy-preserving conditions. When applied to water quality emergency response generation, expert evaluation results showed that the proposed method significantly outperformed baseline method in terms of event root-cause tracing accuracy and the feasibility of generating emergency plans. [ **Conclusion** ] The findings demonstrate that Fed-WaterKG-LLM effectively balances knowledge sharing and privacy protection, enhancing intelligent decision-making for cross-domain water quality emergency events.

**Keywords** water quality emergency disposal large model knowledge graph federated learning privacy safety

保障饮用水安全是全社会高度关注的重大问题<sup>[1-2]</sup>。近期,我国城市自来水安全问题再次凸显,在跨业务域水质应急管理中(如居民自来水出现异味),信息分散、响应迟缓和数据孤岛现象普遍存在。水源地和水厂拥有原水及出厂水的实时监测数据(如COD、氨氮、余氯等);管网公司掌握管网流量、压力及管网节点水质信息;监管部门掌握环境监测数据及历史水质污染事故案例;公众与媒体提供用户反馈、舆情动态等信息。受限于隐私与安全壁垒,这些跨业务域数据难以实现高效整合与共享,直接制约了水质应急事件的溯源、预警与决策效率<sup>[3]</sup>。因此,如何在保障数据安全的前提下实现跨区域水质知识的融合与应用,是当前亟需解决的关键问题。

现有研究已在水质应急管理中引入知识图谱与联邦学习等技术。一些研究<sup>[4]</sup>利用知识图谱中的关系推理实现污染事件溯源,能够表达基本的因果链条,但多聚焦于单一水体或单一业务域场景,无法扩展至多业务域水质应急协同场景。时序知识图谱研究对实体关系的动态演化建模<sup>[5-6]</sup>,但普遍依赖集中式数据处理,难以适应水务行业长期存在的数据孤岛问题。联邦学习方面,多流特征感知模型可缓解非独立同分布数据带来的性能下降<sup>[7]</sup>,多服务器隐私保护架构可以增强模型更新过程的可追责性与安全性<sup>[8]</sup>。此外,将联邦学习与隐私计算结合的技术已在医疗大数据协同分析中应用<sup>[9]</sup>。然而,此类方法缺乏对水务场景多源异构实体——关系结构的统一建模与融合能力。针对这一挑战,已有研究<sup>[10-11]</sup>从异构图谱嵌入、结构一致性建模等角度探索分布式知识融合。例如,共享潜在嵌入矩阵被用于实现跨域图谱的一致性表示<sup>[11]</sup>,但未提供跨域关系重要性建模机制,也未与推理链生成或应急推理模型结合。综上所述,现有研究在关系推理、时序演化建模以及联邦协同等方向均取得了一定进展,但总体仍呈现出分散化特征,缺乏能够同时面向多业

务域异构知识、兼顾隐私保护并支持应急推理的统一方法体系,难以满足复杂水质事件的跨域协同处置需求。

针对上述问题,文章提出了一种融合联邦知识图谱隐私计算与大模型推理的水质应急处置方法(Fed-WaterKG-LLM),该方法以水厂、管网、二次供水等全流程供水为研究对象,面向跨业务域的水质应急协同场景。具体而言,在知识层面,为解决不同业务域实体关联强度不均带来的融合偏差,Fed-WaterKG-LLM通过动态量化实体在各客户端中的关系重要性,实现跨域实体的加权聚合,从而在保持本地差异的同时获得更具代表性的全局表示;在安全层面,考虑到实体嵌入在跨域传输过程中可能存在泄露风险,模型在联邦训练阶段引入同态加密机制,使得知识嵌入能够在密文状态下参与聚合,满足水质数据的高安全性要求;在推理层面,为应急处置提供可解释的分析链条,Fed-WaterKG-LLM利用检索增强生成(RAG)机制<sup>[12]</sup>,将融合后的知识图谱作为大模型上下文知识来源,引导大模型生成因果链与应急方案建议。本文的知识图谱融合未涉及实体消歧过程,仅针对相同实体在不同业务域中的多源关系信息进行加权聚合,突出本研究在相同实体层面的语义融合与权重分配优化上的创新性与实用价值。

## 1 研究方法

### 1.1 Fed-WaterKG-LLM 架构设计

为实现跨业务域水质知识的安全融合与应急处置,本文提出了Fed-WaterKG-LLM方法。该方法通过整合多源知识图谱、保障数据隐私,并结合大模型推理能力,实现应急方案的快速生成。如图1所示,Fed-WaterKG-LLM包含4个核心模块:基于关系权重的实体聚合模型、联邦-本地模型信息融合、同态加密支持的安全计算、水质应急处置方案生成。

在数据流转过程中,水厂、管网和二次供水等客

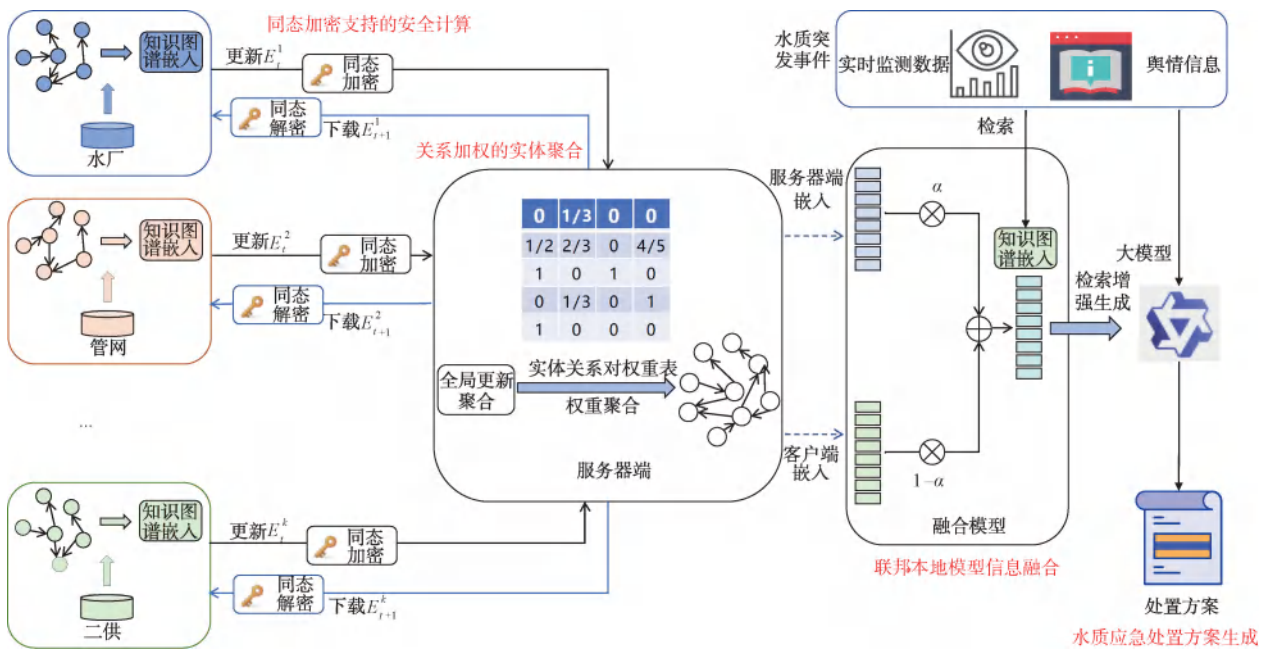


图1 Fed-WaterKG-LLM 架构设计

Fig. 1 Architecture Design of Fed-WaterKG-LLM

户端首先在本地形成结构化知识图谱,这些图谱记录了运行监测参数、设备信息、事件记录及历史处置经验等内容。具体而言,水厂端主要涉及原水与出厂水水质指标(COD、氨氮、浑浊度、pH、余氯等)、工艺运行状态及加药数据;管网端包含流量、压力、节点余氯、阀门状态及回流监测信息;二供端涵盖水箱水位、水质监测值、加压泵站运行及维护记录;此外,还包括历史水质突发事件的应急预案记录和专家建议文本等。为确保隐私与一致性,所有数据在进入模型前均经过脱敏处理,仅保留用于知识表示与建模的核心特征信息。

在模型交互阶段,首先各客户端(如水厂、管网、二次供水等)利用本地知识图谱训练实体嵌入表示,并通过同态加密技术对实体嵌入向量加密后上传至服务器端。接下来服务器端根据关系权重对实体嵌入向量进行聚合处理,形成全局知识表示,再将聚合结果解密后返回给各客户端。最后客户端将全局知识图谱嵌入表示与本地知识图谱嵌入表示进行加权融合,获得可用于下游任务的统一知识图谱表示。该过程实现了从本地训练、加密上传、全局聚合、结果回传的安全数据流转,既保障了数据隐私,又实现了知识共享。

在应用层面,当发生水质突发事件时,系统实时采集各业务域监测数据及用户反馈的舆情等信息,将其映射成语义向量表示,并从融合后的统一知识图谱中检索异常事件相关的子图,作为大语言模型的外部知识。然后以检索增强生成机制<sup>[13]</sup>增强大模型在特定水质事件处置的推理能力,实现应急路径快速推理与处置方案生成,从而提升响应速度与方案的科学性与可行性。

在实体聚合过程中,通过分析重叠实体在各客户端中的关系数量权重进行加权聚合,既保留了各客户端的特征信息,又有效过滤了冗余噪声。联邦-本地融合模型则通过权衡全局模型与本地模型的特征,既保持了全局模式的普适性,又增强了本地知识图谱个性化需求的处理能力。隐私保护方面,采用同态加密技术确保数据传输安全,在保护用户隐私的同时保证了知识图谱嵌入的有效性。基于此,Fed-WaterKG-LLM能够在保障安全的前提下实现跨域知识协同,并通过大模型驱动的智能推理,有效支撑复杂水质应急场景下的科学决策。

## 1.2 关系加权的实体聚合

基于关系权重的实体聚合模块是联邦学习在知识图谱数据隐私安全表示中的关键组件,它主要由

服务器端和客户端 2 个部分构成。服务器端的核心职责是将各客户端上传的实体嵌入表示进行有效聚合,并将聚合后的实体嵌入表示反馈给各客户端。而客户端则需利用本地数据对实体嵌入表示和关系嵌入表示进行持续更新,并将更新后的实体嵌入表示回传至服务器端,以实现知识图谱的动态演化和优化。

在服务端,为了精准地将客户端的本地实体嵌入映射到全局实体嵌入空间,构造出一组变换矩阵,如式(1)。

$$\mathbf{G}_k \in \{0,1\}^{N \times N_k} \quad (1)$$

其中: $\mathbf{G}_k$ ——第  $k$  个客户端的全局实体嵌入矩阵;

$k$ ——客户端的总数;

$N$ ——所有客户端的总实体数;

$N_k$ ——第  $k$  个客户端中的实体数。

若  $(\mathbf{G}_k)_{ij} = 1$  则表示全局实体嵌入表示中的第  $i$  个实体对应客户端中的第  $j$  个实体;若  $(\mathbf{G}_k)_{ij} = 0$  则表示为全局实体嵌入表示中的第  $i$  个实体在客户端中没有对应的第  $j$  个实体。

Fed-WaterKG-LLM 将客户端中实体的权重系数定义为如式(2)。

$$\mathbf{W}_k \in [0,1]^{N \times 1} \quad (2)$$

其中: $\mathbf{W}_k$ ——第  $k$  个客户端的权重系数。

其值为全局实体嵌入表示中的第  $i$  个实体对应客户端中的第  $j$  个实体中该实体的三元组数量占全局该实体三元组数量的比例,这一权重系数的设定充分考虑了不同客户端对同一实体信息贡献的差异性。对于重叠实体,服务端根据其在各客户端中出现的关系数量为其设置权重,若某个客户端对某一重叠实体拥有更多关系信息,则赋予该客户端更高的权重系数,而非简单地进行平均处理。客户端的实体嵌入  $E$  更新如式(3)。

$$E = \sum (\mathbf{G}_k E_k) \times \mathbf{W}_k \quad (3)$$

其中: $E_k$ ——第  $k$  个客户端的实体嵌入。

对于客户端,各个客户端从服务器端接收到实体嵌入表示后,使用本地自己的三元组和关系嵌入表示通过知识图谱嵌入表示方法来更新实体嵌入表示,同时也会更新对应的关系嵌入表示。目前常用

的知识图谱嵌入表示方法如表 1 所示。各个客户端的知识图谱嵌入表示任务采用基于负采样的交叉熵损失函数如式(4)。

$$L = \sum_{(h,r,t) \in b} -\log\{\sigma[f_r(h,r,t) - \alpha]\} - \sum_{i=1}^{k^-} s(h,r,t_i^-) \log\{\sigma[\alpha - f_r(h,r,t_i^-)]\} \quad (4)$$

其中: $L$ ——损失函数;

$b$ ——客户端数据集的一个批次;

$(h,r,t)$ ——知识图谱中存在的三元组;

$(h,r,t_i^-)$ ——知识图谱中不存在的三元组;

$k^-$ ——负样本的数量;

$\alpha$ ——间隔超参数;

$\sigma$ ——sigmoid 函数;

$s(h,r,t^-)$ ——负样本的权重;

$f_r(h,r,t)$ ——分数函数(即知识图谱嵌入表示方法),用来表示知识图谱中三元组存在的可能性。

$s(h,r,t_j^-)$  定义如式(5)。

$$s(h,r,t_j^-) = \frac{\exp[\beta f_r(h,r,t_j^-)]}{\sum_i \exp[\beta f_r(h,r,t_i^-)]} \quad (5)$$

其中: $\beta$ ——温度系数。

常用的知识图谱嵌入表示分数函数如表 1 所示。

表 1 常用的知识图谱分数函数  
Tab. 1 Common Knowledge Graph Scoring Functions

知识图谱嵌入方法	分数函数	向量空间
基于复数旋转的嵌入模型 (RotatE) <sup>[14]</sup>	$\text{Re}[h^T \text{diag}(r) t]$	$h, r, t \in \mathbb{C}^d$
基于平移的知识图谱嵌入模型 (TransE) <sup>[15]</sup>	$-\ h + r - t\ $	$h, r, t \in \mathbb{R}^d$
双线性对角矩阵模型 (DistMult) <sup>[16]</sup>	$-\ h \times r - t\ $	$h, r, t \in \mathbb{C}^d$
基于复数表示的知识图谱嵌入模型 (ComplEx) <sup>[17]</sup>	$h^T \text{diag}(r) t$	$h, r, t \in \mathbb{R}^d$

注: $h, r, t$ ——知识图谱中三元组的向量;Re—复数的实数部分;diag—对角矩阵; $\mathbb{C}$ —复数域; $\mathbb{R}$ —实数域。

### 1.3 联邦-本地模型信息融合

虽然通过联邦学习的方式可以融合各客户端的信息,从而达到全局最佳的效果,但可能忽略了各客户端的局部特性。因此, Fed-WaterKG-LLM 通过将

本地知识图谱嵌入表示和联邦知识图谱嵌入表示进行融合,得到最佳的知识图谱嵌入表示。

对于三元组  $(h, r, t)$ ,  $f_r^l(h, r, t)$  表示本地知识图谱嵌入表示通过分数函数计算得到的向量,  $f_r^s(h, r, t)$  表示联邦知识图谱嵌入表示通过分数函数计算得到的向量,通过将  $f_r^l(h, r, t)$  和  $f_r^s(h, r, t)$  进行融合得到最终的向量。为了便于表示,将  $f_r^l(h, r, t)$  记成  $v_l$ ,  $f_r^s(h, r, t)$  记成  $v_s$ , 使用加权平均的方法进行融合,即将 2 个向量按一定比例加权平均如式(6)。

$$\mathbf{v}_{\text{fused}} = \alpha v_s + (1 - \alpha) v_l \quad (6)$$

其中:  $v_s$  ——  $f_r^s(h, r, t)$ ;

$v_l$  ——  $f_r^l(h, r, t)$ ;

$\mathbf{v}_{\text{fused}}$  —— 联邦-本地模型三元组融合向量。

融合后的向量既包含了全局联邦知识图谱的信息,又包含了本地个性化的知识图谱信息。通过加权平均的方式根据不同的重要性对向量进行加权,使得最终结果能够更好地反映各个向量的特性,权重的分配使得向量加权平均的结果更加可解释。

#### 1.4 同态加密支持的安全计算

由于服务器端在聚合各个客户端的实体嵌入表示后会直接发送回各个客户端,攻击者可以利用接收到的实体嵌入表示进行模型反向工程,试图重建原始数据或识别特定用户的特征,这个过程可能会存在数据隐私泄露的风险。Fed-WaterKG-LLM 为了保护各客户端的数据隐私,在客户端向服务器端传输本地的实体嵌入表示时进行同态加密,在服务器端向客户端传输聚合好的实体嵌入表示时进行同态解密,保护各个客户端数据的隐私。

同态加密是指满足密文同态运算性质的加密算法,即数据经过同态加密之后,对密文进行特定的计算,得到的密文计算结果,再进行对应的同态解密后的明文等同于对明文数据直接进行相同的计算,实现数据的“可算不可见”。

在 Fed-WaterKG-LLM 场景中,同态加密的对象是知识图谱中实体的词嵌入向量。各客户端在本地完成实体嵌入的加权计算后,只需在聚合阶段执行向量间的加法操作,因此对加法同态性有明确需求,而不涉及矩阵乘法等复杂运算。在此背景下,Paillier 部分同态加密能够较好地满足需求,既保证了隐私安全,又避免了全同态加密所带来的高额计

算开销。

首先进行密钥生成,选择 2 个大的素数  $p$  和  $q$ , 其计算如式(7)。

$$\begin{aligned} n &= pq \\ \lambda &= l_{\text{cm}}(p-1, q-1) \end{aligned} \quad (7)$$

其中:  $n$  —— 模数, 2 个素数的乘积;

$l_{\text{cm}}$  —— 最小公倍数。

接下来生成密钥对,选择随机数  $g = n + 1$ , 生成公钥为  $(n, g)$ , 私钥为  $\lambda$ 。生成密钥后,对客户端向服务器端传输本地的实体嵌入表示进行加密,如式(8)。

$$c = Q(m) = g^m \times r^n \times \text{mod}(n^2) \quad (8)$$

其中:  $g$  —— 生成公钥时选择一个随机数;

$r$  —— 加密时选择一个随机数;

$m$  —— 实体嵌入向量中的单个元素;

$Q$  —— 加密函数;

$\text{mod}$  —— 取模运算;

$c$  —— 加密后的密文数据。

对加密后的实体嵌入表示聚合后,通过私钥  $\lambda$  和公钥为  $(n, g)$  进行解密,其公式如式(9)。

$$D(c) = \frac{L[c^\lambda \text{mod}(n^2)]}{\lambda} \times \text{mod}(n) \quad (9)$$

其中:  $D(c)$  —— 对密文数据  $c$  解密的函数。

解密过程中使用的辅助函数如式(10)。

$$L(u) = \frac{u-1}{n} \quad (10)$$

其中:  $L(u)$  —— 解密过程中使用的辅助函数;

$u$  —— 由密文指数运算得到的中间变量。

通过在传输过程中的同态加密以及解密过程,可以在保护各客户端的原始数据隐私的同时,对数据进行聚合运算。

此外,由于各个客户端只负责将实体嵌入表示上传给服务器端,因此各个客户端的原始关系信息以及关系嵌入表示是各自独立的,不会被服务器端和其他客户端知道,保证了各客户端知识图谱中关系的隐私性。并且,各个客户端并没有将原始的三元组进行传输,也直接保证了各个客户端三元组的数据隐私。而对于实体嵌入表示,通过基于同态加密的隐私保护算法对传输过程中的实体表示嵌入进

行加密解密,从而更好地保护各客户端知识图谱中实体的隐私性。

### 1.5 水质应急处置方案生成

在水质应急事件发生后,如何快速、准确地生成可执行的应急处置方案,是协同决策的核心目标。Fed-WaterKG-LLM 综合利用传感器实时数据、文本信息(如监测报告、公众投诉、专家意见)以及多源知识图谱,构建一体化的检索与生成框架。该框架包含了多源数据采集与处理、上下文检索、应急方案生成 3 个核心阶段,通过串联这 3 个阶段, Fed-WaterKG-LLM 能够将实时监测数据与深厚的领域知识相结合,最终输出结构严谨、推理清晰、可操作性强的应急处置方案。

多源数据采集与处理阶段是将异构的多源信息统一表示为模型可处理的格式,系统实时获取各监测点的关键水质参数[如浑浊度( $x_1$ )、pH( $x_2$ )、余氯( $x_3$ )、氨氮( $x_4$ )等],并收集文本类输入  $T$ (事件描述、初步报告、公众反馈),通过特征工程,将这些数据构建为一个统一的查询向量,如式(11)。

$$X = \{x_1, x_2, \dots, x_i, \text{Embedding}(T)\} \quad (11)$$

其中: $i$ ——水质参数的个数;

$T$ ——事件描述、历史案例等文本类输入;

$X$ ——输入信息;

$\text{Embedding}(\cdot)$ ——将文本转化为向量表示。

$X$  包含了当前应急事件的客观状况,为后续的知识检索提供了精准的查询依据。

上下文检索阶段是从联邦-本地融合后的知识图谱中精准定位与当前事件最相关的专业知识,作为后续大语言模型提示工程中的上下文信息。利用知识图谱嵌入向量表示  $E$ ,将  $X$  映射为查询表示  $q$ ,随后在图谱嵌入空间中进行多跳相似检索,筛选出与当前状况高度相关的实体、处置措施及其关联关系,构建一个较为精炼的知识子图  $G_{\text{sub}}$ ,如式(12)。

$$G_{\text{sub}} = \underset{e_i \in G}{\text{argmax}} \text{Sim}(q, e_i) \quad (12)$$

其中: $G_{\text{sub}}$ ——与应急事件高度相关的子图;

$e_i$ ——知识图谱节点的嵌入;

$\text{Sim}(\cdot)$ ——相似度度量函数(如余弦相似度);

$\text{argmax}$ ——选择相似度最高的 Top- $k$  个节点及其关联边,构成知识子图。

此过程确保了最终提供给大语言模型的上下文信息,是高度相关且具备较高的知识密度,为生成可靠处置方案奠定了知识基础。

应急方案生成阶段的核心旨在利用检索到的知识  $G_{\text{sub}}$  与  $X$  进行融合,通过精心设计的提示工程,驱动大语言模型生成最终的应急处置方案  $R$ 。提示工程的构建是确保生成内容准确性、逻辑性和实用性的关键,其核心在于构建一个结构化的提示模板,如下所示。

#### 1) 系统角色定义

你是一名资深的多业务域水质应急处理专家,任务是基于提供的实时数据和上下文信息,生成一份科学、可执行的应急处置报告。

#### 2) 实时事件数据

传感器读数:浑浊度= $x_1$ , pH= $x_2$ , 余氯= $x_3$ , 氨氮= $x_4$ 。

文本描述: $T$ 。

3) 上下文信息:从  $G_{\text{sub}}$  中提取的关键三元组如下。

污染物信息:(污染物 A, 属性, 高毒性);(污染物 A, 中和剂, 药剂 B)等。

处置规程:(规程 C, 适用条件, pH 值>9);(规程 C, 操作步骤, 关闭阀门)等。

历史案例:(案例 D, 主要措施, 关闭闸门 E 与投加活性炭)等。

4) 逻辑推理指令:请严格遵循以下推理步骤进行思考,并最终输出报告。

因果分析:结合实时数据与领域知识,分析导致当前水质异常的最可能原因。

影响评估:预测污染物的扩散趋势和对供水系统的影响范围。

措施制定:根据检索到的处置规程和历史案例,制定具体的工程调度与水处理工艺调整方案。

监测建议:提出后续需要重点监控的指标和频次。

5) 格式化输出要求:报告必须严格按照事件原因初步分析、受影响范围预测、应急处置、后续监测与评估计划进行组织。

该模板系统性地整合了以下 5 个关键要素。  
(1) 系统角色定义:明确指定大语言模型扮演“资深

水质应急处理专家”的角色,限定其输出必须专业、严谨且具备可操作性;(2)实时事件数据:清晰列出从第一阶段获取的传感器读数(如  $x_1$ 、 $x_2$  等)和  $T$ ,为模型提供决策的事实依据;(3)上下文信息:将知识图谱  $G_{\text{sub}}$  提取为结构化的文本,如异常指标、工艺流程、历史案例等关键三元组信息,为模型注入领域知识;(4)逻辑推理指令:通过思维链技术,指令化地要求模型遵循“因果分析、影响评估、措施制定、监测建议”的步骤进行思考,引导其进行结构化、可解释的推理;(5)格式化输出要求:明确规定报告的最终输出格式,要求必须包含“事件原因初步分析”“受影响范围预测”“应急处置”及“后续监测与评估计划”等标准化章节,确保信息的完整性和条理性。

应急处置方案  $R$  生成如式(13)。

$$R = L_{\text{LM}}[\text{Prompt}(G_{\text{sub}}, X)] \quad (13)$$

其中:  $\text{Prompt}(\cdot)$ ——将输入信息填充至提示工程模版;

$L_{\text{LM}}$ ——大语言模型。

该框架通过上述模版设计,确保了生成的应急处置方案既植根于可靠的领域知识,又能针对具体的实时情境进行动态推理,最终输出兼具科学性、准确性和可操作性的高质量报告。

## 2 试验设置

### 2.1 数据采集与处理

在联邦知识图谱研究中,常用的基准数据集为 FB15k-237-Fed3,该数据集通过将 FB15k-237 按照关系划分为多个客户端,用于模拟多方协同但数据不可集中共享的联邦知识图谱场景,为算法性能评测提供了标准化环境。然而,此类通用基准数据集主要基于开放领域知识构建,难以反映水务行业知识共享隐私保护等方面的特定需求。本文基于实验室构建的模拟数据集开展研究,所有数据均为仿真生成,不涉及真实用户隐私及敏感信息。通过模拟跨业务域水质应急事件等场景,构建了包含水源监测、管网运行、二次供水检测、环境监管及舆情动态等多维特征的协同决策研究环境。

本研究采用自主构建的模拟水质知识图谱数据集,通过程序化生成以下关键要素:模拟水体(河流、湖泊、地下水)的水质参数(COD、氨氮、余氯、pH等)、流量及地理位置信息;虚拟管网拓扑结构(管径、材质、连接关系)、压力监测、流量变化及关键节

点水质数据;模拟小区水箱、加压泵站等设施的运行状态、水质变化及维护记录;合成环境监测站数据、历史污染事件案例库及应急响应预案;基于大模型生成技术模拟社交媒体、投诉平台等渠道的舆情文本,包括用户报告的水质异常、异味描述等。

试验设计方面,将模拟知识图谱按业务域分为3个独立子集,每个子集对应一个虚拟客户端的训练集(80%)、验证集(10%)和测试集(10%)。划分时严格保持模拟客户端间的数据隔离性,确保各虚拟客户端仅能访问自身分配到的仿真数据,以此模拟实际场景中的隐私保护需求。通过该实验室环境验证方法在水务应急协同场景的有效性,为实际应用提供理论参考。

### 2.2 模型设置

本文采用实验室自研的大语言模型作为 Fed-WaterKG-LLM 核心推理引擎,该模型的初始权重来自公开可获得的 DeepSeek-R1-Distill-Qwen-32B 模型参数集合,DeepSeek-R1-Distill-Qwen-32B 模型通过对预训练模型 DeepSeek-R1 进行蒸馏而得,具有较高的性能表现,能够在保持较优推理准确率的同时,有效减小模型规模和资源消耗,达到推理效率与性能的良好平衡,具备较高的代表性与权威性。在此基础上,实验室在私有环境中,利用大量水质应急报告、工艺流程以及专家知识文档等专业语料,进行了指令微调(instruction fine-tuning)使得模型能够更好地适应水质应急处理场景,实现行业知识的深度融合与优化。

知识图谱嵌入方法是将知识图谱中的实体和关系映射到低维向量空间的技术,旨在保留图谱中的结构信息和语义关系。这些方法在链接预测、实体分类和关系抽取等下游任务中具有重要应用。为了验证所提方法的鲁棒性,选取了4种常见的知识图谱嵌入方法进行试验评估:RotatE<sup>[14]</sup>、TransE<sup>[15]</sup>、DistMult<sup>[16]</sup>和 ComplEx<sup>[17]</sup>。

试验使用 Adam 优化器,学习率设置为 0.001,批次大小(batch size)设置为 512,负样本数量设置为 256,早停容忍轮数(early\_stop\_patience)设置为 15。所有试验均在配备 8 卡 GeForce RTX 4090 图形处理单元(GPU)的环境中,利用 PyTorch 框架进行。

### 2.3 评价指标

为了验证方法的有效性,在试验中采用知识图

谱融合常用的链接预测任务进行评估,链接预测旨在预测图谱中缺失的实体,即从已有的实体和关系中推断出可能存在的头实体或尾实体。为此,选取了2种链接预测任务核心评估指标:平均倒数排名(MRR)和 Hits@K。

MRR:用于衡量模型在预测缺失链接时的整体表现,计算方法为所有查询的倒数排名的平均值。倒数排名是指正确答案在所有候选答案中的排名位置的倒数,MRR值越高,表明模型预测的准确性越高,能够更快地找到正确的答案。

Hits@K:模型在前K名中预测正确的比例,通常使用K=1、5、10等不同值进行评估。Hits@K反映了模型在有限的候选范围预测正确结果的能力,K值越小,对模型的预测精度要求越高。例如,Hits@1衡量模型在所有候选答案中将正确答案排在第一位的比例,而 Hits@10则衡量正确答案排在

前10位的比例,指标值越高表示模型在K个结果中的准确率越高。

综上所述,本文采用MRR、Hits@1、Hits@5和Hits@10 4种量化指标衡量方法的有效性。

### 3 试验结果与分析

#### 3.1 跨域知识图谱融合结果与分析

试验设定独立训练与本文所提联邦训练方法(Fed-WaterKG)进行对照试验,独立训练将3组数据分别部署在3个独立的客户端上,每个客户端训练自己的知识图谱嵌入模型,并在各自客户端上评估模型效果;联邦训练方法在训练过程中,各客户端的数据保持独立,通过同态加密支持的安全计算、关系加权的实体聚合、联邦-本地模型信息融合进行协作,最终在各个客户端上评估联邦训练后的模型效果,以验证联邦训练方法的性能,试验结果如表2所示。

表2 联邦训练方法在知识图谱链接预测任务的试验结果

Tab. 2 Experimental Results of Federated Training Method in Knowledge Graph Linking Prediction Tasks

客户端	知识图谱嵌入方法	试验设置	MRR	Hits@1	Hits@5	Hits@10
客户端 1	RotatE	独立训练	0.527 9	0.427 9	0.648 9	0.706 3
		Fed-WaterKG	0.559 3	0.446 6	0.697 5	0.760 4
	TransE	独立训练	0.507 6	0.399 7	0.638 6	0.700 8
		Fed-WaterKG	0.551 1	0.443 3	0.685 3	0.745 7
	DistMult	独立训练	0.484 2	0.398 2	0.587 5	0.643 4
		Fed-WaterKG	0.530 1	0.437 1	0.640 0	0.700 2
客户端 2	ComplEx	独立训练	0.485 2	0.396 5	0.590 3	0.652 7
		Fed-WaterKG	0.528 8	0.436 6	0.642 0	0.703 9
	RotatE	独立训练	0.410 3	0.293 2	0.549 6	0.638 8
		Fed-WaterKG	0.431 4	0.297 0	0.600 3	0.688 6
	TransE	独立训练	0.389 5	0.261 3	0.544 3	0.636 2
		Fed-WaterKG	0.420 7	0.282 9	0.590 6	0.684 9
客户端 3	DistMult	独立训练	0.369 6	0.269 6	0.489 1	0.561 6
		Fed-WaterKG	0.411 7	0.301 8	0.542 2	0.624 9
	ComplEx	独立训练	0.370 2	0.267 4	0.492 2	0.578 4
		Fed-WaterKG	0.407 4	0.296 1	0.540 6	0.628 1
	RotatE	独立训练	0.367 4	0.249 7	0.500 5	0.605 0
		Fed-WaterKG	0.389 4	0.258 8	0.545 3	0.650 3
TransE	独立训练	0.337 1	0.203 5	0.493 2	0.599 4	
	Fed-WaterKG	0.360 9	0.219 9	0.530 9	0.638 6	
客户端 3	DistMult	独立训练	0.345 0	0.233 0	0.471 3	0.572 1
		Fed-WaterKG	0.366 8	0.250 5	0.497 4	0.604 3

(续表2)

客户端	知识图谱嵌入方法	试验设置	MRR	Hits@ 1	Hits@ 5	Hits@ 10
平均结果	ComplEx	独立训练	0.328 3	0.217 6	0.450 1	0.555 8
		Fed-WaterKG	0.362 4	0.246 2	0.492 2	0.599 7
	RotatE	独立训练	0.431 9	0.319 7	0.563 9	0.648 4
		Fed-WaterKG	0.456 4	0.329 6	0.612 2	0.698 1
	TransE	独立训练	0.408 4	0.284 6	0.556 5	0.644 0
		Fed-WaterKG	0.441 0	0.311 1	0.600 3	0.688 7
	DistMult	独立训练	0.396 0	0.296 5	0.512 8	0.589 1
		Fed-WaterKG	0.433 0	0.326 2	0.557 4	0.640 9
	ComplEx	独立训练	0.391 4	0.290 4	0.508 3	0.593 5
		Fed-WaterKG	0.429 5	0.322 4	0.555 8	0.641 8

与传统的独立训练方法相比, Fed-WaterKG 在知识图谱嵌入 (包括 RotatE、TransE、DistMult 和 ComplEx) 和评价指标 (涵盖 MRR、Hits@ 1、Hits@ 5 和 Hits@ 10) 方面均展现出显著优势。具体而言, 各项指标较独立训练方法分别平均提升了 3.3%、2.4%、4.6% 和 4.8%。这一优异表现主要源于 2 个关键因素: 首先, 联邦学习实现了对分布式数据的协同利用, 有效解决了单一客户端数据不足的问题; 其次, 通过精心设计的融合机制, 该方法在全局知识共

享与本地特征保留之间达到了理想的平衡。试验证明, Fed-WaterKG 不仅成功克服了传统独立训练中存在的“数据孤岛”问题, 而且在严格保护各客户端数据隐私的前提下, 实现了水质应急管理场景中跨域知识的有效整合与协同优化。

### 3.2 消融试验结果分析

为了验证关系加权的实体聚合与联邦-本地模型信息融合模块在跨域知识图谱融合链接预测任务上的有效性, 文章设计了消融试验进行验证 (表 3)。

表 3 消融试验结果

Tab. 3 Results of Fusion Experiment

知识图谱嵌入方法	试验设置	MRR	Hits@ 1	Hits@ 5	Hits@ 10
RotatE	独立	0.431 9	0.319 7	0.563 9	0.648 4
	关系加权	0.444 1	0.318 4	0.596 0	0.683 8
	联邦-本地融合	0.445 9	0.318 8	0.601 2	0.689 3
	Fed-WaterKG	0.456 4	0.329 6	0.612 2	0.698 1
TransE	独立	0.408 4	0.284 6	0.556 5	0.644 0
	关系加权	0.429 8	0.303 8	0.583 6	0.672 4
	联邦-本地融合	0.430 2	0.302 9	0.591 2	0.675 3
	Fed-WaterKG	0.441 0	0.311 1	0.600 3	0.688 7
DistMult	独立	0.396 0	0.296 5	0.512 8	0.589 1
	关系加权	0.387 6	0.284 8	0.506 4	0.588 4
	联邦-本地融合	0.421 8	0.314 9	0.543 8	0.631 7
	Fed-WaterKG	0.433 0	0.326 2	0.557 4	0.640 9
ComplEx	独立	0.391 4	0.290 4	0.508 3	0.593 5
	关系加权	0.396 4	0.294 7	0.513 7	0.594 9
	联邦-本地融合	0.421 8	0.314 9	0.543 8	0.631 7
	Fed-WaterKG	0.429 5	0.322 4	0.555 8	0.641 8

表 3 展示了不同消融设置下的试验结果。从整体来看, Fed-WaterKG 在 4 种知识图谱嵌入方法均取得了最佳表现, 验证了所提出方法在跨域水质应

急知识图谱融合中的有效性。以 RotatE 为例, Fed-WaterKG 的 MRR 达到 0.456 4, 较独立训练提升约 2.5%, 较关系加权的实体聚合与联邦-本地融合方

法的 MRR 分别提升了 1.2% 和 1.1%, 在 Hits@ 10 指标上达到 0.698 1, 明显优于其他方法。表明通过关系加权的实体聚合与联邦-本地双重建模机制, 可以在保证信息充分融合的同时, 有效减少冗余噪声的干扰。

从不同方法的对比来看, 独立训练由于缺乏全局信息支撑, 整体指标最低; 关系加权方法在 RotatE、TransE 和 ComplEx 下表现出一定提升, 但在 DistMult 中效果不稳定, 说明仅依赖关系权重进行实体聚合可能造成信息冗余; 联邦-本地融合模型在多数情况下优于前两者, 表明在全局与本地特征间进行权衡有助于增强个性化服务能力。然而其提升幅度有限, 未能充分挖掘跨域信息的潜力。相比之下, Fed-WaterKG 通过动态量化实体重要性、全局-本地双重建模实现了在 4 种嵌入方法下的全面超越, 整体结果更为稳定。

### 3.3 隐私保护模型对跨域知识图谱融合效果分析

本研究在模型中引入了基于同态加密的安全计算机制, 尽管隐私保护的引入显著提升了数据安全

性, 但也可能对试验结果和系统性能产生一定影响。为此, 本文选取嵌入方法 TransE, 对隐私保护模型在效果与性能方面的影响进行了评估。

由表 4 可知, 引入同态加密的 Fed-WaterKG 在保障数据隐私的同时, 对试验效果和性能产生了较为轻微的影响。从 MRR 和 Hits@ K 指标来看, Fed-WaterKG 与无隐私保护的模型结果非常接近, 平均 MRR 仅下降 0.22%, Hits@ 1、Hits@ 5 和 Hits@ 10 的降幅也均控制在 0.4% 以内, 说明隐私保护机制并未显著削弱联邦知识图谱表示学习和推理任务上的有效性。在性能方面, 隐私保护模型的推理耗时相较于无隐私保护方案略有增加。平均推理耗时从 0.970 1 s 上升至 0.982 6 s, 增幅不足 2%, 表明同态加密在带来安全性的同时, 其计算时间开销处于可接受范围, 不会对系统整体响应速度造成实质性影响。试验结果表明, 在 Fed-WaterKG 中引入同态加密能够有效提升客户端数据的隐私安全性, 同时保证了模型效果与性能的稳定性, 为跨业务域水质应急知识图谱的构建与应用提供了兼顾安全性与高效性的技术路径。

表 4 隐私保护模型对试验结果和性能的影响

Tab. 4 Impact of the Privacy Protection Model on Experimental Results and Performance

客户端	试验设置	MRR	Hits@ 1	Hits@ 5	Hits@ 10	推理耗时/s
客户端 1	无隐私保护	0.552 7	0.444 2	0.686 7	0.746 3	0.971 2
	Fed-WaterKG	0.551 1	0.443 3	0.685 3	0.745 7	0.987 1
客户端 2	无隐私保护	0.421 5	0.283 7	0.591 7	0.685 7	0.968 3
	Fed-WaterKG	0.420 7	0.282 9	0.590 6	0.684 9	0.979 2
客户端 3	无隐私保护	0.361 7	0.221 3	0.531 6	0.639 4	0.970 8
	Fed-WaterKG	0.360 9	0.219 9	0.530 9	0.638 6	0.981 6
平均结果	无隐私保护	0.443 2	0.314 9	0.601 4	0.689 5	0.970 1
	Fed-WaterKG	0.441 0	0.311 1	0.600 3	0.688 7	0.982 6

### 3.4 Fed-WaterKG-LLM 决策支持结果分析

为进一步验证 Fed-WaterKG-LLM 在跨业务域水质应急场景中的适用性, 本研究设计了利用联邦知识图谱与大模型生成应急处置方案的试验。试验模拟了 5 类典型水质应急事件, 涵盖取水口污染、水厂投药异常、管网回流污染、二供消毒过量以及跨域复合事件。各业务域(如水源、水厂、管网、二次供水)作为独立客户端, 监测指标、公众舆情作为输入, 系统在此基础上通过 Fed-WaterKG-LLM 联合训练知识图谱表示, 调用大模型生成因果链推理与处置方案。

试验分别从事件溯源能力及应对方案生成能力 2 个维度进行评估: 事件溯源能力评估能否准确定位出最先出现异常的业务域及对应的异常指标, 采用自动评估方案; 应对方案生成能力评估能否针对不同业务环节提出具有针对性、可执行性和协调性的措施, 采用专家打分(1~5 分)。专家打分环节邀请了 3 位行业专家参与, 包括 1 位来自省级水务集团的高级工程师(具有 10 年以上供水运行及应急管理经验)、1 位高校环境工程与智能决策方向副教授, 以及 1 位长期参与智慧水务项目的算法专家。评分流程为每位专家独立审阅模型生成的应急处

置方案,结合实际水务应急响应流程,从针对性(是否针对具体问题提出有效措施)、可行性(是否具备工程可实施性)、协调性(是否体现多部门联动逻辑)3个维度进行量化评分,最后取平均值作为最终得分。

试验结果如表5所示,Fed-WaterKG-LLM在事件溯源与方案生成2个方面均表现出显著优势。首先,在业务域识别和异常指标推理分别达到了0.95和0.91的准确率,显著优于单纯依赖大模型或独立训练的知识图谱方法,表明引入关系权重的联邦聚合机制能够更有效地整合跨区域数据特征,避免了信

息融合中的偏置问题,从而提高了业务域识别与异常指标推理的准确性。其次,在应对方案的针对性、可行性和协调性上,Fed-WaterKG-LLM的专家打分均分超过4.5分,尤其在针对性指标上获得了4.8分,说明该方法不仅能够结合不同业务域的特点提出差异化措施,还能在跨域协同的情境下提供较强的可操作性和执行指导。相比之下,独立训练知识图谱方法虽然在推理准确性方面有所提升,但缺乏全局一致性,导致应急方案的协调性不足;单一LLM方法则存在回复过于通用的问题,难以准确适配具体场景。

表5 水质应急事件处置人工评价结果

Tab. 5 Manual Evaluation Results of Emergency Response to Water Quality Incidents

方法	应急事件溯源准确率	异常指标推理准确率	处置方案针对性	处置方案可行性	处置方法协调性
LLM	0.70	0.68	3.4	3.6	3.2
图谱独立训练+LLM	0.82	0.79	3.7	4.0	3.4
Fed-WaterKG-LLM	0.95	0.91	4.8	4.6	4.4

以管网回流污染事件为例,某区域监测数据中,出厂水余氯质量浓度为0.8 mg/L(正常),而管网末梢余氯质量浓度降至0.05 mg/L,舆情平台出现自来水有异味的投诉。Fed-WaterKG-LLM融合水厂、管网公司、监管端与公众端的多源数据后输出的应急处置方案如下。

(1)事件原因初步分析:出厂水余氯质量浓度保持在0.8 mg/L(正常),但管网末梢余氯质量浓度降至0.05 mg/L(异常下降),舆情平台出现“自来水有明显异味”投诉。联邦知识图谱推理链识别到的因果路径为出厂水余氯正常→末梢余氯突降→管网回流水力倒灌→二次供水箱受污染风险。根据跨业务域聚合权重判断,最可能的诱因为该区域GXX节点压力突降导致局部回流。综合判定事件由局部压力异常触发的回流现象所致。

(2)受影响范围预测:基于联邦知识图谱与GNN推理结果推断受影响的区域为高风险区段GXX—GXX(回流水段);潜在扩散方向为GXX→下游居民区;二供水箱风险点为S2水箱存在倒灌污染可能。可能受影响用户数约为2300户(按历史流量与节点分布估算)。综合判断风险主要集中在GXX—GXX区段,并可能向二供系统扩散。

(3)应急处置:①工程调度措施,立即关闭GXX下游阀门,阻断回流路径,对GXX—GXX区段进行分区冲洗( $\geq 1.5$ 倍常规流量),实施局部水力隔离,

恢复该区域正常压力,启动二供S2水箱反压监测,必要时执行强制清洗与临时停用。调度运行人员现场排查压降原因,复核周边施工与异常启停记录;②水处理工艺调整临时提高该片区末梢余氯控制目标至0.4~0.6 mg/L。将监测频次由2 h/次提升至30 min/次。若冲洗后仍存在余氯衰减异常,建议在GXX—GXX增设临时加氯点。加强异常指标(浑浊度、氨氮、UV<sub>254</sub>)的短周期复核。

(4)后续监测与评估计划:①持续监测周期为12~24 h;②重点监控指标为余氯、浑浊度、压力、二供水箱反压值;③建议频次为30 min/次(前6 h),1 h/次(后续);④恢复判定标准为末梢余氯 $\geq 0.2$  mg/L且连续3次浑浊度 $\leq 1$  NTU。

系统在接收到各客户端数据后,首先从融合后的统一知识图谱检索潜在因果链:出厂水余氯正常→管网末梢余氯异常下降→管网回流→二供倒灌风险,并根据联邦聚合权重判断最可能的异常源为管网节点的压力突降。随后,构建结构化的应急处置方案提示模板,大模型严格按照提示模板中的推理指令与格式化输出要求自动生成应急处置方案,Fed-WaterKG-LLM生成的方案不仅给出了清晰的因果链推演过程,更输出了多项可直接指导现场行动的决策措施。例如:立即关闭节点下游阀门并实施分区冲洗;启动二供水箱反压检测,防止进一步回流;临时调高末梢余氯监控频率;通过公众渠道发布

临时告知。

总体来看, Fed-WaterKG-LLM 在跨域水质应急场景下实现了知识共享、安全协同与智能决策的有机结合, 验证了其在提升应急管理效能方面的应用潜力与推广价值。

#### 4 结论

本文针对跨业务域水质应急场景中信息孤岛、隐私保护与智能决策不足等问题, 提出了联邦知识图谱隐私计算与大模型应急处置方案生成融合方法。通过关系加权的实体聚合与联邦-本地双重建模机制, 实现了全局知识与本地特征的有效融合; 结合同态加密技术, 保障了跨域数据协作过程中的隐私安全; 在决策环节, 利用大模型与知识图谱结合的因果推理链生成方法, 有效提升了水质应急事件的溯源准确率与应急处置方案生成能力。试验结果表明, Fed-WaterKG-LLM 在知识共享质量、隐私保护和智能决策三个方面均取得了优异表现, 在知识图谱链接预测任务 MRR、Hits@K 指标上较独立训练提升 2% 以上; 采用隐私保护后, 平均 MRR 仅下降 0.22%, Hits@K 降幅均控制在 0.4% 以内, 表明 Fed-WaterKG-LLM 能够兼顾数据嵌入过程中的安全性与知识表示的性能; 在水质应急处置方案生成任务中实现了更高的应急事件溯源准确率与方案可行性, 展现了良好的实际应用潜力;

未来的研究可进一步探索联邦知识图谱与大模型融合的智能演化机制, 推动模型从集中式协同向自治式协同演进。一方面, 引入自适应联邦优化与跨域知识迁移, 以提升多源场景下的泛化与鲁棒性; 另一方面, 深化知识图谱与大模型的语义对齐, 实现知识驱动的可解释推理与动态更新。

#### 参考文献

[ 1 ] 林思远, 王健, 路佳, 等. 基于供水安全保障的南宁市多水源系统规划与设计[J]. 净水技术, 2025, 44(5): 78-85.  
Lin S Y, Wang J, Lu J, et al. Planning and design of nanning's multi-water sources system based on water supply security[J]. Water Purification Technology, 2025, 44(5): 78-85.

[ 2 ] 徐鹏, 朱俊明, 曾珏. 全局视角下“十五五”水安全保障目标体系研究[J]. 中国水利, 2025(14): 7-12.  
Xu P, Zhu J M, Zeng J. Study on the water security target framework for the 15th Five-Year Plan from a global perspective [J]. China Water Resources, 2025(14): 7-12.

[ 3 ] 构建区域协同应急机制 有效提升整体监测能力——湖北省

开展突发环境事件区域联合应急监测实战演练[J]. 环境与生活, 2021(7): 8-8.

Constructing regional coordinated emergency response mechanism to effectively improve the overall monitoring ability—Hubei Province carried out regional joint emergency monitoring drills for sudden environmental incidents [J]. Green Living, 2021(7): 8-8.

- [ 4 ] 毛文山. 基于知识图谱关系推理的水污染事件追踪溯源方法研究[D]. 兰州: 兰州交通大学, 2020.  
Mao W S. Research on water pollution event tracing method based on knowledge graph relational reasoning [D]. Lanzhou: Lanzhou Jiatong University, 2020.
- [ 5 ] 陆佳民, 张晶, 冯钧, 等. 时序知识图谱构建研究综述[J]. 计算机科学与探索, 2025, 19(2): 295-315.  
Lu J M, Zhang J, Feng J, et al. Survey on construction method of temporal knowledge graph [J]. Journal of Frontiers of Computer Science & Technology, 2025, 19(2): 295-315.
- [ 6 ] 曾朝晖, 杨阳, 陈晓方, 等. 时序知识图谱构建关键技术及研究进展[J]. 控制理论与应用, 2025, 42(5): 865-874.  
Zeng C H, Yang Y, Chen X F, et al. Key technologies and research progress of temporal knowledge graph construction [J]. Control Theory & Applications, 2025, 42(5): 865-874.
- [ 7 ] 黄苛明, 刘苗. 面向个性化联邦学习的多流特征感知网络[J/OL]. 计算机工程, 1-12[2026-01-29]. <https://doi.org/10.19678/j.issn.1000-3428.0252410>.  
Huang K M, Liu M. A Multi-Stream Feature Perception Network for Personalized Federated Learning [J/OL]. Computer Engineering, 1-12[2026-01-29]. <https://doi.org/10.19678/j.issn.1000-3428.0252410>.
- [ 8 ] 郭瑞, 李非凡, 张应辉, 等. 多服务器可追责的隐私保护联邦学习方案[J/OL]. 计算机科学与探索, 1-22[2026-01-29]. <https://link.cnki.net/urlid/11.5602.tp.20250904.1718.002>.  
Guo R, Li F F, Zhang Y H, et al. Multi-Server Accountable Privacy Protection Federated Learning Scheme [J/OL]. Journal of Frontiers of Computer Science and Technology, 1-22[2026-01-29]. <https://link.cnki.net/urlid/11.5602.tp.20250904.1718.002>.
- [ 9 ] 张晋伟, 王维. 基于联邦学习与隐私计算的医疗大数据共享平台构建研究[C]//天津市电子学会, 天津市仪器仪表学会. 第三十九届中国(天津)2025'IT、网络、信息技术、电子、仪器仪表创新学术会议论文集. 2025: 46-49.  
Zhang J W, Wang W. Research on the construction of a medical big data sharing platform based on federated learning and privacy computing [C]// Tianjin Electronic Society, China Instrument and Control Society. Proceedings of the 39th China (Tianjin) 2025'IT, Network, Information Technology, Electronics, and Instrumentation Innovation Academic Conference. 2025: 46-49.
- [ 10 ] 李春娟. 基于联邦学习的异构知识图谱嵌入关键技术研究 [D]. 长春: 长春工业大学, 2025.

- Li C J. Research on key technologies of heterogeneous knowledge graph embedding based on federated learning[D]. Changchun: Changchun University of Technology, 2025.
- [11] 王依科, 张东月. 基于共享潜在嵌入矩阵的联邦知识图谱嵌入[J]. 计算机技术与发展, 2025, 35(11): 154-160.  
Wang Y K, Zhang D Y. Federated knowledge graph embedding based on latent embedding matrix sharing [J]. Computer Technology and Development, 2025, 35(11): 154-160.
- [12] Siriwardhana S, Weerasekera R, Wen E, et al. Improving the domain adaptation of retrieval augmented generation (RAG) models for open domain question answering[J]. Transactions of the Association for Computational Linguistics, 2023, 11: 1-17. DOI: 10.1162/tacl\_a\_00530.
- [13] Edge D, Trinh H, Cheng N, et al. From local to global: A graph RAG approach to query-focused summarization [PP/OL]. V2. arXiv (2025-02-19) [2025-10-28]. <https://doi.org/10.48550/arXiv.2404.16130>.
- [14] Sun Z Q, Deng Z H, Nie J Y, et al. RotatE: Knowledge graph embedding by relational rotation in complex space [PP/OL]. V1. arXiv (2019-02-26) [2025-10-28]. <https://doi.org/10.48550/arXiv.1902.10197>.
- [15] Bordes A, Usunier N, Garcia-Duran A, et al. Translating embeddings for modeling multi-relational data: Proceedings of the 27th International Conference on Neural Information Processing Systems, Nevada, December 5-10, 2013 [C]. United States: Curran Associates Inc., 2013.
- [16] Yang B S, Yih W T, He X D, et al. Embedding entities and relations for learning and inference in knowledge bases [PP/OL]. V4. arXiv (2015-08-29) [2025-10-28]. <https://doi.org/10.48550/arXiv.1412.6575>.
- [17] Trouillon T, Welbl J, Riedel S, et al. Complex embeddings for simple link prediction: Proceedings of the 33th International Conference on Machine Learning, New York, June 19-24, 2016 [C]. United States: PMLR, 2016: 2071-2080.

## (上接第 37 页)

- [19] Taoufik N, Boumya W, Achak M, et al. The state of art on the prediction of efficiency and modeling of the processes of pollutants removal based on machine learning [J]. Science of the Total Environment, 2022, 807: 150554. DOI: 10.1016/j.scitotenv.2021.150554.
- [20] Zounemat-Kermani M, Matta E, Cominola A, et al. Neurocomputing in surface water hydrology and hydraulics: A review of two decades retrospective, current status and future prospects [J]. Journal of Hydrology, 2020, 588: 125085. DOI: 10.1016/j.jhydrol.2020.125085.
- [21] Smirnov D A, Mokhov I I. From Granger causality to long-term causality: Application to climatic data [J]. Physical Review E, 2009, 80: 016208. DOI: 10.1103/PhysRevE.80.016208.
- [22] Kodra E, Chatterjee S, Ganguly A R. Exploring Granger causality between global average observed time series of carbon dioxide and temperature [J]. Theoretical and Applied Climatology, 2011, 104(3/4): 325-335.
- [23] McGraw M C, Barnes E A. Memory matters: A case for granger causality in climate variability studies [J]. Journal of Climate, 2018, 31(8): 3289-3300.
- [24] 朱得康. 新型格兰杰因果检验法的研究及其应用 [D]. 青岛: 山东科技大学, 2020.  
Zhu D K. Research and application of the new granger causality test method [D]. Qingdao: Shandong University of Science and Technology, 2020.
- [25] Tiwari A K, Adewuyi A O, Adeleke M A, et al. A time-varying Granger causality analysis between water stock and green stocks using novel approaches [J]. Energy Economics, 2023, 126: 107010. DOI: 10.1016/j.eneco.2023.107010.
- [26] Liu P, Wang J, Sangaiah A K, et al. Analysis and prediction of water quality using LSTM deep neural networks in IoT environment [J]. Sustainability, 2019, 11(7): 2058. DOI: 10.3390/su11072058.
- [27] Ali Shah S, Ai S T, Rack W. Interconnections, trend analysis and forecasting of water-air temperature with water level dynamics in Blue Moon Lake Valley: A statistical and machine learning approach [J]. Journal of Environmental Management, 2025, 379: 124829. DOI: 10.1016/j.jenvman.2025.124829.
- [28] Tiwari A K, Adewuyi A O, Adeleke M A, et al. A time-varying Granger causality analysis between water stock and green stocks using novel approaches [J]. Energy Economics, 2023, 126: 107010. DOI: 10.1016/j.eneco.2023.107010.
- [29] 吕智慧. 格兰杰因果关系的检验方法及其应用 [D]. 长春: 东北师范大学, 2020.  
Lü Z H. Granger causality tests and their applications [D]. Changchun: Northeast Normal University, 2020.