

基于 GraphRAG 的 网络安全知识图谱构建方法

Construction Method of Network Security Knowledge Graph Based on GraphRAG

屠梓浩, 龚 健, 刘 磊, 马先龙, 周婉君, 王 超 (中国联通网络运营事业部, 北京 100037)

Tu Zihao, Gong Jian, Liu Lei, Ma Xianlong, Zhou Wanjuan, Wang Chao (China Unicom Network Operations Division, Beijing 100037, China)

摘 要:

针对通信网络架构复杂化引发的网络安全系统性挑战,提出了一种网络安全知识图谱构建方法。依托 GraphRAG 技术框架与大语言模型协同,以图模式 Schema 驱动对网络安全文档数据进行预处理、文档分块实现知识抽取,并通过结构和语义双重感知技术构建出结构化强、关系精确可溯的网络安全知识图谱。最后采用多粒度混合检索、加权融合重排等多种调优技术,实现知识的精准检索。评测结果表明,该方式构建的网络安全知识图谱更适配通信领域复杂的场景特征,为该领域实现高性能知识检索和智能问答服务提供了适合的技术与实践路径。

关键词:

通信网络安全;检索增强生成;GraphRAG;大语言模型;知识图谱;智能问答

doi:10.12045/j.issn.1007-3043.2026.09.009

文章编号:1007-3043(2026)09-0047-07

中图分类号:TN918

文献标识码:A

开放科学(资源服务)标识码(OSID):



Abstract:

To address the systemic network security challenges arising from the complex communication network architecture, a network security knowledge graph construction method is proposed. By integrating the GraphRAG framework with large language models, it achieves knowledge extraction through graph schema-driven preprocessing of document data and document chunking. A highly structured knowledge graph with precise and traceable relationships is constructed via dual structural and semantic perception. Finally, accurate knowledge retrieval is realized using multi-granularity hybrid retrieval, weighted fusion reranking, and various optimization techniques. Evaluations verify that the constructed network security knowledge graph better adapts to the complex scenario characteristics of the communication field, and it provides appropriate technical and practical approaches for high-performance knowledge retrieval and intelligent Q&A services in this field.

Keywords:

Communication network security; Retrieval-augmented generation; GraphRAG; LLM; Knowledge graph; Intelligent Q&A

引用格式:屠梓浩,龚健,刘磊,等. 基于 GraphRAG 的网络安全知识图谱构建方法[J]. 邮电设计技术,2026(9):47-53.

0 引言

随着数字化浪潮的持续推进,电信网络已由相对封闭的垂直架构,演进为多层次、多技术域融合的复杂系统^[1]。5G、云网融合等新技术在显著提升网络灵活性与业务创新能力的同时,也进一步加剧了网络架构的复杂性,使网络安全面临严峻的系统性挑战^[2]。为此,亟需构建能够深度映射网络复杂性的网络安全

知识图谱。该类“安全地图”可将传统点状告警分析升级为面向场景的综合研判,为自动化响应提供推理支撑,并将跨域安全经验沉淀为可计算知识体系,推动安全运营由被动防御向主动预测转型,从而为复杂网络环境提供关键性的认知基础。然而,当前相关研究还处于初期阶段,亟需发展具备场景适配性高与优良性能的网络安全图谱构建技术,推动行业标准化^[3]。

检索增强生成(RAG)技术在知识图谱构建与智能问答场景中,一定程度上缓解了大模型语言(LLM)的幻觉问题,但仍存在以下痛点:其一,知识以零散文

收稿日期:2026-07-21

本块形式存储,缺乏明确实体关系关联,导致信息链路不可追溯;其二,依托向量空间的单跳相似性检索,缺乏多跳推理能力,难以应对复杂问题,检索精度受限;其三,对网络安全领域多层次、异构化场景的适配性不足,易出现核心知识漏检与无关信息召回;其四,文本分块语义碎片化,难以满足该领域对知识准确性与逻辑严谨性的高要求^[4]。

针对上述不足,提出了一种 Schema 先行的 GraphRAG^[5]构建方法。该方法以领域知识作为前置契约,指导图结构的构建与信息组织;设计兼顾结构特征与语义信息的双重感知社区检测算法,形成四层知识树结构;进一步构建支持多粒度可追溯的混合检索机制,并结合加权融合与重排序策略,实现高精度查询。在网络安全文档上的实测结果表明,该方法能够更有效地捕捉复杂语义关联与上下文依赖关系,显著提升问答与推理的准确性、可解释性与逻辑一致性,从而有效缓解传统 RAG 方法的痛点问题,更适用于通信领域中规则严谨与结构复杂的推理场景。

1 网络安全知识图谱构建过程

1.1 文档预处理

网络安全知识图谱构建的数据主要来自于通信行业相关企业的非公开结构化和非结构化文档。所涉及文档主要分为五大类:应急预案、运行维护规程、技术方案、行业规范及管理制度。

文档预处理质量在很大程度上决定了后续实体关系抽取^[6]及知识图谱构建的整体效果。因此,研究首先将各类原始文档统一转换为 Markdown 格式。该格式具有轻量化、可读性强及良好系统兼容性等优势;可将文档中的表格信息统一转换为 HTML <table> 结构,通过工具序列化生成含 <tr>、<td> 标签的标准化代码,并支持合并单元格等复杂结构,为后续信息的精准解读奠定基础;对于结构松散文档,则通过优化子标题层级、酌情补充目录结构等方式,为基于章节的文本切分提供清晰且稳定的边界。

1.2 知识抽取

知识抽取是图检索增强生成框架中的核心,本环节采用 Schema 驱动抽取方法,并利用 LLM 从 Markdown 文本中提取结构化语义知识,为图谱构建提供“社区—关键字—关系—属性”四层知识树,从而保障输出结果的准确性与一致性。知识抽取的核心流程如图 1 所示。该流程最终形成可精准建模、推理可溯

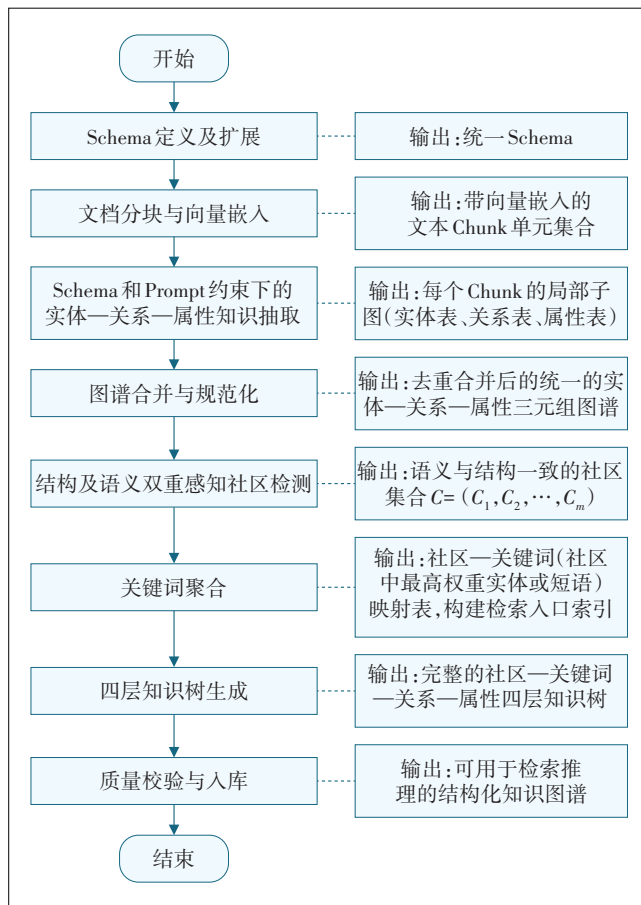


图 1 知识抽取流程

的知识图谱。其所使用的基础配置选型如下:大模型采用 Qwen3-235B-A22B,本地图存储采用 JSON 与 Neo4j5.23 组合,向量检索使用 FAISS,向量模型选用智源研究院(BAAI)开源的 bge-m3。

1.2.1 Schema 定义与扩展

本环节制定知识语法规则,用于约束抽取范围并规范知识结构。Schema 具体设计流程如下:首先界定网络安全领域核心知识范畴,以三元组为核心,定义实体类型、关系的定义域与值域、属性的类型及约束,并为各类型添加唯一 ID 等信息,生成初始 Schema 结构;其次从五大类文档中抽取少量高质量样本,基于结构定义,基于 LLM 进行语义分析,形成种子 Schema 集合;再通过人工审核,合并同义类型、补充缺失类型、删除冗余约束;并统计新增文档未匹配 Schema 的高频候选类型,再次审核合并;最后 LLM 检查并消解冲突以保障一致性,最终形成包含 31 类实体、41 类关系、66 类属性的网络安全领域专业 Schema。网络安全领域图模式 Schema 结构设计如表 1 所示。

表1 网络安全领域图模式 Schema 结构设计

实体类型	厂商、规范、安全威胁、安全策略、通信技术、系统、人物、命令、处理方案、告警、网络类型、仪器仪表设备、组织部门、维护项目、通信网络业务、功能等
关系类型	属于、关联、发布、备份、存在、面临、符合、网络安全、包含、容灾方案、接收、发送、维护子项目、设备故障、故障管理、质量管理要求、故障处理等
属性类型	设备流控影响、设备流控取值说明、设备流控配置指导、设备流控原理、变更操作管理审批、故障历时管理、故障处理原则、网络优化、设备维护技术指标等

1.2.2 文档分块与向量嵌入

本环节旨在将文档切分为语义完整的 Chunk、并为各 Chunk 生成向量嵌入 (Vector Embedding)^[7], 用于后续相似性检索与社区聚类, 最终输出带嵌入向量的 Chunk 集合。为克服传统静态分块忽略章节语义, 易产生碎片的问题, 本研究结合 Markdown 结构与主题进行切分: 以#号为分隔符确保主题明确, 若相邻块大小之和未超 chunk_size 则合并以保连贯; 同时将 HTML 表格标签 (<table>...</table>) 设为不可分割单元, 兼顾文档结构与表格完整性。分块后得到语义完整的 Chunk 及来源 ID、位置偏移等元数据。采用 bge-m3 向量模型生成固定维度稠密向量以支撑相似性计算^[8]。最终将嵌入向量与元数据绑定为键值对入库, 便于检索时同步召回 Chunk。

1.2.3 图模式约束下的实体—关系—属性知识抽取

本环节按照 Schema 约束, 根据预设的提示词引导 LLM 从每个 Chunk 中抽取实体、关系、属性, 设计充分考虑了通信专业特性及输出规范 (见表 2)。

在知识抽取过程中, LLM 的参数配置对图谱的准确率和稳定性具有重要作用。配置如下: Temperature 设为 0.1 以降低随机性, 确保模型倾向选择概率最高的词; do_sample 设为 False, 禁用随机采样, 与低温设置协同保证输出一致性; Top-K 降低为 10, Top-P 降低为 0.5, 形成双重质量门禁, 共同限制模型仅在高质量

表2 知识抽取提示词 Prompt 输出规范

规范维度	规范要求
角色定义	网络通信领域的专业信息提取者
实体类型定义	定义了网元、网络类型、故障等实体类型的具体标准
名称一致性	定义严格的输出顺序格式要求, 强调名称一致性
实体处理规则	明确增加了实体提取与处理的详细规范
表格处理原则	定义了针对表格数据的专项处理规则
推理过程要求	推理过程随机性要求直接输出结果, 严格禁止任何形式的假设、推理与思维链
执行流程	按照实体抽取顺序规范执行
示例丰富度	3 个网络通信运维实际案例

候选词中进行选择, 在兼顾平衡稳定性与准确性的同时保留少量多样性。输出结果为每个 Chunk 的局部子图。

图谱合并与规范化环节, 是对知识抽取环节所取得的各 Chunk 三元组列表进行数据清洗, 主要包含实体合并、关系规范化与属性校验。最终输出全局统一的三元组图谱 $G_{\text{triple}} = (E, R, D)$, 其中 E 为实体集合, R 为关系集合, D 为属性集合。

1.2.4 结构及语义双重感知社区检测

双重感知社区检测是知识抽取的核心聚合能力, 它通过量化实体的拓扑结构与语义特征双维度, 解决传统方法重结构、轻语义问题, 最终输出支撑四层知识树的高质量社区集合 $C = \{C_1, C_2, \dots, C_m\}$ 。其中, 拓扑连通性采用 Jaccard 相似度衡量实体与社区结构的重叠度, 语义相似性则通过余弦相似度衡量实体嵌入与社区质心向量匹配度, 两者加权融合形成双重感知评分^[9]。具体流程如下。

a) 实体嵌入向量生成: 输入全局三元组图谱 $G_{\text{triple}} = (E, R, D)$ 的实体集合 E , 对每个实体 e , 首先为其 1 跳邻居三元组 (中心实体+关系+邻居实体) 生成基础嵌入向量 (e_0, r_i, e_i) , 然后融合为一个表征向量, 再通过均值或注意力等聚合策略, 得到融合结构与语义的嵌入向量 e_{emb} , 最后经归一化消除模长差异, 输出与 bge-m3 模型维度一致的归一化嵌入向量^[10]。

b) 聚类初始化: 用 K-Means 对实体嵌入进行聚类, 得到初始社区 $C_1(0) \dots C_k(0)$ ^[11]。参数设置为最小社区数 $\beta=10$ (防过碎)、最大社区数 $\eta=200$ (防过粗), 采用欧氏距离保证嵌入空间中的聚类稳定性。

c) 迭代融合: 通过双重感知评分计算、社区中心选择、社区合并判断 3 个步骤, 计算实体与社区间亲和力, 融合拓扑连通性, 迭代聚类至社区收敛。双重感知加权融合评分函数如式 (1) 所示:

$$\varphi(e_i, C_m^{(t)}) = \omega_1 \times J(e_i, C_m^{(t)}) + \omega_2 \times \text{Cos}(e_i_{\text{emb}}, C_{m_cent_emb}) \quad (1)$$

其中 $\omega_1 + \omega_2 = 1$ (默认 $\omega_1 = 0.4, \omega_2 = 0.6$); e_i 为待评估实体 (如“AMF 网元”); $C_m^{(t)}$ 为第 t 轮迭代中的第 m 个社区 (如“IMS 核心网社区”), $C_{m_cent_emb}$ 为社区 C_m 的质心向量, 即社区内所有实体嵌入向量按权重 (抽取置信度、关系强度) 计算的加权平均向量, 是社区在向量空间中的抽象语义参考点; $J(e, C)$ 为实体 e 和社区 C 之间的拓扑连通性的 Jaccard 相似度, $\text{Cos}(e_{\text{emb}}, C_{cent_emb})$ 为实体嵌入与社区质心向量之间的语义

余弦相似度。该函数输出每个实体与对应社区的亲和度得分,得分取值范围为 $[0, 1]$,得分越高归属感越强。Jaccard 相似度用于衡量 2 个集合的相似程度。提取实体 e 关系集合 $R_e = \{r_1, r_2, \dots, r_k\}$ 与社区 C 所有成员实体的关系并集 $R_C = \{r_1, r_2, \dots, r_m\}$, 则 e 和 C 的 Jaccard 相似度计算公式如式(2)所示:

$$J(e, C) = |R_e \cap R_C| / |R_e \cup R_C| \quad (2)$$

结果的取值范围为 $[0, 1]$,值越大表示 e 和 C 间的结构亲和度越高。

d) 社区中心选择:为每个社区选择双重感知评分最高实体作为 e_{center} ,代表该社区核心语义与结构,并作为社区合并基准。

e) 社区合并判断:在迭代过程中,计算两社区 $e_{center1}$ 与 $e_{center2}$ 的评分差值,若是差值小于阈值 τ (默认为 0.05) 则合并两社区;当某一轮迭代中不再发生合并,或迭代次数达到最大次数(默认 50) 时,则终止迭代,完成聚类收敛。

f) 社区摘要生成:用 LLM 生成各社区名与描述,输出包含名称、描述、成员列表的社区集合 $C^{[12]}$ 。以社区中心为语义锚点,提取 1~2 跳核心三元组、关键属性、实体结构重叠度等结构素材,同时提取社区质心向量的核心语义主题、实体与质心余弦相似度排序、高频关键词等语义素材。LLM 根据 Schema 和 Prompt 生成粗摘要,再经质心向量校验语义一致性后,将其压缩为符合 Token 限制的标准摘要,与社区节点绑定,完成四层知识树中社区层与关键词层的关联映射。

1.2.5 关键词聚合

本环节任务是构建“社区—关键词”映射表,形成可追溯知识链路,以支撑粗粒度快速检索推理。基于双感知社区检测所输出的稳定社区,包括社区 ID、质心向量、中心实体、成员实体集合、实体关系三元组及属性,作为关键词聚合的基础。从结构与语义双维度构建候选关键词池,结构候选方面,提取中心实体 1~2 跳核心关联实体及社区高频实体;语义候选则根据实体与社区质心向量的余弦相似度排序选取 Top-N 实体,并结合社区摘要中主题词进行补充。对候选池去重后,形成社区候选关键词总池。

随后,根据业务场景调整各项指标权重,按式(3)计算各候选关键词综合权重:

$$\text{Weight}(\text{key}) = \alpha \times \text{Cos}(\text{key_emb}, C_{\text{cent_emb}}) + \beta \times \text{Freq}(\text{key}) + \gamma \times \text{Rel_Strength}(\text{key}) \quad (3)$$

其中, $\text{Cos}(\text{key_emb}, C_{\text{cent_emb}})$ 表示关键词与社

区质心向量的余弦相似度, $\text{Freq}(\text{key})$ 为关键词在社区三元组中的出现频次, $\text{Rel_Strength}(\text{key})$ 为关键词关联关系的平均强度评分, $\alpha + \beta + \gamma = 1$ 。权重计算完成后,对关键词的嵌入向量进行 K-Means 聚类(例如聚类数设为 10),同一聚类内保留权重最高的 1~2 个关键词以避免重复;并过滤掉低权重或与社区主题无关的关键词。最终输出标准化关键词集合,包含关键词 ID、文本、综合权重、所属社区 ID、对应实体 ID 及语义标签,并为每个关键词生成嵌入向量,从而构建语义索引作为检索入口。

1.2.6 四层知识树生成

本环节将关键词集合作为知识树的关键词层,与社区层(社区 ID)、关系层和属性层上下关联,构建“社区—关键词—关系—属性”四层关联结构知识树,实现知识分层组织与双向可追溯,从而支撑高效检索与多跳推理。构建过程的输入包括双感知稳定社区数据、社区摘要、标准化关键词集、高置信度实体三元组及属性数据。各层核心功能与构建逻辑如下。

a) 社区层:作为语义锚点,支撑粗粒度检索。为各稳定社区创建节点,关联其社区 ID、摘要、中心实体、质心向量及边界信息;基于社区质心向量相似度保留 Top-N 相似社区并建立关联边;利用 LLM 生成语义标签,作为顶层入口。

b) 关键词层:连接社区与核心知识,提升检索效率。将关键词集与社区层节点绑定,关联社区 ID、实体 ID 及权重;优先展示高权重关键词,构建“关键词—社区”映射索引,保障双向快速定位。

c) 关系层:存储核心关系,支撑细粒度推理。提取关键词对应实体的 1~2 跳核心三元组,与关键词层关联,实现关键词与核心知识节点间的多对多映射。

d) 属性层:完善知识维度。基于实体属性抽取结果构建节点,并建立与对应实体的关联边,提供实体细节信息。

完成上述建构后,进行质量与一致性校验、入库及索引构建,至此网络安全知识图谱构建完成。

2 网络安全知识图谱知识检索

2.1 知识检索逻辑与流程

知识检索是 Schema 引导的 LLM 检索过程,本研究设计的核心流程为查询解析、多粒度分层检索、结果融合重排、推理验证、溯源输出,依托四层知识树实现高效和精准检索。流程如图 2 所示。

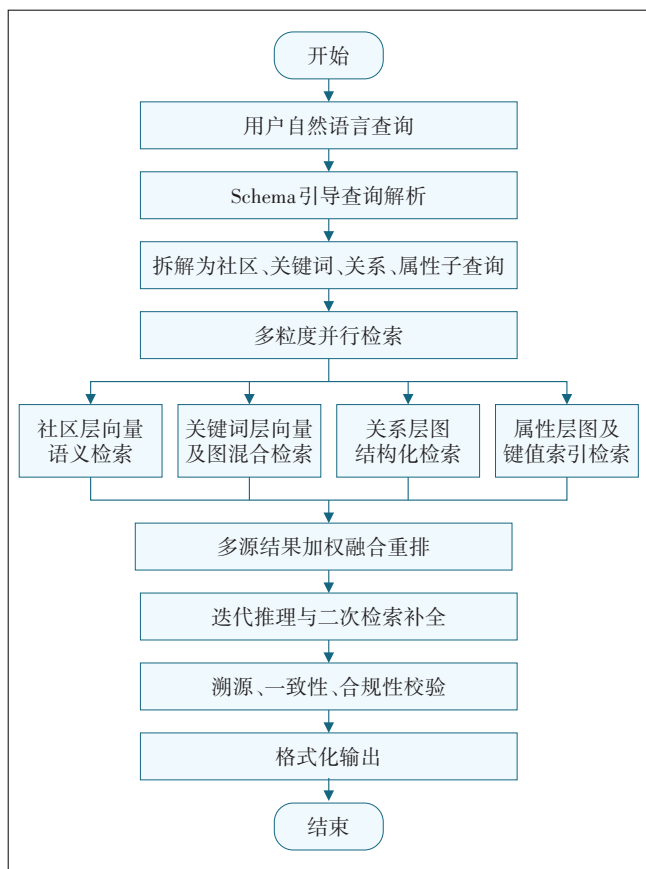


图2 知识检索流程

2.2 检索步骤

2.2.1 查询接收与 Schema 引导解析

本步骤接收用户自然语言查询(如5G基站的传输链路告警关联哪些KPI指标?)。首先基于Schema识别查询中的核心实体(“5G基站”“传输链路”“KPI指标”)、关系类型(如“告警关联”)及属性需求(如KPI指标详情)。随后将复杂查询拆解为并行子查询:社区级定位含目标实体的关联社区,关键词级匹配关系路径(如“告警关联”),属性级提取KPI指标详情;最后为各子查询生成对应的检索语句(向量检索或图查询Cypher语句)。

2.2.2 多粒度并行检索

本步骤分发子查询并行执行4类检索,实现“粗筛+精排”协同机制。

a) 社区层:采用向量语义检索,根据用户查询生成查询向量,与社区质心向量余弦匹配,召回Top-N相似社区(含社区ID、相似度得分、摘要),输出候选列表。

b) 关键词层:采用混合检索,结合查询向量与关键词嵌入向量的匹配及文本比对,输出候选关键词列

表(含关键词ID、权重、关联社区ID)。

c) 关系层:采用图数据库结构化检索,遍历候选社区内的实体关系路径,筛选出高置信度的核心关系三元组。

d) 属性层:采用图结构与键值索引检索,按实体ID提取属性,输出属性列表。

2.2.3 多源融合与重排

本步骤采用加权融合算法,权重设置与业务场景及四层知识树的检索优先级保持一致。融合规则如式(4)所示:

$$\text{Score}(\text{result}) = \alpha \times \text{社区相似度} + \beta \times \text{关键词权重} + \gamma \times \text{关系置信度} + \delta \times \text{属性置信度} \quad (4)$$

其中, $\alpha + \beta + \gamma + \delta = 1$ (示例取值: $\alpha = 0.2, \beta = 0.3, \gamma = 0.4, \delta = 0.1$)。然后根据综合得分对检索结果进行逻辑排序,过滤重复、低得分结果,保留Top-N(如Top-10)候选结果,对于存在语义冲突的结果,保留高置信度且经过多来源验证的结果。

2.2.4 迭代推理及二次检索补全

本步骤基于融合结果执行多跳推理,补全推理链条(如补全“5G基站A到传输链路B到告警C到KPI指标D”的路径);对于缺失的信息,触发二次检索从属性层进行补全,经多轮迭代验证后,调用LLM整合推理链条,生成结构化回答草稿。在输出环节,依次进行溯源验证、一致性校验和合规性检查,最终生成结构化回答,内容包含核心结论、推理路径、关联知识节点及来源链接等。

3 评测

3.1 网络安全知识图谱评测方法

本环节从知识抽取和知识检索2个维度评测网络安全图谱构建效果,以保障“构建—评估—优化”成效。知识抽取通过实体抽取准确率、召回率、F1值评测数据准确性;知识检索通过QA问答结果打分测评准确性。

3.1.1 数据准备

从预处理环节五大类文档中选取测试样本。人工整理出标准节点清单(247个)及QA测试集(317条,含全局类、事件类、比较类、多跳推理4类问题集)。

3.1.2 调优方法

3.1.2.1 知识图谱构建阶段的调优方法

a) Schema精细化分层设计与动态约束优化:明确实体与关系定义域及值域,减少跨层级的无效关联;

提升 Schema 扩展的置信度阈值,避免引入噪声;添加行业强规则约束,在抽取阶段过滤违规三元组,降低 FP(混淆矩阵,假正例)。

b) 文档分块大小优化:通过合理划分文本粒度,平衡语义完整性与检索精准性,语义完整优先,使粒度适配检索目标,避免简单按固定 Token 长度切割文档。

c) 社区检测与知识树构建优化:提升社区知识同质性以及检索准确率和召回率。根据场景调整结构与语义权重(即调整式 1 的参数)。缩小最小社区实体数,保留长尾知识社区;限制最大社区数,避免社区过粗检索精度下降;将“社区—关键词—三元组”的关联边权重与三元组置信度挂钩,避免低置信度三元组进入核心检索链路。

3.1.2.2 知识检索策略调优方法

a) 多粒度检索融合策略优化:平衡粗筛与精排环节。向量检索双阈值动态调整:设置余弦相似度双阈值,以低阈值召回足量候选结果保障召回率,再通过较高阈值过滤低相关结果以提升准确率。向量检索与图检索协同:先执行向量检索,召回 Top-N 结果(如 Top-10)完成粗筛;再利用图检索对召回结果做多跳路径遍历,实现精排。关键词权重分配:依据式(4)提升核心实体对应关键词权重,优先检索核心知识。

b) 查询解析与子查询优化:提升检索效率与精准性。基于 Schema 引导查询拆解,绑定实体与关系类型;对模糊查询、生成子查询分别进行检索。子查询按“核心实体>关系>属性”的优先级顺序执行,权重高的任务优先处理。

c) 结果融合与重排优化:提升结果质量与完整性。多源结果加权融合:调整式(4)融合权重,使结果偏向目标指标。冲突结果消解:按“高置信度>多源验证>权威对齐”的优先级排序,减少误检(FP)。迭代检索补全:对召回结果中的缺失信息,触发二次检索进行补全。

3.2 评测结果

3.2.1 图谱构建效果及调优评测

采用 LLM 相似度评分的方式,将知识抽取所召回节点类型及内容与参考答案进行评分。评分规则如下:满分 10 分为精准匹配,即所有参考答案节点类型和内容均准确召回;0 分为节点类型和内容完全未命中或结果为空;每档分差为 1 分,按照命中节点类型和内容的比例进行打分。优化评测过程如图 3 所示。

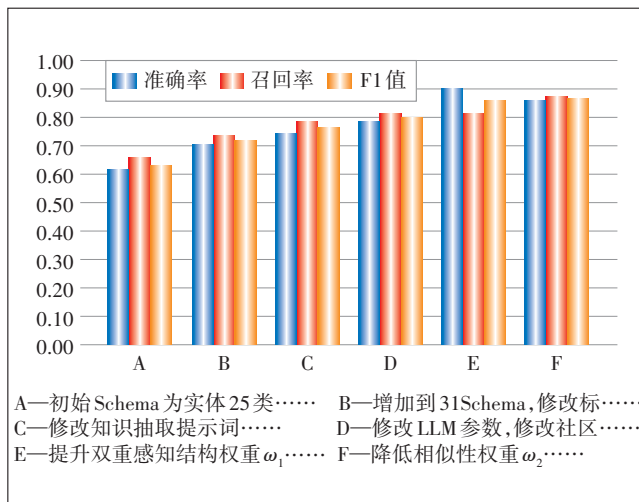


图 3 知识抽取调优评测准确率、召回率、F1 分数柱状图

经过调优,知识抽取准确率、召回率和 F1 分数分别提升到 0.85、0.87 和 0.86,达到较为理想的性能水平,验证了图谱构建和调优方式的有效性。

3.2.2 知识检索 QA 问答评测

QA 知识问答评测基于 317 条测试用例,依托自研知识问答系统完成接口自动化调用。评分规则是针对每道测试问题分别设定核心事实、必要点与次要点,总分设置为 10 分,满分要求核心事实、必要点、次要点全部正确;5 分及以上则要求核心事实完全正确,再结合必要点和次要点的正确比例综合计分;5 分以下说明核心事实开始模糊或内容错误,按各类要点正确比例评分;最低 0 分,为无正确信息。基于此评分规则,对研究构建的 GraphRAG 网络安全知识图谱智能问答系统与自有 RAG 知识库系统进行 QA 问答能力对比评测,具体测试用例如下。

问题事实类:“固定核心网交换设备中,长途交换设备日常维护测试项目中,交换机运行状况监视的维护周期为多久?”参考答案为:“实时”。RAG 系统答案:“交换机运行状况监视的维护周期是实时,每日早、下午与晚忙时段。”评测得分为 7 分,评分依据为核心事实回答正确、必要点无错误,但额外掺杂非必要冗余信息,次要点错误数为 3 项,核心事实错误数与必要点错误数均为 0;GraphRAG 系统的答案为“在固定核心网交换设备中,长途交换设备日常维护测试项目里,交换机运行状况监视的维护周期是实时”,该回答完整贴合标准内容,核心事实、必要点与次要点均无错误,最终取得满分 10 分。317 条测试用例评测对比如表 3 所示。

表3 RAG 与 GraphRAG QA 知识问答评测对比

构建方式		RAG				GraphRAG			
问题类型		比较类	多跳推理类	全局类	事实类	比较类	多跳推理类	全局类	事实类
问题数量/个		48	56	51	162	48	56	51	162
得分个数/个	10分	12	9	16	58	22	16	19	80
	9分	6	12	2	26	9	12	8	28
	8分	16	18	9	39	12	13	16	34
	7分	5	7	4	23	4	6	1	15
	6分	-	1	2	1	-	1	2	-
	5分	-	1	3	3	1	2	2	2
	4分	9	7	4	6	-	6	2	3
	3分	-	-	3	1	-	-	-	-
	2分	-	-	4	1	-	-	-	-
	1分	-	1	1	-	-	-	-	-
	0分	-	-	3	4	-	-	1	-
单类均分		7.77	7.7	6.65	8.25	8.96	8.11	7.96	8.96
总体均分		7.82				8.65			
准确率/%		70.34				84.86			

数据表明,本研究方式构建的图谱在知识问答评测中的平均得分为8.65,准确率达84.86%,较对比测试的RAG系统(平均得分为7.82和准确率为70.34%),分别提升10.61%和20.64%,有效提升了问答性能。

4 结语

针对通信网络架构日益复杂化给网络运行安全所带来的系统性挑战,本研究提出了面向网络安全领域的知识图谱构建和调优方法。该方法依托 GraphRAG 技术框架与大语言模型协同,以图模式 Schema 为驱动,实现网络安全知识图谱的构建。评测结果表明,该方法能有效缓解传统 RAG 方法存在的痛点问题,更加适配通信网络安全领域中多层次结构、多技术域融合、设备异构互联等场景特征,为网络安全等通信领域实现高性能知识检索和智能问答等服务提供了技术路径与实践参考,具有良好的应用前景。

未来研究将从新增业务文档,拓展图谱广度和深度,构建更全面的知识库,适配资源核查、自动排障等通信场景;完善调优与评测机制,GraphRAG 较传统 RAG 在资源算力消耗上没有优势,需要通过持续优化系统架构、数据结构、模型适配性(LLM 和 Embedding 模型)等关键技术提升问答系统性能与资源利用率,同时进一步拓展新的评测维度;探索知识图谱构建与

新兴技术结合(如 Agentic RAG 技术),推动跨领域业务融合等方面的研究。

参考文献:

[1] Chaudhary, Kumarn, Zeadallys. Network service chaining in fog and cloud computing for the 5G environment; data management and security challenges[J]. IEEE Communications Magazine, 2017, 55(11): 114-122.

[2] 张爱玲. 移动运营商网络安全管理平台架构分析与探索[J]. 科技通报, 2017, 33(2): 90-94.

[3] 钟晓峰, 杨国正, 单连勇. 网络安全知识图谱研究综述[J]. 信息对抗技术, 2024, 3(5): 19-29.

[4] 凌峰. 大模型 RAG 应用开发: 构建智能生成系统[M]. 北京: 清华大学出版社, 2025.

[5] Edge D, Trinh H, Chen G N, et al. From local to global: a graph RAG approach to query-focused summarization[EB/OL]. [2026-05-26]. <https://arxiv.org/abs/2404.16130>.

[6] Detroja K, Bhensdadia C, Bhatt B S. A survey on relation extraction[J]. Intelligent Systems with Applications, 2023, 19: 200244.

[7] Mikolov T, Chen K, Corrado G, et al. Efficient estimation of word representations in vector space[EB/OL]. [2026-05-26]. <https://arxiv.org/abs/1301.3781>.

[8] Sarmah B, Mehta D, Hall B, et al. HybridRAG: integrating knowledge graphs and vector retrieval augmented generation for efficient information extraction[C]//Proceedings of the 5th ACM International Conference on AI in Finance. Brooklyn: Association for Computing Machinery, 2024: 608-616.

[9] Chen Y, Zhao P, Li P, et al. Finding communities by their centers[J]. Scientific Reports, 2016, 6: 24017.

[10] Wang K, Liu Y, Xu X J, et al. Knowledge graph embedding with entity neighbors and deep memory network[EB/OL]. [2026-05-26]. <https://arxiv.org/abs/1808.03752?context=cs>.

[11] Arthur D, Vassilvitskiis. K-means++: the advantages of careful seeding[C]//SODA '07: Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms. New Orleans: Society for Industrial and Applied Mathematics, 2007: 1027-1035.

[12] Yang X R, Ma B, Li S Y, et al. Large language model-based idempotent summarization method for educational text[J]. Computer Engineering, 2024, 50(7): 32-41.

作者简介:

屠梓浩, 正高级工程师, 硕士, 主要从事网络 AI 平台、知识图谱知识库、AI 智能体研究建设和管理工作; 龚健, 高级工程师, 学士, 主要从事网络 AI 平台、知识图谱知识库建设和管理工作; 刘磊, 工程师, 学士, 主要从事网络 AI 平台和知识库测评工作; 马先龙, 工程师, 学士, 主要从事网络 AI 平台和知识库建设工作; 周婉君, 工程师, 学士, 主要从事网络 AI 平台和知识库建设工作; 王超, 工程师, 学士, 主要从事网络 AI 平台和知识库建设工作。