

GraphRAG 기반 금융보안 규제 해석 프레임워크

주성용 박준영 오병훈 이재우
중앙대학교 중앙대학교 중앙대학교 중앙대학교
산업보안학과 예술공학부 융합보안학과 산업보안학과
jusy4901@cau.ac.kr june295921@cau.ac.kr danny9807@cau.ac.kr jaewoolee@cau.ac.kr

A GraphRAG Framework for Financial Security Regulation

Seongyong Ju Junyoung Park Byunghoon Oh Jaewoo Lee
Chung-Ang Chung-Ang Chung-Ang Chung-Ang
University, University, University, University,
Department of Department of Department of Department of
Industrial Security Art&Technology Security Convergence Industrial Security

요약

금융 규제는 법률 조항 간의 참조, 예외, 벌칙 등 복잡한 인과관계로 얽혀 있지만 기존 RAG는 분절 단위 텍스트 검색으로 인해 맥락을 반영하지 못한다. 본 연구는 규제 조항 간 인과관계를 'Causal Regulatory Graph'로 모델링하는 GraphRAG 기반 규제 해석 프레임워크를 제안한다. 사용자 질의 관련 조항뿐만 아니라, 인과적으로 연결된 조항까지 컨텍스트로 확장하여 LLM에 제공한다. 국내 LLM과 FSKU 벤치마크를 활용하여 실험한 결과 GraphRAG 방식이 표준 RAG 대비 근소하게 더 높은 정확도를 보였다. 이는 LLM이 조항 간 맥락과 상관 관계를 기반으로 추론하게 하였을 때 유도된 결과로, 복잡한 규제 준수 문제에 있어 관계 중심의 지식 검색이 더 효과적임을 시사한다.

1. 서론

금융업은 다양한 법률과 규제 아래 복합적으로 운영되고 있다. 전자금융거래법, 개인정보보호법, 정보보호관리체계 인증기준 등의 금융, 인증 관련 규제는 단일 조항의 독립적 해석보다 다수 조항의 상호 참조나 예외, 면제 등의 인과적 연결을 형성하여 교차적으로 적용되기도 한다. 이처럼 한 조항의 변화는 다른 조항의 적용, 범위 등에 영향을 미치는 상호 의존성이 존재하여 규제 해석의 복잡성을 가중시킨다. 최근 금융 당국의 보안 감사나 사고 대응 과정을 살펴보면, 특정 행위가 어떤 규제 적용을 받는지나 예외나 처벌 조항으로 이어지는 관계를 추적할 수 있는 체계가 요구되는 실정이다. 그러나 현재 금융 관련 규제 질의 응답 시스템의 대부분은 RAG(Retrieval Augmented Generation)기반의 문장 단위 검색을 사용하고 있다[1]. 이는 각 문서의 개별 조항 의미 파악에는 유용하나, 규제 조항 간의 인과 관계나 영향성을 반영하여 모델링은 어렵다. 따라서 사용자 입장에서 규제의 부분적 내용만 얻을 수 있고, 규제의 논리적 구조나 맥락은 이해하기 어려운 한계점이 존재한다. 이에 텍스트 중심 모델에서 나아가 관계 중심 모델로 확장한 GraphRAG 기반 금융보안 규제 대응 시스템을 제안한다. 이는 규제 조항 간의 인과적 연결을 Casual Regulatory Graph 형태로 모델링하여 해석의 논리적 일관성과 투명성을 강화할 수 있다[2]. 이렇게 기존 RAG 시스템의 한계점을 극복하며 규제 해석 분야에서 LLM의 활용도를 높이는 연구로 발전할 수 있다.

2. 관련 연구

2.1 RAG 외부지식검색

LLM의 생성 과정에 검색 기능을 결합한 방식으로, 외부 지식 문서로부터 관련 텍스트 청크를 검색한 뒤 이를 생성 모델에 입력하여 응답을 생성한다[1]. 이러한 접근은 외부 지식을 활용할 수 있다는 장점이 있으나, 문장 단위 검색에 의존해 문맥적 연관성과 논리적 일관성을 충분히 반영하지 못한다.

2.2 GraphRAG 구조적 지식을 활용한 추론

기존의 국소적인 정보 검색의 한계를 해소하기 위한 접근법으로, LLM이 문서로부터 개체와 관계를 추출하여 그래프를 구성하고 커뮤니티별로 요약한다[4]. 이를 계층적으로 통합하여 전역적인 요약을 생성하고 검색에서 관련 커뮤니티를 병렬적으로 탐색하고 결과를 통합하여 전역적인 응답을 생성하는 방식이다.

3. 제안

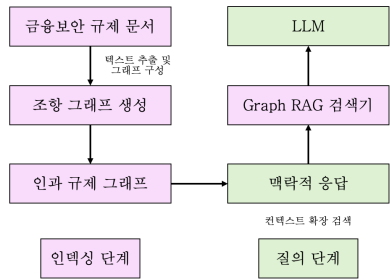
3.1 규제 인과관계 그래프

프레임워크 기반이 되는 규제 인과관계 그래프는 분산된 법률을 노드, 엣지로 구조화된 지식 그래프로 재구성한다.

- 노드(Node): 규제 문서를 조항 단위로 파싱하여 개별 노드로 정의한다. 각각의 노드는 조항의 구성요소인 번호, 내용 등의 데이터를 가진다.
- 엣지(Edge): 조항 간의 단순한 참조가 아니라 인과적 관계를 표현한다. 적용, 위반, 예외 등 관계 유형을 기반으로

로 조항 간 맥락을 분석하여 방향성을 가진 엣지를 생성한다.

3.2 그래프 기반 확장 검색(GraphRAG)



(그림 1) GraphRAG 적용 LLM 응답 프레임워크

그림1은 LLM에 GraphRAG를 적용하였을 때 그래프 구성 및 적용이 이루어지는 단계를 도식화 한 것이다.

GraphRAG는 기존 RAG의 단순한 문서 유사도 기반 검색을 확장하여, 그래프 구조를 활용한 맥락적 검색을 수행한다. 우선 시드 노드 식별 단계에서는 사용자 질의와 의미적으로 유사한 노드를 검색하여 탐색의 출발점으로 선정한다. 이후 그래프 탐색 기반 컨텍스트 확장 단계에서는 시드 노드로부터 인과 관계 그래프를 따라가며 관련된 노드를 탐색한다. 이때 단순히 인접 노드만을 고려하는 것이 아니라, 사전에 정의된 핵심 키워드의 인과 경로를 따라 동적으로 확장함으로써 하위 그래프를 추출한다. 이렇게 구성된 하위 그래프는 질의의 맥락적 의미를 보강하는 컨텍스트로 사용되어, 언어모델이 보다 구조적이고 일관된 응답을 생성할 수 있도록 지원한다.

4. 실험

4.1 데이터셋

FSKU(Financial Security Knowledge Understanding): 금융보안원이 개발한 금융보안 업무 적합도 벤치마크[3]이다. 객관식·주관식을 기반으로 AI 모델의 성능을 비교 선택할 수 있도록 구성된 질문, 답변 데이터를 사용하였다.

4.2 RAG 기반 실험

본 실험의 기준선으로 법령 데이터를 대상으로 관계 그래프를 적용하지 않은 기초형 RAG구조를 구현하였다. 각 법령 문서를 조항 단위로 분할하고 불필요한 기호를 제거해 청크 단위 코퍼스를 구축하였다. 질의는 동일한 토큰나 이저로 전처리한 뒤, BM25 점수를 기준으로 상위 5개 조항을 검색해 컨텍스트를 구성하였다.

4.3 GraphRAG 기반 실험

본 실험에서는 법령 데이터를 대상으로 조항 간의 관계 정보를 반영한 GraphRAG 구조를 구현하였다. 각 법령은 조항 단위로 분할하고 ‘적용’, ‘예외’, ‘벌칙’, ‘면제’, 등의 관계 유형을 중심으로 그래프를 구축하였다. 그래프의 노드는 법령 조항을, 엣지는 조항 간 관계 유형을 나타내며, 약 3,000개의 노드와 3,000개의 엣지를 포함한다. RAG 인덱싱 단계에서는 E5-ko 임베딩을 활용하여 FAISS 벡터

인덱스를 구축하였다. GraphRAG 질의 단계에서는 동일한 임베딩을 사용해 상위 3개의 시드 노드를 검색한 뒤, 1-hop 범위 내 인접 노드를 확장하여 문맥 컨텍스트를 구성하고 이를 시스템 프롬프트로 언어모델에 입력하였다.

4.3 실험 결과

모델	성능 지표
gemma-ko-7b	0.3048888594
EXAONE-3.5-7.8B	0.5090990482
RAG EXAONE	0.5561523137
Graph RAG EXAONE	0.5787359189

5. 결론

본 연구는 금융 규제의 복잡성을 해석하기 위해, 조항 간 인과관계 모델링으로 GraphRAG 기반 규제 해석 프레임워크를 제안하였다. 특히 소비권 AI 확보가 중요한 금융 규제 도메인에서 국내 제작 LLM을 기반으로 실험을 진행하였다. 실험 결과로 GraphRAG는 표준 RAG보다 FSKU 벤치마크에서 더 높은 지표를 보였다. 이는 GraphRAG가 인과적으로 연결된 조항을 구조화된 컨텍스트와 함께 제공하여, LLM이 텍스트 청크 검색에서 나아가 맥락과 상관 관계를 기반으로 추론하도록 유도했기 때문이다. 다만, 실험 결과 그래프는 키워드 기반으로 구축되어 복잡한 관계를 놓칠 수 있고, 탐색 범위 또한 1-hop으로 제한적이다. 특히 RAG 자체의 텍스트 임베딩 기반 검색이 어느 정도는 유지되므로 베이스 라인과 유사한 성능으로 수렴할 수 있다. 또한 FSKU 벤치마크 지표만으로 ‘해석의 논리성’ 자체를 평가하기 어렵다는 한계점이 존재한다. 그럼에도 그래프에서 추출한 ‘구조적 맥락으로 관련 조항들의 연결망을 함께 제공하는 것이 표준 RAG보다 규제 해석 정확도를 향상시킴을 입증했다. 이는 복잡성을 지닌 규제 해석에 있어, 관계 중심의 지식 검색이 더 효과적임을 시사한다.

참고 문헌

- [1] LEWIS, Patrick, et al. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in neural information processing systems*, 2020, 33: 9459-9474.
- [2] EDGE, Darren, et al. From local to global: A graph rag approach to query-focused summarization. *arXiv preprint arXiv:2404.16130*, 2024.
- [3] 금융보안원(2024). “FSKU(Financial Security Knowledge Understanding) 벤치마크 개발 및 공개.” 금융보안원 보도자료
- [4] Yao, Y., Liu, H., Xiao, C., Chen, Z., & Wang, W. Y. (2024). *GraphRAG: Enhancing Large Language Models with Graph-Based Retrieval and Reasoning*. *arXiv preprint arXiv:2404.16130*.