Machine Learning Models in Predicting Bond Default Risk Based on LIME and SHAP [Электронный документ] / arXiv preprint arXiv:2502.19615, 2025, режим доступа: https://arxiv.org/abs/2502.19615 (дата обращения: 10.02.2026).

Вульфин Алексей Михайлович, д.т.н., проф. кафедры вычислительной техники и защиты информации, vulfin.alexey@gmail.com.
Кириллова Анастасия Дмитриевна, к.т.н., доцент каф. ВТиЗИ, kirillova.ad@ugatu.su.
Садыкова Эльвина Евгеньевна, аспирант каф. вычислительной техники и защиты информации, elvina.sadykova.01@mail.ru.

УДК 004.7.056; 004.7:004.8

МОДЕЛЬ АНАЛИЗА СЛАБОСТРУКТУРИРОВАННЫХ ДАННЫХ ОБ УГРОЗАХ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ НА ОСНОВЕ ГРАФОВ ЗНАНИЙ

А.И. Луцкович, А.М. Вульфин, А.Д. Кириллова
Уфимский университет науки и технологий, г. Уфа

Ключевые слова: граф знаний, расширенная дополненная генерация, тактики и техники, анализ угроз информационной безопасности.

Для решения задач обработки слабоструктурированных данных о техниках, тактиках и процедурах (TTP) сегодня все чаще привлекаются технологии искусственного интеллекта (ИИ), связанные с использованием Large Language Models (LLMs) [1].

Одним из эффективных подходов к уменьшению эпизодов «галлюцинаций» и устаревших данных является предложенная в 2020 году технология Retrieval Augmented Generation (RAG) – генерация ответов с расширенным поиском [2]. Архитектура RAG-системы (рисунок 1) в общем виде состоит их двух основных частей – поисковика и генератора. Поисковик на основании получаемого запроса осуществляет поиск наиболее близких к смыслу запроса релевантных документов в векторной базе данных (БД), где хранятся документы компании, относящиеся к вопросам обеспечения ее информационной безопасности (ИБ). Полученные таким образом документы содержат максимально полезную контекстную информацию, необходимую для последующей обработки запроса модулем генератора. Генератор, используя полученную от поисковика информацию (запрос + найденный контекст), генерирует с помощью LLM ответ на поступивший в систему запрос пользователя.

Примеры успешного применения RAG-систем в задачах обработки и анализа слабоструктурированных данных в домене ИБ приводятся в достаточно большом числе публикаций (например, [3, 4]).

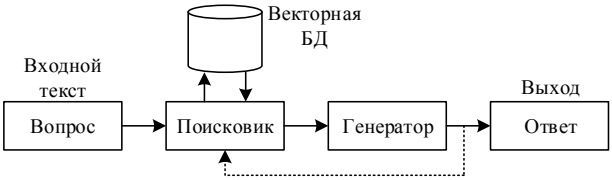

Рисунок 1 – Архитектура RAG-системы

Цель работы – проектирование, реализация и экспериментальная оценка эффективности GraphRAG-подхода для анализа данных в области ИБ.

В отличие от классического RAG, где контекст формируется исключительно из фрагментов документов, система приоритизирует графовые факты банка данных угроз (БДУ ФСТЭК) (объекты, последствия, способы реализации, меры) и использует текстовые доказательства как дополнительный слой обоснования.

Граф знаний BDU-KG построен на основе интеграции нескольких предметных онтологий, объединенных в единую модель данных. Онтологическая модель включает шесть взаимосвязанных доменов БДУ ФСТЭК, MITRE ATT&CK, MITRE D3FEND, CWE-CVE-CAPEC, TTP-документы. Онтологическая модель BDU-KG реализована в графовой СУБД Neo4j в виде размеченного графа свойств.

Вычислительный эксперимент проведен для сценария Crosswalk-QA «Построение сквозных сопоставлений». На вход подается фрагмент TTP-отчета, требуется: извлечь TTP и другие сущности, сопоставить их с УБИ, сформировать объяснение через подграф «TTP – ATT&CK – БДУ – БДУ-меры – D3FEND». Для сценария сформирован набор из 250 тестовых вопросов с эталонной разметкой. Наборы вопросов сформированы стратифицировано по типам запросов и целевым сущностям.

Для оценки устойчивости результатов GraphRAG к выбору LLM проведен эксперимент с матрицей 3x3: три LLM (Gemma3:27B, Llama-3.1-70B-Instruct, Qwen2.5-32B-Instruct) и три embedding-модели (BGE-M3, multilingual-e5-large, LaBSE) (таблица 1, лучший результат выделен жирным).

Таблица 1 – Качество ответов Crosswalk-QA

Embedding / LLM	Gemma3:27B	Llama-3.1-70B	Qwen2.5-32B
	EM / F1 / Faithfulness	EM / F1 / Faithfulness	EM / F1 / Faithfulness
BGE-M3	**0,63 / 0,74 / 0,90**	0,61 / 0,72 / 0,88	0,59 / 0,70 / 0,87
multilingual-e5-large	0,58 / 0,69 / 0,88	0,56 / 0,67 / 0,86	0,54 / 0,65 / 0,85
LaBSE	0,55 / 0,66 / 0,86	0,53 / 0,64 / 0,84	0,51 / 0,62 / 0,83

Метрика качества ответа на уровне идентификаторов Exact Match (EM) принимает значение 1 при полном совпадении множества идентификаторов в ответе с эталонным и 0 в противном случае. F1-score

представляет собой гармоническое среднее точности (доля корректных идентификаторов среди выданных) и полноты (доля найденных идентификаторов среди эталонных). Метрика RAGAS Faithfulness оценивает долю утверждений в ответе, подкрепленных информацией из контекста; низкие значения указывают на присутствие «галлюцинаций».

Прирост EM для всех комбинаций, что подтверждает устойчивость преимущества GraphRAG с межонтологическими связями. Влияние выбора embedding-модели превышает влияние выбора LLM, что указывает на критичность качества retrieval для GraphRAG-систем. Комбинация Gemma3:27B и BGE-M3 обеспечила наилучшие абсолютные показатели.

Практическая значимость определяется возможностью применения разработанной системы для поддержки принятия решений аналитиками центров мониторинга информационной безопасности при разборе TTP-отчетов, формировании рекомендаций по мерам защиты в соответствии с требованиями российского законодательства и атрибуции инцидентов с использованием международных таксономий.

Список использованных источников

1.Котенко И.В., Абраменко Г.Т. Использование больших языковых моделей для поиска угроз кибербезопасности на основе методов глубокого обучения: анализ современных исследований // Правовая информатика. 2024. №. 3. С. 32-42.

2.Lewis P. et al. Retrieval-augmented generation for knowledge-intensive NLP tasks // Advances in neural information processing systems. 2020. Vol. 33. P. 9459-9474.

3.Singh J.P., Agrawal S. Automating threat intelligence analysis with retrieval augmented generation rag for enhanced cybersecurity posture // International Journal of Science and Research (IJSR). 2024. Vol. 13, no. 5. P. 251-255.

4.Rajapaksha S., Rani R., Karafili E. A RAG-based question-answering solution for cyber-attack investigation and attribution // European Symposium on Research in Computer Security. 2024. P. 238-256.

Луцкович Андрей Иванович, ст. препод. каф. ВТиЗИ, andrey.lutskovich@gmail.com.
Вульфин Алексей Михайлович, д.т.н., профессор каф. ВТиЗИ, vulfin.alexey@gmail.com.
Кириллова Анастасия Дмитриевна, к.т.н., доцент каф. ВТиЗИ, kirillova.ad@ugatu.su.

УДК 004.7.056; 004.7:004.8

МУЛЬТИМОДАЛЬНАЯ НЕЙРОСЕТЕВАЯ СИСТЕМА ОБНАРУЖЕНИЯ ФИШИНГОВЫХ ВЕБ-САЙТОВ

А.В. Минко, А.М. Вульфин, А.Д. Кириллова
Уфимский университет науки и технологий, г. Уфа

Ключевые слова: фишинг, мультимодальное обнаружение, позднее слияние, объяснимый искусственный интеллект, кибербезопасность

Традиционные методы обнаружения (черные списки, статический