УДК 004.056.52

ВНЕДРЕНИЕ МЕХАНИЗМОВ РОЛЕВОГО КОНТРОЛЯ ДОСТУПА В RAG-ПАЙПЛАЙНЫ

Патрикеев Р.О. (ИТМО)
Научный руководитель - кандидат технических наук, доцент Коржук В.М. (ИТМО)

Введение. Стремительное развитие генеративных языковых моделей и архитектур дополненной генерации (Retrieval-Augmented Generation, RAG) создало основу для нового класса корпоративных систем, способных работать с распределенными массивами знаний и обеспечивать интеллектуальный доступ к ним. Однако практическое внедрение таких решений выявило системную проблему: большинство RAG-пайплайнов разрабатывались без изначального учёта механизмов разграничения доступа к данным. В результате контроль безопасности реализуется как внешняя надстройка — через фильтрацию запросов, изоляцию сервисов или постобработку ответов, — что не обеспечивает строгих гарантий предотвращения утечек конфиденциальной информации и не учитывает специфику семантического поиска [1, 2].

Современные исследования демонстрируют наличие фундаментального компромисса между безопасностью, производительностью и управляемостью данных. Использование отдельных индексов для различных ролей упрощает контроль доступа, но приводит к дублированию векторных представлений и росту затрат на хранение. Напротив, единый индекс с пост-фильтрацией результатов снижает точность поиска и увеличивает вычислительные задержки [3]. На практике ситуация дополнительно осложняется необходимостью интеграции с корпоративными системами управления доступом (IAM), где изменение политик требует повторной обработки данных. Это вызывает рост latency при выполнении запросов, необходимость повторной генерации эмбеддингов (re-embedding) при изменении списков доступа (ACL) и проблемы высокой кардинальности ролей, когда количество комбинаций прав начинает ограничивать масштабируемость системы. Таким образом, актуальной научной задачей становится разработка методологии, позволяющей интегрировать ролевой контроль доступа непосредственно в архитектуру RAG, а не применять его как внешний механизм.

Цель работы. Целью работы является повышение защищенности RAG систем путем внедрения ролевого контроля доступа на разных этапах этапах обработки и выдачи данных.

Основная часть. Предлагаемый подход рассматривает контроль доступа как сквозной процесс, сопровождающий данные на всех стадиях жизненного цикла — от подготовки корпуса документов до формирования ответа модели. На этапе индексации текстовые фрагменты обогащаются метаданными, отражающими допустимые роли и политики доступа. Это позволяет связать семантическое представление информации с её контекстом безопасности уже на уровне векторного индекса и избежать необходимости последующей полной переработки данных [4].

На этапе поиска используется гибридная стратегия, сочетающая предварительное ограничение выборки по метаданным и последующую проверку результатов. Предварительная фильтрация уменьшает размер поискового пространства

и снижает задержки, тогда как дополнительная проверка обеспечивает корректность при динамическом изменении прав доступа. Для снижения избыточности хранения применяется адаптивное логическое разделение индексов по ролевым группам, позволяющее учитывать кардинальность ролей без физического дублирования данных [3, 5]. На этапе генерации вводятся механизмы валидации ответа, предотвращающие косвенное раскрытие информации через агрегированный контекст, что особенно важно для RAG-систем, использующих объединение нескольких источников знаний [6].

Новизна подхода заключается в трактовке RBAC не как внешнего слоя авторизации, а как структурного элемента RAG-архитектуры, влияющего на процессы индексирования, извлечения и генерации текста.

Выводы. Предложенная модель позволяет:

- обеспечить соблюдение политик доступа без существенной деградации качества семантического поиска;
- снизить влияние изменений ACL на необходимость повторной индексации данных;
- контролировать рост задержек за счет ранней фильтрации и оптимизации поискового пространства;
- учитывать высокую кардинальность ролей при масштабировании корпоративных систем;
- минимизировать риски косвенных утечек информации при генерации ответов.

Практическая реализация предполагает апробацию в пилотных корпоративных системах, разработку методических рекомендаций по безопасному построению RAG-пайплайнов и создание библиотек интеграции с векторными базами данных и IAM-инфраструктурами.

Список использованных источников

1. Lewis P. et al. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks // *Advances in Neural Information Processing Systems*. 2020.
2. OWASP Foundation. *OWASP Top 10 for Large Language Model Applications*. 2024.
3. Zhong H., Lentz M., Narodytska N. et al. HONEYBEE: Efficient Role-Based Access Control for Vector Databases via Dynamic Partitioning // arXiv:2505.01538, 2025.
4. Bhatt S. et al. Enterprise AI Must Enforce Participant-Aware Access Control // arXiv:2509.14608, 2025.
5. Sandhu R., Coyne E., Feinstein H., Youman C. Role-Based Access Control Models // *IEEE Computer*. 1996.
6. Sun S. et al. SMA: Auditing Membership Leakage in Retrieval-Augmented Generation Systems // arXiv:2508.09105, 2025.
7. Manning C., Raghavan P., Schütze H. *Introduction to Information Retrieval*. Cambridge University Press, 2008.
8. Knuth D., Mitzenmacher M. et al. Challenges in Secure and Scalable Vector Search Systems // *Communications of the ACM*. 2023.

Автор _____ Патрикеев Р.О.

Научный руководитель _____ Коржук В.М.