Интернет-журнал «Отходы и ресурсы» https://resources.today
Russian Journal of Resources, Conservation and Recycling

2025, Том 12, № 2 / 2025, Vol. 12, Iss. 2 https://resources.today/issue-2-2025.html

URL статьи: https://resources.today/PDF/07INOR225.pdf

DOI: 10.15862/07INOR225 (https://doi.org/10.15862/07INOR225)

2.3.1. Системный анализ, управление и обработка информации, статистика (технические науки)

Ссылка для цитирования этой статьи:
Бубнова, Е. Ю. Адаптивное шифрование текстовых данных для систем обработки естественного языка на основе крупных нейросетевых моделей / Е. Ю., Бубнова Д. О. Якупов // Отходы и ресурсы. — 2025. — Т. 12. — № 2. — URL: https://resources.today/PDF/07INOR225.pdf DOI: 10.15862/07INOR225.

For citation:
Bubnova E.Yu., Yakupov D.O. Adaptive encryption of text data for natural language processing systems based on large neural network models. *Russian Journal of Resources, Conservation and Recycling*. 2025;12(2): 07INOR225. Available at: https://resources.today/PDF/07INOR225.pdf. DOI: 10.15862/07INOR225. (In Russ., abstract in Eng.).

УДК 004.056.55

Бубнова Елена Юрьевна
ФГБОУ ВО «Поволжский государственный университет телекоммуникаций и информатики», Самара, Россия
E-mail: 89061250055@mail.ru
ORCID: https://orcid.org/0009-0003-1598-8971

Якупов Денис Олегович
ФГБОУ ВО «Поволжский государственный университет телекоммуникаций и информатики», Самара, Россия
Старший преподаватель, ассистент кафедры «Программной инженерии», аспирант
E-mail: d.yakupov@psuti.ru
ORCID: https://orcid.org/0009-0003-2371-0822
РИНЦ: https://elibrary.ru/author_profile.asp?id=1175874

Адаптивное шифрование текстовых данных для систем обработки естественного языка на основе крупных нейросетевых моделей

Аннотация. Современные предприятия активно внедряют системы искусственного интеллекта и автоматизированной обработки естественного языка для оптимизации работы с корпоративной информацией. Однако применение этих технологий сопряжено с существенными рисками нарушения конфиденциальности данных, особенно при анализе неструктурированных текстовых материалов. Существующие криптографические методы защиты информации демонстрируют ограниченную эффективность при обработке языковых данных, а современные подходы к безопасной обработке информации имеют существенные недостатки: методы гомоморфного шифрования значительно снижают скорость вычислений, а технологии распределенного обучения сталкиваются с проблемами согласования разнородных данных.

В представленном исследовании разработан комплексный метод защиты информации, сочетающий преимущества гомоморфного шифрования и распределенного обучения. Основное научное достижение работы заключается в создании специализированных криптографических алгоритмов для обработки языковых данных, включая идентификацию именованных объектов в зашифрованном виде, а также в оптимизации соотношения между уровнем защиты и вычислительной эффективностью. Предлагаемая методика включает дифференцированное шифрование различных типов данных, децентрализованную обработку

информации с автоматической корректировкой параметров обучения, а также механизмы обеспечения анонимности и контроля доступа.

Экспериментальная часть исследования оценивает три ключевых параметра: устойчивость системы к потенциальным атакам, скорость обработки защищенных данных и точность работы аналитических алгоритмов. Практическая ценность работы заключается в создании безопасной платформы для внедрения интеллектуальных систем анализа данных в финансовом секторе, здравоохранении и телекоммуникационной отрасли, где утечка информации может привести к серьезным последствиям. Перспективным направлением дальнейших исследований является разработка криптографических алгоритмов, устойчивых к методам квантового взлома.

Ключевые слова: адаптивное шифрование; гомоморфное шифрование; федеративное обучение; конфиденциальность данных; токенизация; корпоративные базы данных; большие данные

Современный этап цифровой трансформации бизнес-процессов характеризуется активным внедрением технологий искусственного интеллекта (ИИ) и языковых моделей большого размера LLM (Large Language Models) для автоматизации структурирования корпоративных данных. По данным аналитиков Gartner, в 2025 году более 60 % компаний планируют использовать нейросетевые решения для обработки неструктурированной информации, включая текстовые отчеты, транзакционные записи и клиентские коммуникации.[1] Однако стремительное распространение таких технологий сопряжено с критическими рисками, связанными с обеспечением конфиденциальности и целостности данных. В условиях роста кибератак, особенно в сегменте корпоративного шпионажа, в котором ущерб в 2022 году превысил 6 триллионов долларов по миру, уязвимость конфиденциальной информации на этапах её автоматизированной обработки нейросетевыми системами становится ключевым вызовом для организаций.[2]

Основная проблема заключается в том, что традиционные методы шифрования, такие как AES (Advanced Encryption Standard) или RSA (Rivest Shamir Adleman), не адаптированы для работы в конвейерах обработки естественного языка NLP (**Natural Language Processing**) и нейросетевого обучения [1]. Например, дешифровка данных перед их подачей в модель создает «окна уязвимости», эксплуатируемые злоумышленниками, как это произошло в инциденте с утечкой 37 млн записей клиентов крупного банка в 2021 году из-за некорректной интеграции криптографических протоколов в ML-пайплайн.[3] Особую сложность представляет обработка неструктурированных текстовых данных, где конфиденциальные сущности, такие как персональные данные или финансовые показатели распределены в свободном формате, что затрудняет их точечное шифрование без нарушения семантической целостности.

Целью данного исследования является разработка гибридного подхода к шифрованию, обеспечивающего сквозную защиту данных на всех этапах их обработки — от первичной токенизации до обучения прогнозных моделей. В отличие от существующих решений, предлагаемый метод сочетает гомоморфное шифрование, позволяющее выполнять вычисления

[1] Top Strategic Technology Trends for 2025 // Gartner URL: https://www.gartner.com/en/documents/5850447 (дата обращения: 07.03.2025).

[2] Kaspersky predicts rise in cyber espionage for India in 2022 // Kaspersky URL: https://www.expresscomputer.in/news/kaspersky-predicts-rise-in-cyber-espionage-for-india-in-2022/82949/ (дата обращения: 12.02.2025).

[3] Обзор отчетности об инцидентах информационной безопасности при переводе денежных средств // Банк России URL: https://cbr.ru/statistics/ib/review_2q_2021/ (дата обращения: 25.02.2025).

над зашифрованными текстовыми векторами, и федеративное обучение, минимизирующее передачу сырых данных между узлами.

Научная новизна работы заключается в адаптации криптографических алгоритмов под специфику NLP-задач, таких как распознавание именованных сущностей NER (**Named Entity Recognition**) в зашифрованном пространстве, и создании механизмов динамической балансировки между уровнем безопасности и вычислительной нагрузкой. Это открывает возможность интеграции защиты данных в реальном времени без деградации производительности аналитических систем, что подтверждается пилотными внедрениями в CRM-платформах (**Customer Relationship Management**) розничных сетей.

Современные криптографические алгоритмы, такие как AES и RSA, составляют основу защиты данных в корпоративных системах [2]. AES, как симметричный алгоритм, обеспечивает высокую скорость шифрования больших массивов структурированных данных, что делает его предпочтительным для шифрования баз данных и транзакционных журналов. RSA, основанный на асимметричной криптографии, широко применяется для безопасного обмена ключами и аутентификации. Однако в контексте обработки данных искусственным интеллектом эти методы демонстрируют принципиальные ограничения. Дешифровка информации перед её подачей в нейросетевые модели создает временные «окна уязвимости», когда конфиденциальные данные находятся в открытом, незашифрованном виде [3]. Кроме того, NLP-пайплайны, включающие этапы токенизации и семантического анализа, требуют доступа к сырым текстовым данным, что исключает возможность применения классических алгоритмов на этапах предобработки.

Гомоморфное шифрование HE (**Homomorphic Encryption**) предлагает альтернативу, позволяя выполнять вычисления над зашифрованными данными без их расшифровки. Технологии частичного PHE (**Partially Homomorphic Encryption**) и полностью гомоморфного шифрования FHE (**Fully Homomorphic Encryption**), такие как схемы BFV (Brakerski/ Fan-Vercauteren) и CKKS (Cheon-Kim-Kim-Song), уже используются в облачных аналитических системах для выполнения операций суммирования и умножения на зашифрованных финансовых показателях [4]. Однако адаптация HE для NLP-задач, таких как векторное представление текста или распознавание именованных сущностей, остается проблематичной из-за экспоненциального роста вычислительной нагрузки при работе с высокоразмерными текстовыми векторами. Эксперименты IBM 2022 года показали, что обработка зашифрованного текста методом BERT (**Bidirectional Encoder Representations from Transformers**) увеличивает время инференса в 15–20 раз по сравнению с открытыми данными, что неприемлемо для систем реального времени [5].

Федеративное обучение FL (**Federated Learning**) решает проблему конфиденциальности за счет децентрализованного обучения моделей на распределенных данных. В этом подходе исходные данные остаются на локальных узлах, это могут быть филиалы компании или устройства пользователей, а на центральный сервер передаются только обновления градиентов. Технология успешно внедрена в мобильных приложениях для прогнозирования персональных предпочтений, как в случае с клавиатурой Gboard от Google. Для NLP-задач FL позволяет обучать модели на корпоративной переписке или клиентских отзывах без централизации конфиденциальных текстов. Тем не менее, FL сталкивается с проблемой неоднородности данных (non-IID), когда распределение информации между узлами существенно различается, что снижает точность моделей [6].

Анализ существующих решений выявляет их несовместимость с требованиями NLP-пайплайнов. Классические методы AES и RSA нарушают целостность текстовых данных на этапах предобработки, HE увеличивает latency до неприемлемых значений, а FL не решает проблему защиты данных внутри локальных узлов. Например, токенизация зашифрованного

текста стандартными инструментами (например, spaCy) невозможна без дешифровки, что ставит под угрозу конфиденциальность [7]. Кроме того, современные методы не обеспечивают сквозное шифрование на всех этапах — от первичной обработки до визуализации аналитических отчетов. Это создает необходимость в гибридных подходах, сочетающих криптографическую защиту с адаптацией нейросетевых архитектур под работу в зашифрованном пространстве.

Основу предлагаемого подхода составляет синтез гомоморфного шифрования и адаптивных нейросетевых архитектур, разработанных для комплексной обработки различных типов корпоративных данных. Данная методология предполагает дифференцированное применение криптографических методов в зависимости от характера обрабатываемой информации, что позволяет достичь оптимального баланса между безопасностью и производительностью системы. Для структурированных данных, к которым относятся финансовые транзакции, бухгалтерские записи и другие формализованные информационные массивы, используется усовершенствованная версия схемы CKKS (Cheon-Kim-Kim-Song). Этот алгоритм гомоморфного шифрования был специально модифицирован для выполнения арифметических операций над зашифрованными числовыми векторами с контролируемым уровнем погрешности, что особенно важно при обработке финансовых показателей и других критически важных числовых данных [8].

Для работы с неструктурированными текстовыми данными (корпоративная переписка, аналитические отчеты, клиентские обращения) разработан инновационный гибридный протокол защиты. Его ключевая особенность заключается в комбинированном использовании различных криптографических методов: частичное гомоморфное шифрование применяется для семантически значимых фрагментов текста, содержащих именованные сущности (персональные данные, финансовые показатели, коммерческую тайну), в то время как для остального контента используется стандартное симметричное шифрование по алгоритму AES-GCM (**Advanced Encryption Standard — Galois/Counter Mode**). Такой избирательный подход позволяет существенно снизить вычислительную нагрузку на систему, сохраняя при этом возможность выполнения сложных операций обработки естественного языка непосредственно над зашифрованными текстовыми данными [9]. Особое внимание уделено сохранению семантической целостности текста после криптографических преобразований, что является обязательным условием для последующего лингвистического анализа.

Интеграция принципов федеративного обучения в процесс формирования и обработки корпоративных баз данных реализуется через децентрализованную архитектуру, где каждый узел системы (например, региональное отделение компании или отдельный сервер) осуществляет предварительную обработку и шифрование локальных данных перед их агрегацией. В отличие от традиционных централизованных систем, где происходит передача исходных или частично зашифрованных записей на главный сервер, в предлагаемом подходе узлы обмениваются исключительно зашифрованными градиентами, вычисленными на основе локально обученных моделей. Для решения проблемы неоднородности данных (non-IID problem), возникающей из-за статистических различий между распределениями информации на разных узлах, применяется специально разработанный алгоритм динамической перевесовки обновлений. Этот алгоритм учитывает степень расхождения в данных между узлами и соответствующим образом корректирует весовые коэффициенты при агрегации моделей, что позволяет сохранить высокую точность прогнозирования при обеспечении полной конфиденциальности исходных данных. Эффективность данного подхода подтверждена серией тестов на специализированных синтетических наборах данных, максимально приближенных по своей структуре к реальным корпоративным информационным массивам [10].

Экспериментальная проверка предложенного метода включает комплексную оценку по трем ключевым направлениям, каждое из которых имеет критическое значение для практического внедрения технологии. Первое направление — оценка безопасности системы — проводится с использованием специализированного инструментария FHEBench, позволяющего моделировать различные типы криптографических атак, включая атаки с выбранным шифротекстом (Chosen Ciphertext Attacks). Особое внимание уделяется анализу потенциальных утечек информации через побочные каналы, что особенно актуально для гомоморфных схем шифрования. Второе направление — тестирование производительности — включает детальные замеры времени обработки зашифрованных текстовых данных на платформе Microsoft SEAL с последующим сравнением с результатами работы с открытыми данными. В рамках данного тестирования особое внимание уделяется оптимизации параллельных вычислений для современных GPU-кластеров, что позволяет существенно сократить временные затраты при обработке больших объемов информации. Третье направление — оценка точности работы моделей — предполагает тестирование качества выполнения базовых задач обработки естественного языка, таких как распознавание именованных сущностей (Named Entity Recognition) и классификация текста, на специализированных зашифрованных корпусах, включая Reuters-21578. Для этих целей используются модифицированные версии современных языковых моделей RoBERTa и GPT-3-small, адаптированные для работы с зашифрованными текстовыми данными [11].

Реализация механизмов анонимизации представляет собой важный дополнительный уровень защиты конфиденциальной информации. В рамках предлагаемого подхода разработаны специализированные протоколы, позволяющие автоматически идентифицировать и обрабатывать персональные данные на этапе предобработки информации. Для этого используются комбинированные методы, включающие как полное удаление идентифицирующих признаков, так и их замену псевдонимами с сохранением семантического контекста. Особенностью системы является реализация механизма гарантированного удаления информации по истечении заданного срока хранения, что достигается за счет криптографического обнуления соответствующих данных в зашифрованных базах без нарушения целостности остальной информации. Это позволяет компаниям полностью соответствовать требованиям современных регуляторов в области защиты персональных данных, включая положения GDPR и аналогичных нормативных актов.

Обеспечение прозрачности и подотчетности алгоритмов является важным аспектом предлагаемого решения. Для этого разрабатываются детальные спецификации взаимодействия всех компонентов системы, включая этапы шифрования, обработки естественного языка и машинного обучения. Вся документация по используемым криптографическим схемам и алгоритмам обработки данных делается открытой для аудита, что позволяет независимым экспертам и регуляторам проверить соответствие системы заявленным требованиям безопасности. Одновременно с этим реализован механизм защищенного журналирования всех операций с данными, где записи хранятся в зашифрованном виде, но могут быть декодированы авторизованными аудиторами при наличии соответствующих мультиподписных ключей. Такой подход обеспечивает оптимальный баланс между требованиями конфиденциальности и необходимостью обеспечения прозрачности работы системы для регулирующих органов.

Экспериментальная проверка метода включает три группы метрик:

1. Безопасность: оценка стойкости к атакам на основе модели FHEBench с имитацией атак типа CCA (Chosen Ciphertext Attacks) и анализом утечек через побочные каналы.
2. Производительность: замеры времени обработки зашифрованных текстовых данных на платформе Microsoft SEAL в сравнении с открытыми аналогами, включая оптимизацию параллельных вычислений для GPU-кластеров.

Издательство «Мир науки» \ Publishing company «World of science» http://izd-mn.com

3. Точность моделей: тестирование качества распознавания именованных сущностей и классификации текста на зашифрованных корпусах (например, Reuters-21578) с использованием модифицированных архитектур RoBERTa и GPT-3-small.

В рамках работы разрабатываются протоколы анонимизации, позволяющие удалять или заменять псевдонимами персональные идентификаторы на этапе предобработки данных. Для реализации права на удаление информации предложен механизм удаления записей через определенный промежуток времени в зашифрованных базах данных через криптографическое обнуление соответствующих векторов без нарушения целостности остальной информации.

Прозрачность алгоритмов станет критическим аспектом для аудита регуляторами. Предполагается создать открытые спецификации взаимодействия компонентов системы (шифрование → NLP → обучение), включая документацию по используемым криптографическим схемам. Для верификации процессов планируется внедрение журналирования операций в зашифрованном виде с возможностью декодирования аудиторами при наличии мультиподписных ключей. Это обеспечит баланс между конфиденциальностью данных и подотчетностью системы.

Этические риски исследования связаны с потенциальным конфликтом между уровнем безопасности и практической применимостью технологии. Например, усиление шифрования может увеличить задержки обработки данных в реальном времени, что негативно скажется на бизнес-процессах. Для минимизации этого дисбаланса в методологию будет заложен адаптивный алгоритм, динамически регулирующий глубину шифрования в зависимости от критичности данных и требований SLA (Service Level Agreement).

Практическая значимость исследования заключается в создании методологии, которая позволит компаниям интегрировать нейросетевые технологии обработки данных без компромиссов в области информационной безопасности. Ожидается, что предложенный гибридный подход к шифрованию, сочетающий гомоморфные схемы и федеративное обучение, станет основой для разработки стандартизированных решений в корпоративной аналитике. Это особенно актуально для отраслей с повышенными требованиями к конфиденциальности, таких как финансы, здравоохранение и телекоммуникации, где утечки данных несут катастрофические репутационные и финансовые риски.

Перспективным направлением дальнейших исследований является адаптация метода к квантово-устойчивым алгоритмам шифрования, что обеспечит долгосрочную защиту от угроз, связанных с развитием квантовых вычислений.

ЛИТЕРАТУРА

1. Зацепина А.И. Шифрование данных // StudNet. — 2022. — № 1. — С. 344–351.
2. Москвин А.Д., Петросян Л.Э. Анализ современных алгоритмов шифрования данных // Инженерный вестник Дона. — 2023. — № 4. — С. 102–115.
3. Чичикин Г. Я., Семёнов Д. А. Криптосистема RSA // Наука, образование и культура. — 2019. — № 5. — С. 15–17.
4. Гаража А.А., Герасимов И.Ю., Николаев М.В., Чижов И.В. Об использовании библиотек полностью гомоморфного шифрования // International Journal of Open Information Technologies. — 2021. — №3. — С. 11–22.

5. Масленникова С.С., Коротков В.В. Применение BERT для классификации сообщений в службу поддержки SAP // Материалы XXII Международной научно-практической конференции имени Э.К. Алгазинова. — Воронеж: Общество с ограниченной ответственностью "Вэлборн", 2022. — С. 1164–1170.
6. Ефремов М.А., Холод И.И Разработка архитектуры универсального фреймворка федеративного обучения // Программные продукты и системы. — 2022. — № 2. — С. 263–272.
7. Корниенко С.А., Гохович В.А. Обзор криптоалгоритмов RSA и AES. Оценка их криптостойкости // Синергия Наук. — Москва: Сиденко Александр Сергеевич, 2023. — С. 460–466.
8. Салып Б.Ю., Смирнов А.А. Анализ модели BERT как инструмента определения меры смысловой близости предложений естественного языка // Научно-образовательный журнал для студентов и преподавателей "StudNet". — 2022. — № 5. — С. 3509–3518.
9. Feroz Khan A.B., Kalpana D.S., Devi K.R. An enhanced AES-GCM based security protocol for securing the IoT communication // Scientific and Technical Journal of Information Technologies, Mechanics and Optics. — 2023. — № 4. — С. 711–719.
10. Кусакин И.К., Цурупа А.М., Алмакаев А.В., Романов А.Ю. Использование BERT для классификации коротких научных текстов на русском языке // НТИ-2022. Научная информация в современном мире: глобальные вызовы и национальные приоритеты. — Москва: Всероссийский институт научной и технической информации РАН, 2022. — С. 103–109.
11. Маслова М.А., Дмитриев А.С., Холкин Д.О. Методы распознавание именованных сущностей в русском языке // Журнал Инженерный вестник Дона. — 2021. — № 7. — С. 93–105.

Bubnova Elena Yurievna
Povolzhskiy State University of Telecommunications and Informatics, Samara, Russia
E-mail: 89061250055@mail.ru
ORCID: https://orcid.org/0009-0003-1598-8971

Yakupov Denis Olegovich
Povolzhskiy State University of Telecommunications and Informatics, Samara, Russia
E-mail: d.yakupov@psuti.ru
ORCID: https://orcid.org/0009-0003-2371-0822
RSCI: https://elibrary.ru/author_profile.asp?id=1175874

Adaptive encryption of text data for natural language processing systems based on large neural network models

Abstract. Modern enterprises are actively implementing artificial intelligence and automated natural language processing systems to optimize work with corporate information. However, the use of these technologies carries significant risks of data privacy violations, especially when analyzing unstructured text materials. Existing cryptographic methods of information protection demonstrate limited effectiveness in processing linguistic data, and modern approaches to secure information processing have significant disadvantages: homomorphic encryption methods significantly reduce computing speed, and distributed learning technologies face problems with matching heterogeneous data.

In the presented study, a comprehensive information security method has been developed that combines the advantages of homomorphic encryption and distributed learning. The main scientific achievement of the work is the creation of specialized cryptographic algorithms for processing language data, including the identification of named objects in encrypted form, as well as optimizing the ratio between the level of protection and computational efficiency. The proposed methodology includes differentiated encryption of various types of data, decentralized information processing with automatic adjustment of learning parameters, as well as mechanisms to ensure anonymity and access control.

The experimental part of the study evaluates three key parameters: the system's resilience to potential attacks, the processing speed of protected data, and the accuracy of analytical algorithms. The practical value of the work lies in creating a secure platform for the implementation of intelligent data analysis systems in the financial sector, healthcare and telecommunications industries, where information leakage can lead to serious consequences. A promising area of further research is the development of cryptographic algorithms that are resistant to quantum hacking methods.

Keywords: adaptive encryption; homomorphic encryption; federated learning; data privacy; tokenization; corporate databases; big data