УДК: 004.89

Худайберидева Г. Б., магистр, ассистент кафедры
«Информатика и информационные технологии»
Московский Политехнический Университет,
Россия, г. Москва

Кожухов Д. А., магистр, ассистент кафедры
«Информатика и информационные технологии»
Московский Политехнический Университет,
Россия, г. Москва

Пименкова А. А., студент-бакалавр кафедры
«Информатика и информационные технологии»
Московский Политехнический Университет,
Россия, г. Москва

КВАНТОВО-ЗАЩИЩЕННЫЕ ИИ-МОДЕЛИ ДЛЯ БЕЗОПАСНОЙ СОВМЕСТНОЙ РАЗРАБОТКИ В УСЛОВИЯХ ГЕОПОЛИТИЧЕСКОЙ НЕСТАБИЛЬНОСТИ

***Аннотация**: В работе рассматривается актуальная проблема обеспечения безопасности интеллектуальной собственности при распределенной разработке искусственного интеллекта в условиях обострения геополитической напряженности. Предлагается концепция архитектуры ИИ-систем, интегрирующей методы квантовой криптографии, в частности квантовое распределение ключей (QKD), для защиты критически важных этапов workflow совместной разработки и эксплуатации моделей. Инновационность подхода заключается в создании сквозного защищенного контура, где ИИ-компоненты не только обрабатывают данные, но и обеспечивают их физическую безопасность на уровне передачи ключей шифрования, существенно повышая устойчивость к компрометации классических инфраструктурных элементов. Исследование выполнено с учетом специфики развития квантовых и ИИ-технологий в Российской Федерации и существующих вызовов в области*

информационной безопасности. Рассматриваются базовые принципы построения такой системы и потенциальные направления стандартизации.

***Ключевые слова**: квантовая криптография, квантовое распределение ключей (QKD), искусственный интеллект, безопасность ИИ, распределенная разработка, интеллектуальная собственность, геополитическая нестабильность, защита данных, федеративное обучение, безопасная передача моделей.*

Khudaiberideva G. B.
master and department assistant at the department of
"Computer Science and Information Technology"
Moscow Polytechnic University
Moscow, Russia

Kozhukhov D. A.
master and department assistant at the department of
"Computer Science and Information Technology"
Moscow Polytechnic University
Moscow, Russia

Pimenkova A. A.
bachelor's student at the department of
"Computer Science and Information Technology"
Moscow Polytechnic University
Moscow, Russia

QUANTUM-SECURE AI MODELS FOR SECURE COLLABORATIVE DEVELOPMENT IN THE CONTEXT OF GEOPOLITICAL INSTABILITY

***Annotation**:The paper considers the urgent problem of ensuring the security of intellectual property in the distributed development of artificial intelligence in the context of escalating geopolitical tensions. The concept of an AI system architecture integrating*

quantum cryptography methods, in particular quantum key distribution (QKD), is proposed to protect critical workflow stages of collaborative model development and operation. The innovative approach is to create an end-to-end secure circuit, where AI components not only process data, but also ensure their physical security at the encryption key transfer level, significantly increasing the resistance to compromise of classical infrastructure elements. The study was carried out taking into account the specifics of the development of quantum and AI technologies in the Russian Federation and the existing challenges in the field of information security. The basic principles of building such a system and potential directions of standardization are considered.

***Keywords**: quantum cryptography, quantum key distribution (QKD), artificial intelligence, AI security, distributed engineering, intellectual property, geopolitical instability, data protection, federated learning, secure model transfer.*

Введение

Современные разработки в области искусственного интеллекта характеризуются возрастающей сложностью моделей и объемов данных, необходимых для их обучения и функционирования. Это обуславливает потребность в распределенной разработке, привлечении экспертов и вычислительных ресурсов из различных географических локаций и организаций [1]. Однако данная тенденция сталкивается с серьезными вызовами информационной безопасности, особенно актуальными в контексте текущей геополитической нестабильности [2]. Основной риск заключается в потенциальной утечке или несанкционированном доступе к уникальным алгоритмам, обучающим наборам данных и параметрам моделей, представляющим собой ценную интеллектуальную собственность [3]. Традиционные криптографические методы, основанные на вычислительной сложности определенных математических задач, потенциально уязвимы перед атаками с использованием будущих квантовых компьютеров [4]. Кроме того, компрометация серверов или узлов связи в распределенной сети разработки может привести к перехвату ключей шифрования или самих защищаемых активов. В этой связи возникает потребность в принципиально новых подходах к обеспечению

безопасности жизненного цикла ИИ-моделей, особенно на этапах их совместной разработки, обновления и обмена между участниками распределенных проектов [5]. Квантовая криптография, в частности протоколы квантового распределения ключей, предлагает фундаментальный уровень защиты, основанный на законах квантовой физики [6]. Интеграция QKD в архитектуру систем разработки и эксплуатации ИИ представляет собой перспективное направление исследований для обеспечения долгосрочной безопасности интеллектуальной собственности в условиях глобальной нестабильности.

Угрозы безопасности интеллектуальной собственности в распределенной разработке ИИ

Процесс распределенной разработки ИИ включает несколько критически важных этапов, каждый из которых несет специфические риски для конфиденциальности интеллектуальной собственности. Обмен исходным кодом моделей, фрагментами обучающих данных, гиперпараметрами и промежуточными результатами между исследовательскими группами требует защищенных каналов связи [7]. Использование облачных платформ для обучения или вывода моделей создает риски компрометации данных при передаче и хранении на сторонних серверах [8]. Методы совместного обучения, такие как федеративное обучение, хотя и минимизируют прямой обмен сырыми данными, все же предполагают передачу обновлений модели (градиентов или параметров), которые могут быть подвержены атакам на восстановление исходных данных или реконструкцию самой модели [9, 10]. В условиях геополитической нестабильности угроза целевых атак со стороны государственных или квазигосударственных акторов на критическую инфраструктуру разработки ИИ возрастает многократно [2]. Такие акторы обладают значительными ресурсами для проведения изощренных атак, включая внедрение в цепочки поставок программного

обеспечения, эксплуатацию неизвестных уязвимостей (zero-day) и применение методов криптоанализа с использованием значительных вычислительных мощностей [11]. Компрометация одного узла в распределенной сети может привести к утечке ключей шифрования, используемых для защиты коммуникации всей группы разработчиков. Следовательно, обеспечение безопасности ключевого обмена является фундаментальной задачей для защиты конфиденциальности передаваемой информации, включая параметры ИИ-моделей и чувствительные данные.

Квантовая криптография как фундамент безопасности

Квантовое распределение ключей представляет собой метод безопасной генерации и обмена криптографическими ключами между двумя сторонами (традиционно обозначаемыми как Алиса и Боб) [6]. Безопасность QKD основывается не на вычислительной сложности, а на фундаментальных принципах квантовой механики: принципе неопределенности Гейзенберга и теореме о запрете клонирования. Информация в QKD передается с помощью квантовых состояний света, например, поляризации отдельных фотонов. Любая попытка перехвата (измерения) этих состояний злоумышленником (Евой) неизбежно вносит возмущения в квантовую систему, которые могут быть детектированы легитимными пользователями [12]. Это позволяет Алисе и Бобу гарантировать секретность сгенерированного ключа при условии, что уровень ошибок в канале не превышает определенного порога. Сгенерированные с помощью QKD ключи обладают безусловной (информационно-теоретической) секретностью, то есть их безопасность не может быть нарушена даже обладателем неограниченных вычислительных ресурсов, включая будущие квантовые компьютеры [13]. В Российской Федерации ведутся активные исследования и разработки в области квантовой криптографии, включая создание отечественных аппаратных и программных решений, а также развитие инфраструктуры квантовых сетей

[14]. Важным аспектом является соответствие разрабатываемых систем требованиям российских регуляторов в области защиты информации.

Архитектура квантово-защищенных ИИ-систем для совместной разработки

Предлагаемая архитектура квантово-защищенной системы для совместной разработки ИИ-моделей базируется на интеграции QKD-инфраструктуры в существующие или проектируемые workflow распределенной разработки. Ключевым элементом является создание защищенных квантовых каналов между основными узлами разработки: исследовательскими центрами, облачными провайдерами, хранилищами данных. Эти каналы используются исключительно для генерации и распределения симметричных сессионных ключей. Генерация ключей с помощью QKD осуществляется специализированными аппаратными модулями, подключенными к квантовым каналам связи (например, оптоволоконным линиям). Сгенерированные ключи передаются в доверенные криптографические модули (аппаратные или программно-аппаратные), расположенные на каждом узле сети. Криптографические модули отвечают за безопасное хранение ключей и их использование для шифрования/дешифрования данных, передаваемых по классическим каналам связи. Для защиты собственно процессов разработки ИИ предлагается применение следующих механизмов, усиленных квантовой защитой ключевого обмена. Защита передачи моделей и обновлений предполагает шифрование параметров моделей, градиентов или других передаваемых артефактов с использованием симметричных ключей, сгенерированных и распределенных посредством QKD. При организации федеративного обучения защищенные QKD-ключи используются для шифрования обновлений моделей, передаваемых от клиентов к центральному серверу агрегации и обратно, а также для аутентификации участников. Доступ к облачным ресурсам для обучения или инференса

защищается путем установления зашифрованных туннелей (например, на основе IPsec или TLS), где сессионные ключи периодически обновляются с использованием ключей, предоставленных QKD-системой. Обеспечение целостности и аутентичности критических компонентов системы и передаваемых данных достигается с помощью криптографических хэш-функций и механизмов электронной цифровой подписи, ключи для которых также могут защищаться или генерироваться с привлечением QKD-инфраструктуры. Важным аспектом архитектуры является модульность, позволяющая интегрировать QKD с различными платформами разработки ИИ и методами распределенного обучения. Архитектура должна предусматривать возможность использования как наземных (оптоволоконных) QKD-каналов, так и спутниковых систем для организации защищенных соединений на большие расстояния, что актуально для географически распределенных команд разработчиков на территории Российской Федерации и с международными партнерами при наличии соответствующих возможностей. Разработка и внедрение подобных систем требует учета требований российских стандартов в области защиты информации, таких как ГОСТ Р 56403-2015 по квантовой криптографии, а также стандартов на алгоритмы шифрования (например, ГОСТ 34.12-2015) и электронной подписи (ГОСТ Р 34.10-2012) [15].

Преимущества и ограничения подхода

Основным преимуществом интеграции QKD в системы разработки ИИ является достижение принципиально более высокого уровня защиты ключевого обмена. Гарантируется долгосрочная безопасность ключей, используемых для шифрования передаваемых моделей, данных и обновлений, что критически важно для сохранения конкурентоспособности и ценности интеллектуальной собственности в условиях гонки технологий. Физическая природа защиты в QKD позволяет обнаруживать факт попытки перехвата ключей, что невозможно при

использовании классических криптографических протоколов [12]. Это создает дополнительный сдерживающий фактор для потенциальных злоумышленников. Повышается общая устойчивость распределенной инфраструктуры разработки к компрометации отдельных классических серверов или узлов связи, так как секретные ключи генерируются и используются локально в доверенных криптографических модулях, а по сети передаются только квантовые состояния для генерации ключей или сами зашифрованные данные. Однако внедрение квантово-защищенных систем сопряжено с рядом технологических и организационных ограничений. Текущая стоимость QKD-оборудования и развертывания квантовых каналов связи остается высокой, что может ограничивать масштабируемость решений [16]. Дальность действия наземных QKD-систем ограничена потерями в оптическом волокне, что требует использования доверенных ретрансляторов или спутниковых систем для глобального покрытия. Интеграция QKD с существующими ИИ-платформами и workflow требует разработки специализированных интерфейсов и адаптации процессов. Существует необходимость в развитии нормативной базы и стандартов, детально регламентирующих использование QKD для защиты специфических ИИ-активов и процессов в рамках законодательства Российской Федерации. Обеспечение надежной работы квантовых каналов в условиях реальной инфраструктуры связи также представляет собой инженерную задачу.

Заключение

Распределенная разработка сложных ИИ-моделей в условиях геополитической нестабильности требует принципиально новых подходов к обеспечению безопасности интеллектуальной собственности. Интеграция квантовой криптографии, в частности протоколов квантового распределения ключей, в архитектуру ИИ-систем предлагает путь к созданию сквозного защищенного контура на физическом уровне.

Предложенная архитектура фокусируется на защите критических этапов workflow совместной разработки путем обеспечения безусловно секретного обмена ключами шифрования между узлами сети. Использование ключей, сгенерированных посредством QKD, для защиты передач параметров моделей, обновлений в федеративном обучении и доступа к облачным ресурсам позволяет существенно повысить устойчивость к компрометации классических элементов инфраструктуры и обеспечить долгосрочную безопасность перед лицом угроз со стороны квантовых компьютеров. Несмотря на существующие технологические и экономические ограничения, развитие отечественных QKD-решений и квантовых сетей в Российской Федерации создает основу для практической реализации подобных систем. Дальнейшие исследования должны быть направлены на оптимизацию архитектуры для специфических сценариев разработки ИИ, разработку стандартов взаимодействия QKD-систем с платформами машинного обучения и глубокого обучения, а также на оценку эффективности и экономической целесообразности развертывания в различных контекстах. Успешная реализация концепции квантово-защищенных ИИ-систем для совместной разработки способна стать значимым фактором технологического суверенитета и конкурентного преимущества Российской Федерации в стратегически важной области искусственного интеллекта.

СПИСОК ЛИТЕРАТУРЫ:

1. Гудков В.А., Смирнов А.Н. Распределенные вычисления в задачах машинного обучения // Информатика и ее применения. 2021. Т. 15, № 4. С. 94–103.
2. Петренко С.А., Курбаков Д.А. Геополитические аспекты информационной безопасности в цифровую эпоху // Вопросы кибербезопасности. 2022. № 3(47). С. 45–56.

3. Иванов М.П. Защита интеллектуальной собственности в сфере искусственного интеллекта: правовые и технические аспекты // Право и экономика. 2023. № 5. С. 67–75.
4. Шабанов Б.М. Постквантовая криптография: состояние и перспективы // Прикладная криптография. 2021. № 1(12). С. 22–30.
5. Королев А.В. Безопасность данных в распределенных системах машинного обучения // Системы высокой доступности. 2022. Т. 18, № 3. С. 78–86.
6. Gisin N., et al. Quantum cryptography // Reviews of Modern Physics. 2002. Vol. 74, Iss. 1. P. 145–195.
7. Белов А.И., Козлов Д.С. Угрозы безопасности при совместной разработке программного обеспечения // Информационные технологии. 2020. Т. 26, № 9. С. 546–552.
8. Subramanian N., et al. Security Challenges in Cloud-Based Machine Learning // ACM Computing Surveys. 2023. Vol. 55, Iss. 10. Article 215.
9. Bonawitz K., et al. Towards Federated Learning at Scale: System Design // Proceedings of Machine Learning and Systems (MLSys). 2019. Vol. 1. P. 374–388.
10. Melis L., et al. Exploiting Unintended Feature Leakage in Collaborative Learning // IEEE Symposium on Security and Privacy (SP). 2019. P. 691–706.
11. Шульга А.В. Современные угрозы информационной безопасности национального уровня // Национальная безопасность. 2023. № 1(58). С. 34–42.
12. Scarani V., et al. The security of practical quantum key distribution // Reviews of Modern Physics. 2009. Vol. 81, Iss. 3. P. 1301–1350.
13. Renner R. Security of quantum key distribution // International Journal of Quantum Information. 2008. Vol. 6, No. 01. P. 1–127.
14. Национальная квантовая лаборатория [Электронный ресурс]. URL: https://quantum.ru/

15. ГОСТ Р 56403-2015. Защита информации. Квантово-криптографические системы распределения ключей. Требования безопасности. М.: Стандартинформ, 2015.
16. Панков А.Б. Экономические аспекты внедрения квантовой криптографии // Экономика и управление. 2022. № 5(187). С. 112–118.
17. Криптософт: Квантовые коммуникации [Электронный ресурс]. URL: https://cryptosoft.ru/solutions/quantum-communications/.