САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

На правах рукописи

КИЯМОВ Жасур Уткирович

О ПРОБЛЕМАХ ОПТИМИЗАЦИИ И БЕЗОПАСНОСТИ ДЛЯ МНОГОУРОВНЕВОЙ ВИРТУАЛЬНОЙ СЕТИ

Научная специальность 1.2.2. Математическое моделирование,
численные методы и комплексы программ

ДИССЕРТАЦИЯ
на соискание ученой степени
кандидата технических наук

Научный руководитель:
доктор физико-математических наук, профессор,
Богданов Александр Владимирович

Санкт-Петербург
2023

Содержание

ВВЕДЕНИЕ

Прогресс информационных технологий неразрывно связан с широким применением вычислительно интенсивных приложений в различных сферах деятельности. Приложения, которые требуют значительных вычислительных ресурсов, играют ключевую роль как в научных исследованиях, так и в передовых промышленных отраслях, включая авиастроение, судостроение, биотехнологию, фармацевтику, генетику и другие.

В научных исследованиях использование вычислительно интенсивных приложений стало неотъемлемой частью работы ученых. Они применяются для моделирования сложных физических и биологических процессов, анализа больших объемов данных, предсказания результатов экспериментов и разработки новых теорий. Эти приложения позволяют расширить границы наших знаний и углубить понимание фундаментальных научных принципов.

В промышленности вычислительно интенсивные приложения находят широкое применение в тех передовых отраслях, где требуется высокая степень точности, сложного моделирования и оптимизации. В авиастроении они используются для разработки новых самолетов, прогнозирования и анализа аэродинамических характеристик. В судостроении они помогают оптимизировать дизайн кораблей, улучшить их маневренность и экономичность. В биотехнологии и фармацевтике вычислительные приложения играют решающую роль в разработке новых лекарственных препаратов и моделировании биологических процессов. В генетике они помогают анализировать геномы, исследовать генетические болезни и проводить персонализированную медицину.

Использование вычислительно интенсивных приложений в этих передовых сферах деятельности способствует прогрессу, оптимизации процессов и достижению новых результатов. Они помогают ускорить и улучшить научные

исследования, повысить качество и эффективность промышленных разработок, а также способствуют инновациям и развитию важных отраслей экономики.

В условиях существенных изменений технологической платформы и перехода в новую эпоху, требуется разработка новых подходов к построению информационных систем будущего. В сущности, большинство таких систем являются распределенными, и с ростом их значимости возникают проблемы безопасности. В последнее время блокчейн технология представляется как одна из возможностей для решения этих проблем. Учитывая важность блокчейна в контексте мобильных сетей 6G, необходимо исследовать различные возможности этой технологии и проблемы, связанные с ее внедрением. На данный момент не существует всестороннего обзора, который бы охватывал ключевые аспекты роли блокчейна в 6G. Тем более важно создать такой обзор, учитывая как технические особенности 6G, так и приложения, и варианты использования, предусмотренные этой новой эпохой.

Важно изучить основные идеи, области применения, требования и основные технологии, которые формируют экосистему 6G, чтобы понять тенденции, способствующие развитию будущих приложений 6G, и определить четкие требования к этой новой сети, включая ключевые поддерживающие технологии. В контексте этого исследования блокчейн является одной из ключевых технологий, которая привлекает особое внимание.

Среди многих аспектов применения блокчейна важно выделить высокоуровневое представление о его роли в тенденциях и приложениях 6G, учитывая наблюдаемые тенденции и предполагаемые требования, необходимо понять, что блокчейн может предложить для 6G. Это создает основу для максимально детального изучения использования блокчейна в экосистеме 6G.

Сегодня ясно, что большие распределенные системы наилучшим образом могут быть построены на основе облачных технологий. Облачные вычисления представляют собой модель вычислений, где ресурсы, такие как вычислительная мощность, хранилища данных, сети и программное обеспечение, предоставляются как службы через интернет для удаленного пользователя. Они обеспечивают

доступность выделения ресурсов, динамическую масштабируемость и практически безграничные возможности для решения различных задач. Технология облачных вычислений имеет ряд преимуществ, включая высокую производительность, снижение расходов, высокую доступность и простую масштабируемость.

Однако, при ее практической реализации возникает целый ряд еще нерешенных научных проблем, препятствующих полноценному использованию всех потенциальных достоинств такого подхода. При практической разработке этих технологий необходимо решить целый ряд технических проблем, из которых выделим:

Во-первых, в процессе создания универсальной облачной системы возникает потребность в работе в гетерогенной среде и обеспечении доступа пользователей к их индивидуальным приложениям без ущерба для производительности. Это означает, что система должна быть способна поддерживать различные платформы, языки программирования и программные среды, чтобы пользователи могли работать с приложениями, разработанными под их конкретные требования. При этом необходимо обеспечить высокую производительность и отзывчивость системы, чтобы пользователи могли эффективно выполнять свои задачи.

Во-вторых, безопасность и надежность хранения индивидуальных данных в облачных средах и организация доступа для множества пользователей представляют значительную проблему. Безопасность является одной из важных задач облачных вычислений, так как она влияет на всю систему. Важно обеспечить безопасность программных интерфейсов, используемых для управления ресурсами, виртуальными машинами и сервисами. Эти интерфейсы должны обеспечивать аутентификацию и авторизацию пользователей, а также шифрование данных для защиты от несанкционированного доступа. Кроме того, система должна предоставлять удобный и единообразный авторизованный доступ к ресурсам, учет использования ресурсов и защиту от несанкционированного использования данных и ресурсов.

В-третьих, для практического использования гетерогенной облачной среды в различных областях требуется разработка универсальной системы запуска индивидуальных приложений. Эта система должна обеспечивать возможность запуска и выполнения приложений в гетерогенной облачной среде, где различные пользователи могут работать со своими собственными приложениями без потери производительности. Такая система должна быть гибкой и масштабируемой, чтобы поддерживать разнообразные требования пользователей и эффективно использовать вычислительные ресурсы. Она также должна обеспечивать управление ресурсами и мониторинг производительности для обеспечения оптимальной работы приложений в гетерогенной среде.

Цель исследования. Главной целью данного исследования является повышение эффективности распределенных вычислений в облачной системе. Для достижения этой цели предлагается создание операционного окружения, которое обеспечивает безопасный доступ пользователей к вычислительным ресурсам, а также разработку принципов запуска ресурсоемких приложений в распределенной вычислительной среде на основе технологии облачных вычислений.

Для достижения поставленной цели необходимо решение следующих задач:

1. Создание виртуальных машин, которые реализуют операционное окружение системы блокчейна, включая специальное программное обеспечение. Это позволит создать надежное и безопасное окружение для выполнения вычислений в облачной системе.
2. Разработка методологии запуска приложений в многоуровневых виртуальных средах. Эта методология должна оптимизировать использование ресурсов и повысить производительность гетерогенных программно-аппаратных комплексов, что приведет к более эффективным вычислениям в облачной среде.
3. Разработка подхода к построению операционного окружения пользовательской подсистемы, который обеспечивает безопасный доступ пользователей к ресурсоемким приложениям в гетерогенной распределенной облачной вычислительной среде. Этот подход должен

обеспечить защиту данных и ресурсов от несанкционированного использования.

4. Исследование методов повышения надежности аутентификации и авторизации и разработка методики их применения в гетерогенной облачной среде. Это позволит обеспечить высокий уровень безопасности и предотвратить несанкционированный доступ к ресурсам и данным в облачной системе.

Все эти задачи будут направлены на создание оптимальной и эффективной распределенной вычислительной среды в облачной системе, которая обеспечивает безопасный доступ пользователей к вычислительным ресурсам и повышает производительность при выполнении ресурсоемких приложений.

Предмет исследования. Данное исследование сосредоточено на изучении методов теоретического анализа и экспериментального исследования, связанных с организацией системы доступа пользователей к распределенной вычислительной среде на основе технологии облачных вычислений. Основной упор делается на следующие аспекты:

1. Методы авторизации пользователей по принципу одного окна. Исследуется возможность создания единого механизма авторизации, который позволяет пользователям получать доступ к различным ресурсам в облачной среде без необходимости повторной аутентификации. Это способствует повышению удобства использования системы и уменьшению риска возникновения уязвимостей при множественной аутентификации.
2. Методы построения облачной инфраструктуры с открытым исходным кодом. Исследуется возможность создания облачной инфраструктуры, основанной на открытых стандартах и программном обеспечении с открытым исходным кодом. Это позволяет снизить зависимость от конкретных поставщиков облачных услуг и дает возможность пользователям настраивать и изменять инфраструктуру под свои потребности.

3. Создание эффективной распределенной вычислительной среды на основе технологии облачных вычислений. Изучаются методы, позволяющие оптимизировать использование ресурсов и повысить производительность в распределенной вычислительной среде. Рассматриваются различные алгоритмы планирования ресурсов, механизмы балансировки нагрузки и оптимизации выполнения вычислений.
4. Методы интеграции и консолидации программных продуктов в распределенной вычислительной среде. Исследуются подходы, позволяющие интегрировать и объединять различные программные продукты в облачной среде, чтобы обеспечить их взаимодействие и совместную работу. Рассматриваются вопросы совместимости, стандартизации и интерфейсов для эффективного взаимодействия программных компонентов.

В ходе исследования осуществляется как теоретический анализ, так и проведение экспериментов, для более глубокого понимания этих методов и их применимости в реальных сценариях. Это позволяет улучшить практическую применимость и эффективность разрабатываемых подходов и создать основу для развития более совершенных и надежных систем распределенных вычислений в облачной среде.

Методы исследования. В данной исследовательской работе применяются современные методы, основанные на принципах параллельной и распределенной обработки информации, передачи данных в компьютерных системах, а также на принципах защиты вычислительных систем. Также в работе используются современные технологии проектирования программного обеспечения.

Для анализа и обеспечения надежности информационных систем исследование основывается на теории надежности информационных систем, теории случайных процессов и потоков. Эти методы позволяют провести анализ различных аспектов надежности информационных систем, включая оценку вероятности сбоев и отказов, оценку надежности системы в целом, а также определение оптимальных стратегий обнаружения и восстановления отказов.

Применение современных принципов и технологий позволяет провести исследование, которое способствует достижению поставленных целей работы. Это включает повышение эффективности и производительности информационных систем, обеспечение их надежности и безопасности, а также разработку оптимальных стратегий и методов управления данными и ресурсами.

Научная новизна работы.

1. В рамках данной исследовательской работы была разработана новая методика организации вычислительной системы с использованием многоуровневой виртуальной блокчейн сети. Эта методика направлена на повышение эффективности системы путем применения принципов и механизмов блокчейн технологии. Многоуровневая виртуальная блокчейн сеть позволяет эффективно распределить вычислительные ресурсы и обеспечить безопасность данных, создавая надежную и отказоустойчивую среду для выполнения вычислительных задач. Этот подход представляет собой инновационный способ организации вычислительных систем, который может значительно повысить их эффективность и надежность.

2. В результате данного исследования была разработана новая методика, основанная на существующих методах, которая направлена на повышение степени защиты данных и вычислений в виртуализованной блокчейн среде. Эта методика включает в себя многоуровневую систему защиты, которая обеспечивает надежность и безопасность операций внутри блокчейн сети.

Достоверность научных результатов и выводов подтверждена результатами тестирования алгоритмов и программного обеспечения, а также практическим использованием разработанных алгоритмических и программных методов и средств на действующем программно-аппаратном комплексе факультета ПМ-ПУ СПбГУ. Кроме того, достоверность научных результатов и выводов подтверждена апробацией результатов исследований на ряде научных конференций.

Научные положения, выносимые на защиту.

1. Разработана новая методика и комплекс программ, основанный на этой методике, для создания операционной среды многоуровневой виртуальной

блокчейн сети. Этот подход позволяет значительно увеличить общую производительность гетерогенных программно-аппаратных комплексов. В среднем, производительность повышается на порядок благодаря адаптации архитектуры каждой индивидуальной виртуальной машины под конкретное пользовательское приложение. Таким образом, новая методика и комплекс программ обеспечивают более эффективное использование вычислительных ресурсов и оптимальное функционирование многоуровневой виртуальной блокчейн сети. Это представляет значительный прогресс в области повышения производительности гетерогенных программно-аппаратных комплексов и адаптации виртуальных машин к конкретным пользовательским приложениям.

2. Разработана методика для создания облачной вычислительной системы, которая способствует увеличению общей производительности. Это достигается путем виртуализации не только процессоров, но также памяти и сети обмена данными. Важной особенностью этой методики является динамическая балансировка и управление миграцией процессов, а не данных. Это позволяет оптимально использовать ресурсы системы, обеспечивать эффективное распределение нагрузки и повышать производительность в облачной среде. Таким образом, данная методика представляет собой важный шаг в области оптимизации облачных вычислений и повышения производительности системы.

3. Разработана методика, которая позволяет повысить степень защищенности данных и ресурсов путем внедрения многоуровневой системы защиты. Эта методика способствует обеспечению более высокой степени надежности системы в целом. Путем применения различных уровней защиты, включая аутентификацию, авторизацию, шифрование и другие меры безопасности, данная методика обеспечивает эффективную защиту данных и ресурсов от несанкционированного доступа.

Апробация работы. Основные результаты работы докладывались и обсуждались на национальных и международных научно-технических конференциях.

Публикации

1. Богданов, А. В. Цифровизация здравоохранения: что можно сделать уже сейчас / А. В. Богданов, Н. М. Залуцкая, Н. Л. Щеголева, Н. Р. Зайналов, Ж. У. Киямов, А. Г. Дик // ИНФОРМАЦИОННОЕ ОБЩЕСТВО. — 2022. — № 5. С. 58–70.

2. Bogdanov, A. A Multilayer Approach to the Security of Blockchain Networks of the Future / A. Bogdanov, A. Degtyarev, N. Shchegoleva, V. Korkhov, V. Khvatov, N. Zaynalov, J. Kiyamov, A. Dik // Computational Science and Its Applications — ICCSA 2022 Workshops. — 2022. — P. 205–216. («Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)»; Vol. 13377).

3. Shchegoleva, N. New Technologies for Storing and Transferring Personal Data / N. Shchegoleva, N. Zalutskaya, A. Dambaeva, J. Kiyamov, A. Dik ; O. Gervasi, B. Murgante, S. Misra, A. M. A. C. Rocha, C. Garau (Eds.) // Computational Science and Its Applications — ICCSA 2022. — Cham : Springer Nature, 2022. — Vol. 13380. — P. 680–692. (Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Vol. 13380 LNCS).

4. Bogdanov, A. Comparative analysis and applicability determination for several dlt solutions / A. Bogdanov, V. Korkhov, N. Shchegoleva, V. Khvatov, N. Zaynalov, J. Kiyamov, A. Faradzhov, A. Dik // Proceedings of the 9th International Conference "Distributed Computing and Grid Technologies in Science and Education". — 2021. — Vol. 3041. — P. 592–596. — (CEUR Workshop Proceedings).

5. Degtyarev, A. Risk Model of Application of Lifting Methods / A. Degtyarev, A. Bogdanov, N. Shchegoleva, A. Dik, V. Khvatov, N. Zaynalov, J. Kiyamov, A. Faradzhov // Proceedings of the 9th International Conference "Distributed Computing and Grid Technologies in Science and Education". — 2021. — Vol. 3041. — P. 369–374. — (CEUR Workshop Proceedings).

6. Degtyarev, A. Solving the Problems of Byzantine Generals Using Blockchain Technology / A. Degtyarev, A. Bogdanov, N. Shchegoleva, V. Korkhov, A. Dik,

J. Kiyamov, A. Faradzhov, N. Zaynalov // Proceedings of the 9th International Conference "Distributed Computing and Grid Technologies in Science and Education". — 2021. — Vol. 3041. — P. 573–578. — (CEUR Workshop Proceedings).

7. Bogdanov, A. Testing and Comparative Analysis of the F-BFT-based DLT Solution / A. Bogdanov, N. Shchegoleva, V. Korkhov, V. Khvatov, N. Zaynalov, J. Kiyamov, A. Dik, A. Faradzhov ; O. Gervasi, B. Murgante, S. Misra, C. Garau, I. Blečić, D. Taniar, B. O. Apduhan, A. M. Rocha, E. Tarantino, C. M. Torre (Eds.) // Computational Science and Its Applications — ICCSA 2021. — Cham : Springer Nature, 2022. — P. 31–41. — (Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Vol. 12952 LNCS).

8. Bogdanov, A. Protection of Personal Data Using Anonymization /, A. Bogdanov, A. Degtyarev, N. Shchegoleva, V. Korkhov, V. Khvatov, N. Zaynalov, J. Kiyamov, A. Dik, A. Faradzhov ; O. Gervasi, B. Murgante, S. Misra, C. Garau, I. Blečić, D. Taniar, B. O. Apduhan, A. M. Rocha, E. Tarantino, C. M. Torre (Eds.). — Computational Science and Its Applications — ICCSA 2021. — Cham : Springer Nature, 2021. — P. 447–459. — (Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Vol. 12956 LNCS).

9. Zaynalov, N. Hiding short message text in the uzbek language / N. Zaynalov, U. Narzullaev, A. Muhamadiev, O. Mavlonov, J. Kiyamov, D. Qilichev // 2020 International Conference on Information Science and Communications Technologies, (ICISCT). — Institute of Electrical and Electronics Engineers Inc., 2020. — 9351521.

10. Zaynalov, N., Information Security Issues for Travel Companies / N. Zaynalov, A. Mukhamadiev, B. Ugli, O. Mavlonov, J. Kiyamov, Q Dusmurod. ноя 2019, International Conference on Information Science and Communications Technologies: Applications, Trends and Opportunities (ICISCT 2019). — Institute of Electrical and Electronics Engineers Inc., 2019. — 9011896.

ГЛАВА 1. СОВРЕМЕННЫЙ ПОДХОД К ТЕХНОЛОГИЯМ КОРОТКОВОЛНОВЫХ ОБМЕНОВ ДАННЫХ В РАСПРЕДЕЛЕННЫХ РЕЕСТРАХ

1.1. Использование сетей следующего поколения в качестве распределенной сети

Как блокчейн может улучшить технические аспекты 6G: мобильные сети следующего поколения потребуют значительного улучшения существующих технических аспектов в текущем поколении, а также новых технических строительных блоков. Для каждого рассматриваемого технического аспекта определяются ключевые проблемы и исследуется безопасность данных блокчейна путем обеспечения децентрализованной и криптографической защиты базы данных, где каждый блок информации подтверждается и проверяется сетью участников. Кроме того, были предприняты усилия, чтобы справедливо представить, как плюсы, так и минусы использования блокчейна для рассматриваемых технических аспектов.

Мобильные системы 6G должны быть энергонезависимыми как на стороне инфраструктуры, так и на стороне устройства, чтобы обеспечить бесперебойную связь в любом уголке мира. Развитие возможностей сбора энергии продлит жизненный цикл как устройств сетевой инфраструктуры, так и конечных устройств, таких как устройства IoT [11,40].

Прогресс в области сенсорных технологий и их непосредственная интеграция с мобильными сетями, включая возможности связи с низким энергопотреблением, приведет к созданию передовых сетей 6G, объединяющих связь, обнаружение, управление, локализацию и вычисления в одной системе [63]. Благодаря такой

интеграции сеть 6G будет в состоянии предоставлять услуги зондирования и локализации, а также расширенные функции связи и вычислений. [55].

Обычно устройства IoT потребляют больше энергии для связи, чем для распознавания и обработки данных, но разработка механизмов связи со сверхнизким энергопотреблением и эффективных механизмов сбора энергии может привести к созданию устройств IoT с нулевым потреблением энергии или энергоэффективных устройств. [59].

Развитие технологий беспроводной связи, включая схемы кодирования и антенные технологии, расширит использование доступного спектра и повысит объем передаваемой информации по существующим беспроводным каналам, обеспечивая более надежную передачу большего количества информационных битов. [31,36].

Блокчейн вышел за рамки своего первоначального использования в криптовалютах и все больше становится перспективной технологией в других отраслях, таких как управление цепочками поставок, здравоохранение, интеллектуальное производство, образование, а также другие виды бизнеса и торговли. [3,68]. Блокчейн обладает огромным потенциалом для трансформации способов выполнения одноранговых транзакций, управления журналами, ведения записей, проведения децентрализованных переговоров, заключения торговых соглашений, проведения аудита и соблюдения нормативных требований, разрешения споров, управления доступом и безопасной автоматизации в различных секторах. Следуя общей тенденции, которую мы уже наблюдали в развитии 5G, можно уверенно предположить, что 6G будет включать все более программные, виртуализированные, интеллектуальные и программируемые системы. [12,70]. Тем не менее интересно отметить следующее: с одной стороны, такие концепции, как программирование, виртуализация и программируемость мобильных сетей следующего поколения, приведут к огромным преимуществам, таким как гибкое и открытое управление сетью и оркестровка, мультисервисы и микросервисы, создание виртуальных сетей по запросу. С другой стороны, эти концепции, как правило, усугубляют такие проблемы, как уязвимость системы

безопасности, утечка конфиденциальной информации, несанкционированный доступ к пользовательским данным, подверженное спорам мягкое совместное использование спектра, поддельные или поврежденные сетевые функции программного обеспечения и API-интерфейсы управления, незаконное использование ресурсов и неспособность обеспечить дифференциальную безопасность для дифференцированных услуг [65,69]. Блокчейн может играть важную роль в решении этих проблем. Например, в сетях 6G могут быть доступны программные решения с открытым исходным кодом для реализации различных сетевых функций на разных уровнях, включая ядро, транспорт и доступ. Это может сократить уязвимости системы безопасности, снизить риски утечки конфиденциальной информации и несанкционированного доступа к данным пользователей, а также обеспечить более дифференцированный уровень безопасности для различных типов услуг [56,66]. Использование блокчейна в открытой среде может гарантировать корректное функционирование, управление версиями, целостность и общую безопасность (Рисунок 1), и управление доверием

Рисунок 1 - Использование блокчейна для улучшения технических аспектов универсальной системы связи, предусмотренной 6G

для программного обеспечения, доступного для развертывания в сетях 6G. Это может помочь улучшить технические аспекты мобильных сетей 6G и расширить возможности применения и реализации 6G в различных областях, приложениях и сценариях использования [15].

Некоторые препятствия, существующие в создании парадигмы 3D-сетей, заключаются в следующем:

- Сеть блокчейн может предоставить эффективное решение для управления данными и обеспечения безопасности в гетерогенных сетях в тех случаях, когда существует проблема функциональной совместимости. Блокчейн позволяет децентрализованно управлять версиями и целостностью данных, а также обеспечивать прозрачность и надежность в процессе передачи данных между различными устройствами и системами.
- Добавление высоты в качестве нового измерения может привести к новым уязвимостям, так как злоумышленники могут использовать это для получения доступа к ключам безопасности в распределенной сети.

Таким образом, блокчейн предлагается использовать для обеспечения безопасного и децентрализованного управления гетерогенными сетями, которые расположены на земле, в воздухе и в космосе. Благодаря использованию четко определенных смарт-контрактов и блокчейн-технологии можно создать децентрализованное решение для управления ключами безопасности, аутентификации, авторизации и возможности аудита, а также для обеспечения совместимости между различными типами сетей. Это позволит обеспечить высокий уровень безопасности и защиты от кибератак в интегрированных гетерогенных сетях.

1.2. Управление спектром с динамическим распределением ресурсов

Радиочастотный спектр является дефицитным и незаменимым ресурсом для мобильной связи, его общее управление (распределение, обнаружение и совместное использование) должно выполняться эффективно в сетях 6G.

Традиционный подход к статическому распределению спектра, основанный на политике фиксированного доступа к спектру (ФДС), который используется регулирующими органами связи, может обеспечивать законное использование спектра, но может также приводить к недостаточному использованию выделенного спектра [37]. В последние годы появились различные механизмы динамического управления спектром (ДУС), такие как динамический доступ к спектру (ДДС), лицензированный общий спектр (ЛОС) и система доступа к спектру (СДС) [22]. В целом, система ДУС включает два типа пользователей: первичный пользователь и вторичный пользователь. Основной пользователь — это авторизованный субъект с исключительным правом на выделенный спектр. Он может не иметь возможности полностью использовать спектр все время. Вторичный пользователь — это тот, кто получает доступ к неиспользуемому спектру, совместно используемому основным пользователем. ДУС позволяет вторичным пользовательским устройствам постоянно оценивать среду радиочастотного спектра и автоматически регулировать рабочие частоты, чтобы приспособиться к условиям пропускной способности [17].

Появление на рынке новых операторов мобильной связи приводит к увеличению конкуренции и необходимости более эффективно использовать спектр частот для обеспечения высококачественных услуг. В этом контексте динамическое распределение спектра между операторами может быть одним из способов оптимизации использования ресурсов и обеспечения удовлетворительного уровня обслуживания для конечных пользователей.

Основные проблемы:

- Традиционный (централизованный) способ распределения спектра и управления им сопряжен с многочисленными проблемами. Это высокие административные расходы, уязвимости в системе безопасности, утечка информации о конфиденциальности и трудности с обеспечением прозрачного доступа сети пользователей.
- Аналоговые методы передачи данных для пользователей имеют свои недостатки, такие как ограниченный доступ и возможность атак типа DoS,

которые могут вызвать перегрузку сети и снижение ее производительности.

- В случае совместного использования частных сетей основной проблемой является управление лицензией на приоритетный доступ и общий авторизованный доступ пользователей, чтобы обеспечить их сосуществование путем относительного совместного использования спектра [10]. Развитие выделенной системы доступа SAS будет критически важным для 6G, поскольку предполагается, что частные сети будут продолжать существовать. Однако создание такой системы будет сложной задачей, требующей значительных усилий в разработке и внедрении.

Блокчейн стал многообещающей технологией для динамического распределения доступа и общего управления, в том числе для проведения аукционов и расчетов платежей. Федеральная комиссия по связи уже признала потенциал блокчейна для эффективного управления использованием спектра [51,64]. Использование технологии блокчейн может значительно улучшить мониторинг использования спектра и упростить процесс аудита, что сделает более прозрачным и эффективным внедрение правил совместного использования между местными операторами мобильной связи [58].

Использование распределения по участию нескольких организаций с общими правами доступа к использованию спектра отличается от традиционного способа. Здесь пользователи могут обмениваться правами на использование спектра, чтобы удовлетворить свои потребности. Блокчейн может быть использован для создания децентрализованных торговых платформ с открытым доступом к спектру. Однако открытая торговля ресурсами спектра, основанная на блокчейне, может привести к конфиденциальным проблемам, таким как раскрытие торговых моделей или утечка заявок, сделанных при распределении спектра на основе аукционов. Различные механизмы защиты конфиденциальности, основанные на блокчейне и смарт-контрактах, могут использоваться для защиты торговых организаций от таких угроз конфиденциальности. Чтобы реализовать

такой подход, можно воспользоваться схемой двойного аукциона на платформе блокчейна с использованием смарт-контрактов. Для обеспечения конфиденциальности пользователей во время торгов используется дифференциальная конфиденциальность в дополнение к симметричному шифрованию. Для определения победителя в схеме двойного аукциона используется целочисленное линейное программирование. Чтобы решить проблемы конфиденциальности, административных издержек и единой точки отказа, связанные с традиционной службой широкополосного радиовещания, была предложена распределенная модель с использованием блокчейна. Схема, первоначально предложенная Федеральной комиссией по связи, рассматривается как потенциальный кандидат на соответствие требованиям высокой спектральной эффективности для 6G в качестве нового алгоритма консенсуса, который не только легковесен, но и интегрируется с процессом распределения спектра, т. е. система достигает консенсуса, оценивая стратегию распределения спектра. Кроме того, использование метода кольцевой подписи защищает конфиденциальность, гарантируя отсутствие связи между реальной личностью пользователей и их псевдонимами.

С использованием блокчейна можно реализовать двухуровневую иерархическую архитектуру системы управления ресурсами, которая может быть дополнена искусственным интеллектом. В частности, она демонстрирует использование блокчейна для улучшения функции регулирования и учета финансовых расчетов. Предполагается, что ИИ обеспечивает распознавание шаблонов использования с помощью глубокого обучения с подкреплением и интеллектуального принятия решений. Кроме того, использование иерархического блокчейна снижает административные расходы, уменьшает накладные расходы на вычисления и хранение и сводит к минимуму сложность алгоритмов консенсуса. Тем не менее, накладные расходы блокчейна и методы получения обучающих данных для моделей ИИ с полным соблюдением конфиденциальности по-прежнему сложны.

Блокчейн может обеспечит быстрые финансовые соглашения для совместного использования реальном времени между различными операторами инфраструктуры и между операторами мобильной связи [60]. В частности, токенизация спектра в виде новой виртуальной монеты. Таким образом, их использование позволяет отказаться от финансовых (центральных) бирж и ускорить финансовые расчеты. чтобы максимизировать прибыль, минимизировать убытки и обеспечить безубыточную торговлю.

Использование блокчейн помогает пользователем связи "человек-человек" (H2H) взаимодействовать с мобильными сетями для удовлетворения огромных потребностей в подключении для связи M2M. Архитектура, поддерживаемая блокчейном, обеспечивает конфиденциальность в отношении стоимости совместного использования, обеспечивает оптимизированные стимулы для пользователей H2H через использование смарт-контрактов и оптимизирует распределение реестра для устройств M2M путем разработки консенсуса. В целом, этот метод может быть полезен для динамического управления спектром в будущей конвергенции гетерогенных сетей с трехмерными сетями.

В контексте операторов мобильной связи блокчейн был предложен как услуга для автоматизации безопасного управления и использования спектра между операторами мобильной связи и местными операторами. Такая услуга может обеспечивать:

- выбор владельцев спектра;
- публичный доступ на вторичном рынок;
- прямые платежи P2P за совместное использование и динамическое создание соглашения с использованием смарт-контрактов. [42].

Безопасность всегда была важным свойством телекоммуникаций и сетей. Начиная с отсутствия механизма безопасности и ограничений в сети 1G, аутентификация, анонимность и безопасность на основе шифрования были введены в 2G [30]. В 3G функции безопасности 2G были улучшены за счет введения аутентификации и соглашения о ключе, двусторонней аутентификации, а также

безопасности радиоинтерфейса 3GPP и сетевой аутентификации пользователей. С 4G Enhanced Packet System-AKA (EPS AKA) механизмом доверия и передачей обслуживания, введено управление ключами. 5G включает в себя множество расширенных функций безопасности, таких как улучшенные возможности взаимной аутентификации, скрытый идентификатор подписки, расширяемый протокол аутентификации (EAP-AKA) [28] и 5G механизм идентификации на основе интегрированной схемы шифрования на эллиптических кривых. [26]. В будущих сетях 6G может возникнуть множество проблем безопасности из-за чрезвычайно большой сети гетерогенных сетей и ожидаемых чрезвычайно высоких требований к надежности. Чтобы помешать более изощренным противникам, потребуются передовые и интеллектуальные механизмы безопасности [47,48]. Некоторые из проблем безопасности были определены в предполагаемых сетях 6G:

- Конфиденциальность и целостность будут сложными, поскольку будущая инфраструктура 6G может создавать огромные поверхности угроз с беспроводным подключением и огромным объемом данных, генерируемых в сети.
- Непрерывная доступность услуг будет еще одной проблемой, поскольку более широкая поверхность угроз и широкие возможности подключения повысят риск распределенных атак типа «отказ в обслуживании» (DDoS) [30].
- Механизмы аутентификации и управления доступом должны быть расширены, чтобы соответствовать диверсификации 6G сети, которые являются ресурсоемкими и могут создавать уязвимости данных в связанных услугах. В частности, централизованный способ управления доступом создает серьезные проблемы при проектировании будущих сетей.
- Аудит будет еще одним сложным аспектом безопасности из-за требований к оценке огромного количества больших данных (например, управление сегментами сети между несколькими узлами).

- С появлением искусственного интеллекта в 6G атаки безопасности на основе AI/ML также могут происходить в сетях 6G. Напротив, ИИ можно использовать как инструмент для обнаружения, прогнозирования и смягчения атак на безопасность. Таким образом, развертывание проактивных механизмов безопасности и обнаружение атак нулевым разглашением(ZKP) будут жизненно важными проблемами безопасности в 6G [28].

Блокчейн становится интригующим решением для обеспечения безопасности, подотчетности, наблюдения и управления мобильными сетями. На рисунке 2 приведены возможные атаки на систему безопасности, которые могут произойти в мобильных сетях следующего поколения, и их возможное смягчение с помощью блокчейна. С использованием технологий блокчейн, сети 6G могут противостоять угрозам перехвата благодаря свойствам неизменности, прозрачности, безотказности и распределенного доступа. Блокчейн создает

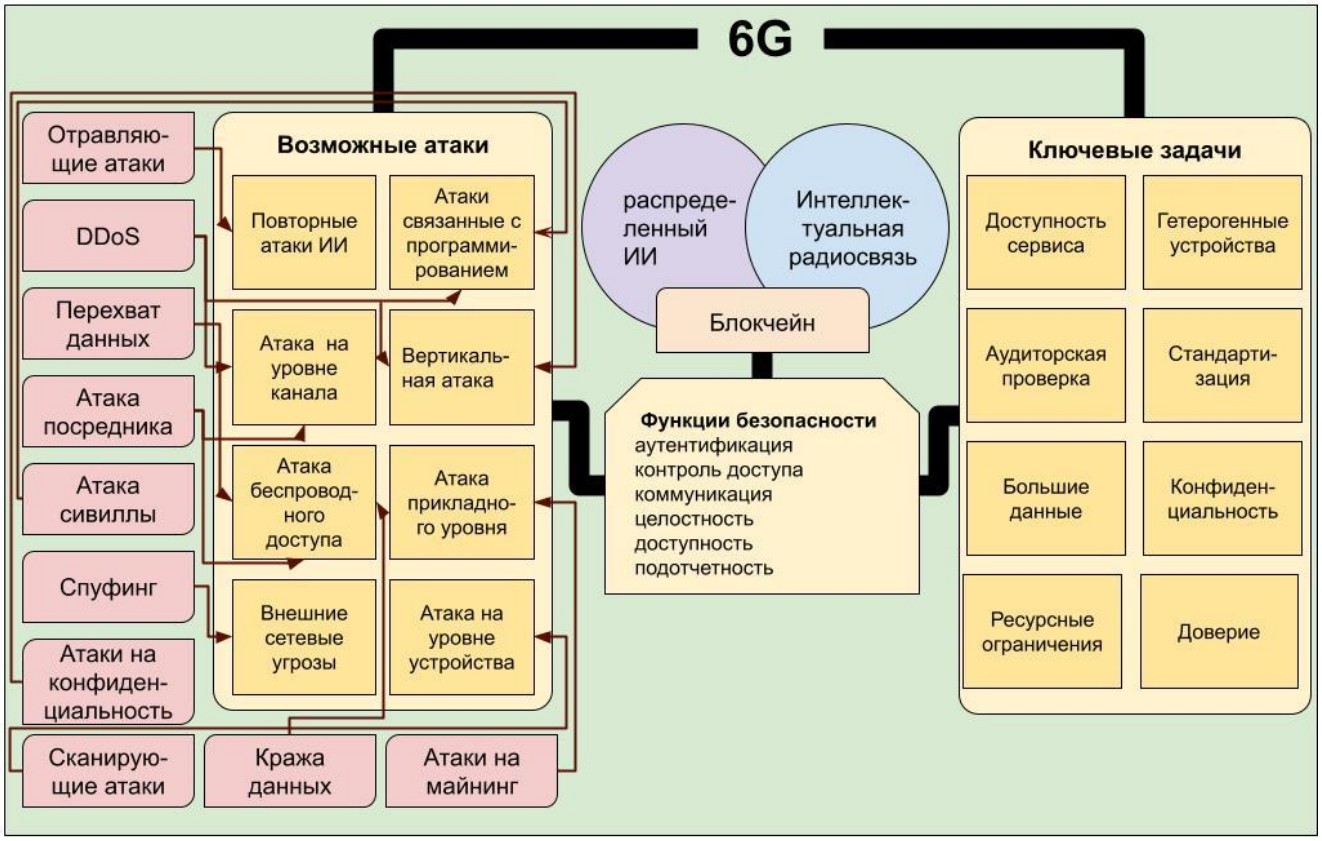

Рисунок 2 - Возможные атаки и проблемы безопасности в сетях 6G, а также функции безопасности, которые могут сопровождать блокчейн.

децентрализованные средства доверия между ненадежными пользователями, так как он может противостоять атакам Сивиллы [9,35,66]. Заметим, что при атаке Сивилла злоумышленник создает большое количество поддельных удостоверений в системе, основанной на одноранговой сети, и получает большее влияние из-за масштаба поддельных удостоверений. Однако, благодаря использованию эффективных протоколов консенсуса, таких как P-BFT, блокчейн обеспечивает уверенность в том, что только настоящие пользователи контролируют систему. [52,64]. Этот подход также сокращает вероятность модификации данных и атак на уровне канала, поскольку только участвующие узлы могут просматривать или добавлять новые транзакции. Благодаря развертыванию хорошо разработанных смарт-контрактов поверх блокчейна, можно решить ключевые аспекты безопасности, такие как аутентификация, контроль доступа и подотчетность.

Кроме того, как указано в [47], использование блокчейна позволяет безопасно обучать модели AI/ML на основе неизменяемых данных, хранящихся в блокчейне, как правило, в гетерогенном сетевом сценарии, укреплять доверие за счет прозрачного хранения связанных данных. Для систем на основе AI/ML, важно иметь доверенное выполнение, чтобы обеспечить безопасность в процессе принятия решений. Совместная оптимизация моделей ML может происходить путем безопасного обмена параметрами модели, без требования доверия между участниками. Смарт-контракты могут использоваться в качестве распределенного механизма для обеспечения безопасности в функциях AI/ML [71]. Например, в системах промышленного IoT с обученной моделью глубокого обучения, разработанный смарт-контракт может быть использован для обнаружения сетевых аномалий и вредоносного трафика. Это обеспечивает анализ сетевого трафика в безопасной распределенной среде, где модели AI/ML могут быть безопасно обучены на основе неизменяемых данных, хранящихся в блокчейне.

Безопасность будет иметь первостепенное значение для укрепления доверия между потенциальными узлами. Таким образом, на рисунке 2 показаны возможные атаки и проблемы безопасности в сетях 6G, а также функции безопасности, предлагаемые блокчейн. Обозначены два ключевых компонента сети 6G, а именно

распределенный ИИ и интеллектуальная радиосвязь, которые наиболее тесно связаны с применимостью блокчейна для обеспечения безопасности.

Атаки на разных уровнях в мобильных сетях следующего поколения 6G, включая граничный уровень, уровень приложений, уровень устройств, беспроводные радиоинтерфейсы, внешнюю сеть и различные вертикали, такие как автономные автомобили, промышленное управление и т. д. Каждый уровень имеет свои уникальные уязвимости и риски для безопасности, которые могут быть использованы злоумышленниками для атак на систему. Поэтому важно обеспечить полную защиту на всех уровнях и сделать мобильные сети 6G более устойчивыми к возможным атакам.

Блокчейн может использоваться для обеспечения аутентификации, авторизации и управления ключами в коммуникационных сетях [53,67,69]. Это общий канал связи, который может установить доверие между заинтересованными сторонами в сети 6G и способствовать их сотрудничеству на общей платформе, даже при возникновении неожиданных сбоев в сети. [27].

Основываясь на вышеупомянутых работах можно сделать вывод, что блокчейн можно использовать в качестве технологии обеспечения безопасности в сетях 6G для решения таких функций безопасности, как аутентификация, контроль доступа, целостность связи, доступность и подотчетность. Кроме того, решения на основе блокчейна могут быть сформированы для защиты сетей 6G от нескольких атак, как показано на рисунке 2. Однако включение таких решений безопасности на основе блокчейна в сети 6G по-прежнему будет проблематичным из-за увеличения неоднородности устройств, конвергенции разнородных сетей в сторону трехмерных сетей, роста больших данных и дополнительных вычислительных ресурсов, необходимых для блокчейна.

Трудности сохранения конфиденциальности при сборе, объединении и обработке данных возможны во многих областях применения 6G, таких как здравоохранение, защита окружающей среды, умный город. Следовательно, конфиденциальность будет важным требованием для сетей 6G. Кроме того,

необходимо обеспечить соответствие таких сетей законам о конфиденциальности ГОСТ Р 59407–2021.

1. Ключевые проблемы:
   - Несанкционированный доступ к виртуализированным сетевым ресурсам может создать высокий риск для конфиденциальности подключенных пользователей, таким образом, безопасный контроль доступа является одной из ключевых проблем, возникающих при сохранении конфиденциальности в обширных наборах данных в сетях 6G.
   - Появление крупномасштабных сенсорных систем с камерами и датчиками, непрерывно собирающими данные из окружающей среды, которые могут привести к нарушению конфиденциальности данных. Например, беспилотные летательные аппараты для наблюдения за толпой, которые собирают информацию о больших собраниях, могут представлять угрозу для частной жизни человека.
   - Из-за огромного количества данных в централизованных точках, таких как серверы агрегации данных, поставщики данных и поставщики услуг, может возникнуть риск утечки конфиденциальной информации [61]
   - Поскольку ИИ являются ключевыми технологиями, поддерживающими 6G, при сборе отдельных личных данных для моделей обучения риск раскрытия данных может привести к нарушению конфиденциальности.
2. Роль блокчейна.

Конфиденциальность — еще один технический аспект, который может быть реализован при использовании блокчейна в сетях 6G [24]. Наличие общего канала связи в виде блокчейна может позволить идентифицировать пользователей сети по псевдонимам вместо прямой идентификации личности [34]. Производительность сохранения конфиденциальности измеряется пользователем и уровнем ресурсов по

отношению к временной сложности и потреблению памяти. Это может обеспечить более высокую конфиденциальность и прозрачность во время сбора данных и управления воздушным движением, позволяя при этом быстро загружать данные с более высокой пропускной способностью и меньшей задержкой между приложениями.

В 6G концепция гиперинтеллектуальности предусматривает интеллектуальные и автономные сети, которые облегчат выполнение ключевых операционных функций [24]. Кроме того, сетевые сервисы и приложения будут сверхгибкими, настраиваемыми и адаптивными, чтобы обеспечить максимальное удобство для пользователей. Эта когнитивная сеть по своей сути нацелена на то, чтобы интегрировать ИИ в качестве родного элемента в будущие сети.

Основные проблемы повсеместного использования интеллекта в сетях 6G можно резюмировать следующим образом:

- Сложность и оптимизация для AI/ML. Разнообразие и сложность сетей 6G приведет к проблемам с точки зрения управления с участием человека и задач оптимизации системы. Решающее значение имеют автономные и самоуправляемые схемы управления и оптимизации. Однако сложность таких интеллектуальных схем и способы их оптимизации представляют собой непростые вопросы. Необходимо управлять общей сложностью и оптимизировать внедрение AI/ML.
- Распределенный характер интеллекта в 6G требует интегрированной работы разрозненных сторон для интеллектуальной инфраструктуры. Это также связано с многосвязной и многодоменной структурой программной и сервисной архитектуры, унаследованной от сетей 5G и расширенной в дальнейшем [16]. Однако надежность агентов AI/ML является нетривиальным атрибутом и может ухудшить производительность разведки в 6G.
- Безопасный обмен данными для всеобъемлющей аналитики. Функция централизованного хранения и управления в приложениях ИИ 6G может

подвергаться угрозам безопасности данных (например, кража или злонамеренное вмешательство из-за компрометации сервера), что приводит к нарушению конфиденциальности и нарушению нормальной работы сети. Это важная задача для крупномасштабных и многопользовательских сетевых систем, таких как 6G.

- Крупномасштабные приложения AI/ML: вездесущий интеллект для различных вариантов использования и приложений приведет к более децентрализованной архитектуре хранения и обмена данными в 6G. Это стремление также связано с кооперативными и федеративными службами и инфраструктурой, гибко связанными вместе для интеллектуальных иммерсионных сред и объектов. Важная проблема заключается в том, как разработать и интегрировать технологии, обеспечивающие такую масштабируемость для вездесущего интеллекта.

Поскольку повсеместное использование ИИ/МО будет реализовано в распределенной и крупномасштабной системе 6G для различных технических аспектов, включая управление сетью, предполагается, что распределенные методы ИИ/МО обеспечивают быстрый контроль и аналитику чрезвычайно большого объема генерируемых данных в сети 6G.

Блокчейн может обеспечить важные функции для повсеместного интеллекта в сетях 6G. Одним из таких аспектов является управление ИИ и обмен данными. Эффективному, надежному и безопасному управлению и совместному использованию результатов обучения препятствует неоднородность и недоверие между узлами. Каждый граничный узел имеет одинаковое количество нейронных клеток и одинаковое количество слоев в их нейронной сети. В реальных сценариях это не всегда возможно. Еще одной важной ролью блокчейна является поддержка контекстной осведомленности и автономности. Гиперинтеллектуальные сети 6G должны быть контекстно-зависимыми для управления сетью, а также для предоставления интеллектуальных цифровых услуг. Это возможно с функциями сверхмасштабного восприятия, которые будут поддерживать такую контекстно-

зависимую работу. Кроме того, они должны быть самоуправляемыми и автономными, чтобы соответствовать KPI производительности.

Блокчейн обеспечит надежное выполнение с помощью смарт-контрактов и совместного использования данных для понимания контекста и, следовательно, автономной работы. Однако сложность решений на основе блокчейна для этой цели может привести к обременению.

Сверхмасштабное зондирование — еще одна важная возможность, поскольку интеграция передовых сенсорных технологий с мобильными сетями в сочетании с возможностями связи с низким энергопотреблением приведет к повсеместному интеллектуальному зондированию и локализованным услугам [8]. Блокчейн также будет служить децентрализованным и заслуживающим доверия слоем данных для сверхмасштабного зондирования и анализа больших данных в 6G. Это важнейшая полезная функция блокчейна, поскольку датчики являются основой интеллектуальных сервисов и сетей. Архитектура безопасности с возможностью обнаружения и отслеживания объектов через блокчейн за счет развертывания алгоритмов искусственного интеллекта на пограничных серверах, в то время как сеть 5G обеспечивает низкую задержку и высокую надежность связи для обслуживания этих чувствительных ко времени приложений. Для сетей 6G ожидаемые услуги требуют очень строгих и стабильных характеристик QoE/QoS. Это возможно только благодаря очень эффективной и интеллектуальной структуре управления ресурсами в сетевом доступе.

Однако нехватка ресурсов и требования к эффективности делают этот вид операций затруднительным. Управление масштабными интеллектуальными поверхностями является особой проблемой управления в 6G. Использование программируемых интеллектуальных поверхностей для покрытия искусственных сооружений позволяет осуществлять интеллектуальное управление средой передачи в беспроводных системах [39]. Такие интеллектуальные поверхности будут играть важную роль в сетях 6G и будут способствовать ухудшению качества радиосигнала. В сети 6G, основанной на блокчейне, блокчейн обеспечивает ключевую возможность облегчения федеративного контроля между

разрозненными интеллектуальными поверхностями. Более того, интеллектуальные функции управления поверхностью из разных административных доменов (например, разных сетевых операторов) могут обмениваться информацией о синхронизации и реконфигурации с использованием блокчейна.

Блокчейн может служить средством реализации всепроникающих интеллектуальных сетей благодаря безопасному обмену данными, неизменности и децентрализованной архитектуре. Децентрализованная парадигма также обеспечивает более надежную интеллектуальную страту для всей сети, начиная от интеллектуальных устройств и заканчивая облачными сервисами. Он может упростить очень крупномасштабные федеративные приложения ИИ/МО для интеллектуальных и автономных цифровых подключенных услуг. Таким образом, в будущем 6G перспективен децентрализованный и безопасный метод обмена данными без так называемой доверенной третьей стороны или посредника, то есть посредством блокчейна. Тем не менее, накладные расходы, масштабируемость и проблемы функциональной совместимости блокчейна особенно очевидны на пути к достижению вездесущей интеллектуальной цели 6G. Например, потребление энергии из-за функций блокчейна, таких как консенсус и криптографические операции, является проблемой, ограничивающей полезность блокчейнов. Это явление свидетельствует о критической необходимости анализа энергетического и экологического воздействия повсеместного внедрения блокчейна в 6G, и поэтому необходимы исследования для его минимизации. Кроме того, необходимо улучшить аспект масштабируемости для реализации сверхмасштабных интеллектуальных решений. Также появляются новые атаки, которые пытаются снизить надежность и точность решений ИИ/МО с использованием блокчейнов, и исследовательскому сообществу необходимо более внимательно прислушиваться к ним в будущем [61].

Поскольку сети 6G будут сверхплотной сетью сетей и будут использовать более высокий частотный спектр (диапазон ГГц/ТГц), границы двухмерных сот сократятся до нескольких метров. Это резкое сокращение количества базовых станций для покрытия и область ожидания хэндоверов для мобильных

пользователей увеличится в геометрической прогрессии. Кроме того, из-за неоптимизированных механизмов передачи обслуживания для различных базовых станций с перекрытием и малым покрытием (таких как ABS Picocell, Femtocell и БПЛА) вероятность отказа передачи обслуживания, эффект вероятности и отказ радиоканала значительно увеличатся [31]. Впоследствии это отрицательно скажется как на пропускной способности, так и на задержке. С другой стороны, появление трехмерных сетей будет поддерживать трехмерную мобильность, что потребует новых методов управления мобильностью, поскольку передача обслуживания также будет происходить в вертикальном направлении [62]. Поскольку 6G стремится обеспечить сверхнадежную связь с малой задержкой, эффективное управление мобильностью имеет очень важное значение.

Вот некоторые из проблем, связанных с управлением мобильностью в сетях следующего поколения:

- Распределенное управление мобильностью вызывает множество вопросов к безопасности, такие как атака с использованием и внедрением псевдоданных, перехват сеанса, DoS-атака, атака «человек посередине» и атаки со стороны злоумышленников. Эти проблемы затрудняют использование цифрового мультиметра для управления мобильностью.
- Чтобы обеспечить безопасную передачу данных для защиты прошлого сеанса с доступом ключа существует угроза скомпрометированного узла привилегий. Помимо прямой секретности, одновременное выполнение других требований (например, взаимная аутентификация, согласование ключей, прослеживаемость и надежность) является серьезной проблемой [17].
- Использование централизованного сервера аутентификации для управления мобильностью приводит к проблемам как с производительностью, так и с безопасностью.
- Роуминговое мошенничество из-за неэффективного и подверженного задержкам обмена данными между домашней сетью и гостевой сетью

приводит к большим экономическим потерям. Такое мошенничество в роуминге может стать серьезной проблемой в будущих сетях 6G из-за уплотнения сетей и популярности локальных/частных сетей.

Сочетание блокчейна и смарт-контрактов может установить доверие между мобильными пользователями и мобильными сетями распределенным и анонимным образом. Различное использование блокчейна для эффективного управления мобильностью в беспроводных сетях обсуждается следующим образом: распределенное управление мобильностью направлено на преодоление проблем (например, единой точки отказа и не оптимальной маршрутизации), связанных с традиционным централизованным управлением мобильностью. Распределенное управление мобильностью страдает от проблем с безопасностью, таких как перехват сеанса, DoS-атака и атаки с помощью мобильных привязок и маршрутизаторов доступа. Эти проблемы в основном связаны с тем, что безопасность реализована централизованно. Хотя их схема показала себя многообещающей, однако высокая избыточность хранилища из-за нескольких распределенных реестров является сложной задачей.

Решение для аутентификации передачи обслуживания для сценария групповой связи в сетях, где необходимо управлять мобильностью транспортных средств: совокупный код аутентификации сообщений, одноразовый пароль и блокчейн, может уменьшить перегрузку при передаче. Предлагаемый протокол использует алгоритм Диффи-Хеллмана на эллиптических кривых для управления ключами сеанса. Протокол используется в блокчейне с управлением базой данных и маршрутизатор доступа к мобильности хранит ключевую информацию об аутентификации транспортных средств. В дополнение к выполнению требований безопасности, таких как конфиденциальность ключа сеанса и иммунитет к ответной атаке и атаке сивиллы, предложенная схема эффективна в контексте задержки передачи обслуживания и вычислительных и коммуникационных накладных расходах. Точно так же использование блокчейна возможно для аутентификации во время передачи обслуживания при защите конфиденциальности пользователя в сетях 5G на основе SDN. Поскольку 6G будет

использовать SDN и NVF для эффективного управления сетями, ожидается, что блокчейн может сыграть важную роль в управлении мобильностью.

Анонимная взаимная аутентификация по ключу имеет надежную отслеживаемость и совершенную прямую секретность с использованием хэш-функций блокчейна. Пользователь готовится к этапу передачи. Сначала он выбирает псевдо-идентификатор, вычисляет различные параметры, чтобы доказать легитимность его личности, и передает параметры аутентификации в точку доступа. Точка доступа проверяет актуальность информации для предотвращения повторных атак и проверяет легитимность параметров аутентификации. Авторы сравнили предложенный механизм с точки зрения размеров сообщений аутентификации и стоимости хранения и пришли к выводу, что оба параметра в предложенной схеме уменьшаются. Используя логику Burrow-Abadi-Needham (BAN) для проверки свойств безопасности, предлагаемой анонимной взаимной аутентификацией. Кроме того, для формальной проверки предложенного механизма аутентификации используются средства проверки моделей.

С тем же намерением обеспечить прямую секретность и удовлетворить требования легкости для мобильных устройств с ограниченными ресурсами, а также защитить конфиденциальность пользователей и устойчивость к взлому, предлагаемая схема использует хэш-функцию и сертификат пользователя, защищенный блокчейном, для выполнения аутентификации при передаче. Когда пользователь регистрируется на сервере аутентификации, последний генерирует уникальный сертификат и загружает его в блокчейн. Когда пользователь переходит на новую точку доступа, пользователь проходит аутентификацию путем перекрестной проверки с исходным сертификатом, доступным в блокчейне. Данная схема обеспечивает условную конфиденциальность, основанную на функции псевдонимности блокчейна, невосприимчивости к различным атакам (таким как повтор, атаки «человек посередине» и пассивное прослушивание) и совершенную прямую секретность сеансового ключа.

Сеть с нулевым касанием и управление услугами (СУУ): инициатива СУУ представляет собой новую парадигму для достижения автоматизации сетей и услуг

E2E. Конечной целью СУУ является реализация стопроцентно автономных сетей, способных к самоконфигурации, самооптимизации, самоконтролю, самовосстановлению и самомасштабированию без какого-либо вмешательства человека. Таким образом, СУУ приводит к гибкому и быстрому предоставлению услуг, обеспечивая экономическую устойчивость в диверсифицированных сетях [36, 44]. Архитектура СУУ сформирована с использованием интерфейсов на основе намерений, операций с обратной связью и методов ИИ/МО для обеспечения полной автоматизации операций управления. Предполагается, что сложная экосистема 6G будет состоять из нескольких операторов и нескольких поставщиков услуг (в виртуальной сети). Таким образом, СУУ будет играть важную роль для сетей 6G в достижении E2E-автоматизации сетевых ресурсов и услуг, которые относятся к разным частям мобильной сети, охватывающим нескольких операторов и поставщиков услуг.

Управления блокчейном и ИИ могут поддерживать механизмы междоменной безопасности и управления доверием в предлагаемой архитектуре СУУ, помогая автоматизировать жизненный цикл службы в многопользовательских и многосторонних средах. Реализация распределенной безопасности в блокчейне осуществляется доверием смарт-контрактов между доверенными и ненадежными сторонами. Этот подход можно дополнительно комбинировать с распределенным искусственным интеллектом для оркестровки когнитивной сети, что приводит к автоматизации процессов сети. Таким образом, блокчейн может стать многообещающей технологией для повышения доверия к управлению сетью с нулевым касанием. Тем не менее, по-прежнему сложно поддерживать надлежащий баланс между конфиденциальностью данных и доверием к автоматизированной системе на основе блокчейна для работы с гарантированной информационной безопасностью.

1.3. Совместное использование ресурсов на базе грид-систем

Способ решения сложных задач, которые требуют большого количества исполнителей (вычислительных ресурсов, а также ресурсов хранения и передачи

данных), работающих параллельно над разными аспектами задачи. Общий термин "грид-системы" охватывает различные подходы и конкретные технологии, используемые для решения таких задач.

Основа грид-систем заключается в обеспечении стабильной работы набора служб на основе широко принятых открытых стандартов и программного обеспечения, называемого "промежуточным программным обеспечением" или "middleware" на английском. Это обеспечивает надежный и унифицированный доступ к географически распределенным информационным и вычислительным ресурсам, включая отдельные компьютеры, кластеры, суперкомпьютерные центры и хранилища информации [72].

Создание грид-систем стало возможным благодаря значительным достижениям в нескольких областях:

- Увеличение производительности массово производимых микропроцессоров. Современные персональные компьютеры достигли производительности, сравнимой с суперкомпьютерами, которые были актуальны десять лет назад.
- Появление быстрых сетевых соединений. Сегодня основные магистрали передачи данных имеют пропускную способность в несколько гигабит в секунду.
- Глобализация обмена информацией через Интернет и Веб.
- Развитие методов метакомпьютинга, научной дисциплины, которая занимается организацией массовых и распределенных вычислительных процессов.

Грид-инфраструктура основывается на предоставлении ресурсов для общего пользования и использовании публично доступных ресурсов. Важным понятием в этой концепции является виртуальная организация (ВО).

Идея грида возникла в ответ на потребность в доступе к большим информационно-вычислительным ресурсам, которые динамически выделяются для решения сложных задач в различных областях, таких как наука,

промышленность, административная деятельность и коммерция. Создание грид-среды включает распределение вычислительных ресурсов по разным расположенным территориально сайтам, на которых установлено специальное программное обеспечение. Это программное обеспечение отвечает за распределение задач между сайтами, прием и возврат результатов пользователю, управление правами доступа пользователей к ресурсам и мониторинг ресурсов.

Общедоступные ресурсы грид-системы включают вычислительные узлы, узлы хранения и передачи данных, данные и прикладное программное обеспечение. Вычислительные ресурсы предоставляют пользователю процессорную мощность и могут быть представлены в виде кластеров или отдельных рабочих станций. Любая вычислительная система, при наличии соответствующего программного обеспечения, может быть потенциальным вычислительным ресурсом грид-системы [73].

Ресурсы хранения данных также используют программное обеспечение, которое реализует унифицированный интерфейс управления и передачи данных. Физическая архитектура ресурсов хранения не является принципиальной для грид-системы, будь то жесткий диск на рабочей станции или система массового хранения данных объемом сотни терабайт. Главным критерием для ресурсов хранения данных является их объем, который в настоящее время измеряется в терабайтах (Тб).

Информационные ресурсы и каталоги представляют собой особый вид ресурсов хранения данных, используемых для хранения метаданных и информации о других ресурсах грид-системы. Информационные ресурсы позволяют структурированно хранить большие объемы информации о состоянии грид-системы и эффективно выполнять задачи поиска ресурсов. Сетевой ресурс играет роль связующего звена между распределенными ресурсами грид-системы.

Понятия "состояние сервиса" и "сервис без состояний" являются ключевыми в теории слабосвязанных сервисов. В контексте слабой связи, преимущества возникают от того, что клиент может использовать любой сервис, способный выполнить его запрос. Если клиент ограничен единственным сервисом,

преимущества слабой связи снижаются. Например, для простых запросов, таких как использование калькулятора или получение информации о курсе акций, клиент запрашивает информацию и получает ее, после чего транзакция считается завершенной, и у клиента нет особой необходимости обращаться к тому же сервису снова. В этом случае связь между клиентом и сервисом является слабой.

Однако, для более сложных запросов, которые требуют нескольких шагов, сервис может сохранять информацию (состояние) о предыдущих шагах в своей локальной памяти. Это позволяет сервису использовать сохраненную информацию при последующих запросах от клиента. В таком случае сервис обладает состоянием (stateful service), и клиент должен обращаться к тому же сервису на следующем шаге. Это может привести к задержкам или отказу в обслуживании, особенно если множество клиентов используют один и тот же сервис или если сервис аварийно прекращает работу между шагами.

Более правильный подход при разработке сервисов основан на "сервисах без состояний" (stateless). Это особенно важно при многошаговой обработке запросов. На каждом промежуточном шаге сервис должен предоставлять клиенту достаточную информацию о состоянии, чтобы другой сервис с соответствующими свойствами мог идентифицировать и продолжить обслуживание. Клиент должен передавать информацию о состоянии любому сервису на следующем шаге. Выбранный сервис должен быть способен принимать и обрабатывать информацию о состоянии, предоставляемую клиентом, независимо от того, обрабатывал ли он сам запрос на предыдущем шаге или это делал другой сервис.

В контексте рисунка 3, клиент делает запрос, который требует три шага обработки и может включать несколько сервисов. Каждый сервис может быть способен обрабатывать любую часть или весь запрос. Важно, чтобы информация о состоянии передавалась между сервисами и клиентом для обеспечения непрерывной обработки запроса.

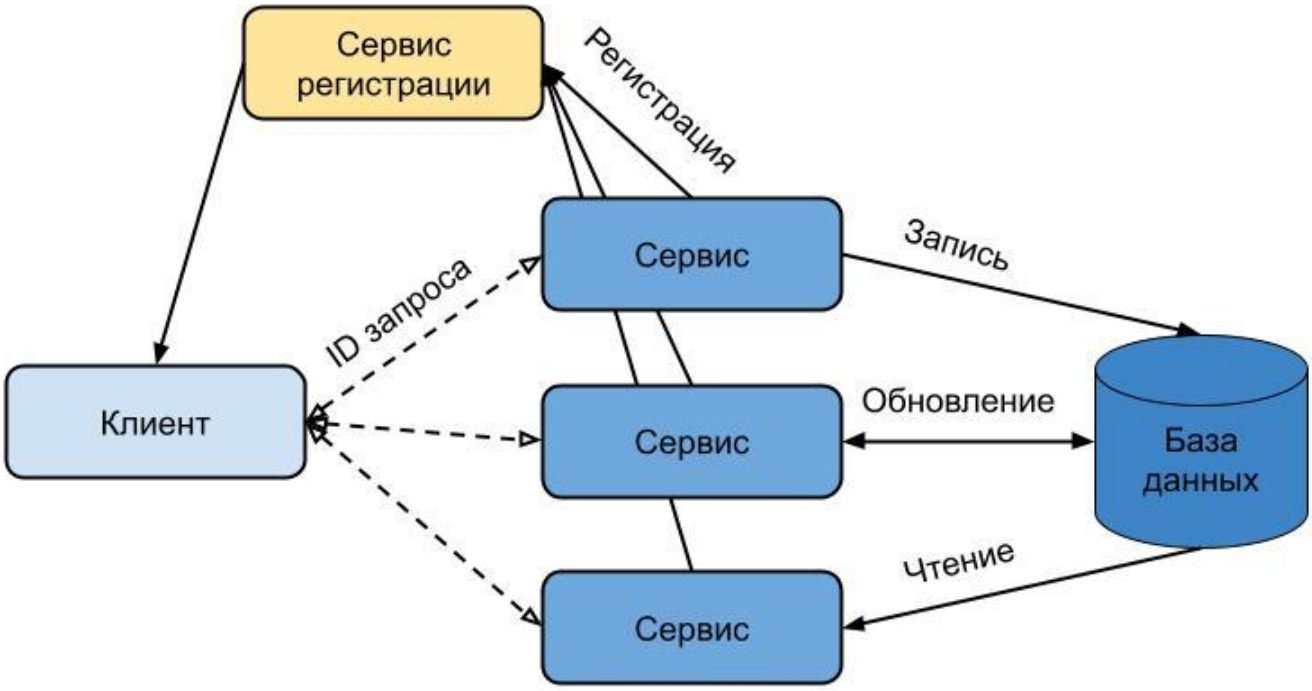

Рисунок 3 - Многошаговое взаимодействие клиента и сервисов

В описанном выше подходе сервис, обрабатывающий первый шаг, сохраняет детали обработки текущего запроса в базе данных и возвращает информацию клиенту вместе с идентификатором запроса. Затем клиент может запросить подтверждение от пользователя перед передачей этого идентификатора другому сервису, который использует его для поиска информации о состоянии в базе данных и инициирует второй шаг обработки. Этот сервис обновляет базу данных и возвращает дополнительную информацию клиенту. Наконец, клиент передает операционный идентификатор третьему сервису вместе с запросом на завершение обработки [72].

В большинстве нетривиальных приложений требуется доступ к информации о состоянии, и вопрос заключается не в том, должны ли состояния существовать, а где они должны храниться. В описанном выше подходе состояние обработки запроса отделено от сервиса, выполняющего обработку, что обеспечивает слабую связь между сервисами и клиентом. Для уменьшения объема информации о состоянии, передаваемой между клиентом и сервисами, существенные детали обработки запроса сохраняются в базе данных. Все участвующие сервисы должны

иметь доступ к базе данных и получать необходимую информацию на основе идентификатора клиента/запроса, который легко передается от клиента сервисам.

Отметим, что работа грид-систем опирается на программное обеспечение промежуточного уровня, которое обеспечивает контролируемый доступ к ресурсам. В начальной стадии развития грид-системы строились на основе специально разработанных общедоступных компонентов или закрытых (проприетарных) технологий. Хотя различные общественные и коммерческие решения демонстрировали успехи в своих областях применения, каждое с собственными преимуществами и ограничениями, у них был ограниченный потенциал в качестве основы для следующего поколения грид-систем, которые должны быть масштабируемыми и интероперабельными, чтобы удовлетворить потребности крупномасштабных научных и производственных проектов.

На рисунке 4 представлена схема простого сервисно-ориентированного грида, в котором сервисы используются как для виртуализации ресурсов, так и для обеспечения других функциональных возможностей грида.

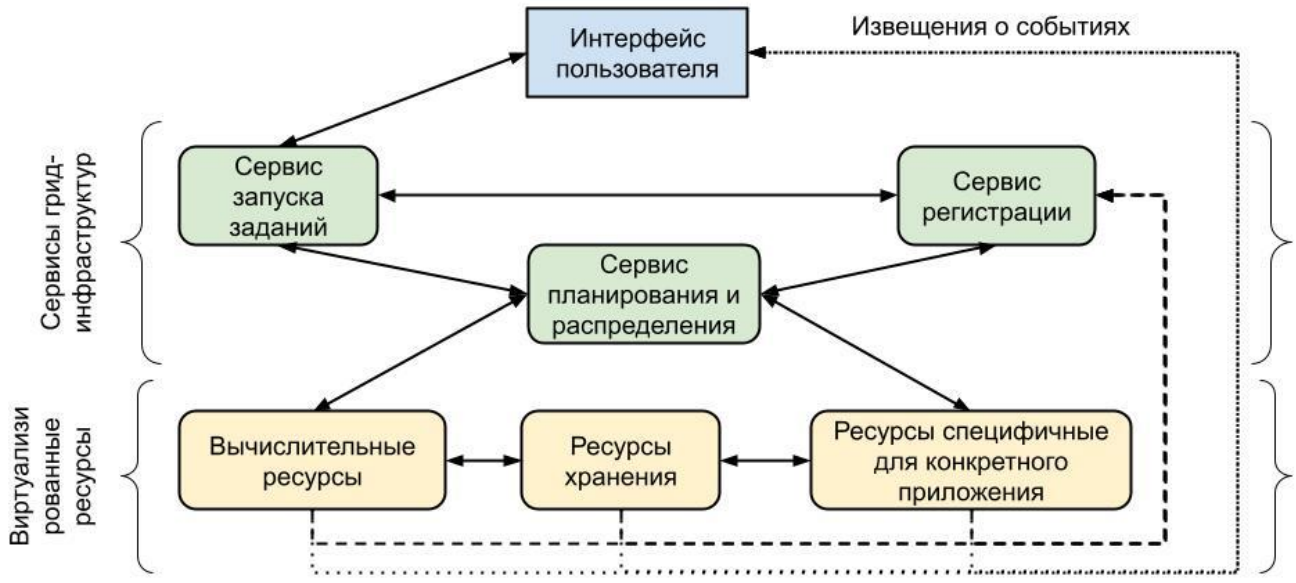

Рисунок 4 - Упрощенная схема сервисно-ориентированного грида

На схеме показана единая консоль и для запуска заданий в грид-среду, и для управления грид-ресурсами. Программное обеспечение интерфейса пользователя (консоли) обращается к сервису регистрации, чтобы получить информацию о

существующих грид-ресурсах. Затем пользователь посредством консоли входит в контакт с сервисами, «представляющими» (виртуализующими) каждый ресурс, чтобы запросить периодическое получение данных о работе ресурсов и получение извещений о существенных изменениях в их состоянии (например, если ресурс становится недоступным или сильно загруженным). Пользователь направляет запрос на запуск задания в службу запуска, которая передает запрос службе распределения заданий (часто называемой «планировщиком»). Служба распределения контактирует со службой, представляющей приложение, и запрашивает информацию о требованиях к ресурсам для выполнения задания. Затем служба распределения запрашивает у службы регистрации информацию о всех подходящих ресурсах в гриде и напрямую контактирует с ними, чтобы убедиться в их доступности. Если подходящие ресурсы доступны, планировщик выбирает наилучшую доступную совокупность ресурсов и передает информацию о них сервису приложения с запросом на начало выполнения. В противном случае планировщик ставит задание в очередь и выполняет его, когда необходимые ресурсы становятся доступными. Когда выполнение задания заканчивается, сервис приложения сообщает о результате планировщику, который извещает об этом сервис запуска заданий. Сервис запуска заданий, в свою очередь, уведомляет пользователя. Отметим, что этот пример для ясности сильно упрощен: функционирование реального грида промышленного уровня является намного более сложным чем показано на схеме. Главным результатом его работы должна быть высокая степень автоматизации и оптимизации использования ресурсов в рамках грид-среды.

Выводы к главе 1

Выявленные задачи распределённого реестра требуют следующих подходов к решениям:

1. В сетях 6G может быть доступно программное обеспечение с открытым исходным кодом для реализации различных сетевых функций на уровне ядра, транспорта или доступа;
2. Двухуровневую иерархическую архитектуру блокчейна можно рассмотреть, как основную систему управления ресурсами;
3. Использование концепции грида, как ответ на появляющиеся потребности в крупных информационно-вычислительных ресурсах, динамически выделяемых для решения громоздких задач;
4. Общие принципы грид-технологии: добавлять пользователей и передавать рабочие ресурсы.
5. Эффективно распределяя ресурсы, грид-технология значительно сокращает время ожидания доступа к ним.

В следующей главе мы рассмотрим способы хранения данных и возможные риски с точки зрения безопасности.

ГЛАВА 2. ДОСТУП РАЗГРАНИЧЕНИЯ ДАННЫХ ПО КЛАССИФИКАТОРАМ

2.1. Комплексные модели безопасности персональных данных

Все сложившиеся процессы управления опираются на сбалансированные системы менеджмента, уровень которых измеряется моделями зрелости. Другим направлением построения комплексных систем безопасности персональных данных является создание фреймворков, нацеленных на построение в организации системы управления персональными данными.

Преобладают два различных подхода: вертикальный – сквозной подход для организации процессов управления, и горизонтальный – ориентация на ту или иную технологическую экосистему (например, Интернет Вещей (IoT), мобильную связь, социальные сети).

Стандарт ISO/IEC 29100:2011 (а также его обновленная версия ISO/IEC 29100:2020) инициирует ряд вопросов, специфичных для обработки персональных данных:

- определяет общую терминологию конфиденциальности;
- определяет участников и их роли в обработке информации, позволяющей установить личность;
- описывает соображения защиты конфиденциальности, основанные на организационных и технических аспектах;
- содержит ссылки на известные принципы конфиденциальности.

Кроме указанного стандарта ISO/IEC выпустило целый ряд документов, описывающих подходы к организации и технической архитектуре вопросов безопасности персональных данных [74]:

- ISO/IEC 27403. IoT security and privacy. Guidelines for IoT-domestics. Стандарт пока имеет статус “draft”;

- ISO/IEC 27550. Privacy engineering for system life cycle processes. Стандарт описывает подход к управлению приватностью через жизненный цикл автоматизированных систем. Все еще находится в статусе “draft”;
- ISO/IEC 27556. User-centric framework for the handling of PII based on privacy preferences. Формирует систему управления персональными данными со стороны пользователя;
- ISO/IEC 27561. Privacy operationalization model and method for engineering (POMME). Операционная модель управления безопасностью частных данных. Готовится к выпуску;
- ISO/IEC 27570. Privacy guidelines for smart cities. Адресует вопросы безопасности персональных данных для умных городов;
- ISO/IEC 29134. Guidelines for privacy impact assessment. Оценивает влияние\ущерб в результате нарушений конфиденциальности персональных данных;
- ISO 31700. Privacy by design for consumer goods and services. Обеспечивает безопасность персональных данных при продаже потребительских товаров и сервисов.

В совокупности инициативы ISO/IEC покрывают широкий круг вопросов, связанных с организацией работы и технической архитектурой для различных систем. В то же время многие из стандартов все еще находятся в разработке, а их многочисленность и несостыковки затрудняют их применение как единой системы руководящих принципов.

Национальный институт стандартов и технологий США (NIST) также предпринял попытку разработать комплексную модель безопасности персональных данных [74]. Среди опубликованных документов и руководств следует отметить следующие:

- NIST.SP.800-122. Guide to Protecting the Confidentiality of Personally Identifiable Information (PII). Руководство по защите конфиденциальности персональных данных. Содержит широкий набор мер, в том числе

организационных, по управлению персональными данными, подготовке их к обезличиванию, документированию процедур и действий персонала, процессу публикации персональных данных, а также действиям в результате утечек;

- NIST.SP.800-53. Security and Privacy Controls for Federal Information Systems and Organizations. Организация контроля процессинга персональных данных для государственных организаций;
- NIST.SP.800-37. Risk Management Framework for Information Systems and Organizations A System Life Cycle Approach for Security and Privacy. Описывает процесс управления риском в организациях с учетом обеспечения безопасности персональных данных;
- NIST.IR.8053. De-Identification of Personal Information. Описание методов обезличивания, подходов к вычислению рисков и общей организации процесса обезличивания;
- NIST PRIVACY FRAMEWORK 1.0 (PRAM): A tool for improving privacy through enterprise risk management. Общая система (фреймворк) управления безопасностью, персональными данными в организациях с учетом организационных процедур, указания для системы приоритезации рисков, набор инструкций и чек листов по известным проблемам безопасности персональных данных.

Обеспеченные документы NIST позволяют рассматривать проблему работы с персональными данными в контексте жизненного цикла в едином организационно-техническом поле.

Безопасность приватных данных через построение системы. Обеспечение безопасности персональных данных через систему последовательного внедрения технических и организационных процедур в практику работы организации (Privacy By Design) – важный компонент общего подхода к построению надежных процедур обработки персональных данных, включая обезличивание данных [57].

Хотя нет единой организации, стоящей за данной концепцией, большинство стандартизируемых и технологических компаний признают этот подход важнейшим в подходе к безопасности персональных данных [13, 41,46].

Эта концепция часто противопоставляется концепции “Приватность по умолчанию” (Privacy By Default), однако в действительности эти подходы дополняют друг друга. Идея Privacy By Design – это верхнеуровневая концепция, декларирующая отсутствие статичного решения вопроса безопасности персональных данных и предлагающая вместо этого использовать динамичный подход, в котором происходит постоянная подстройка под внешние угрозы и вызовы, а система строится организацией шаг за шагом [43].

Основные принципы данного подхода:

- проактивный подход имеет преимущество перед реактивными действиями. Организации должны обеспечивать предварительные действия по управлению безопасностью данных, а не решать уже возникшие проблемы;
- конфиденциальность в качестве настройки по умолчанию. Имеется в виду, что обеспечение конфиденциальности должно быть заложено в каждую систему, оперирующую персональными данными. От физического лица не требуется никаких действий, при их отсутствии его данные остаются в безопасности;
- конфиденциальность встроена в дизайн и архитектуру ИТ-систем, а также в бизнес-практику организации;
- обеспечение полной функциональности при минимизации рисков. Принцип декларирует наличие возможного компромисса, дающего выигрыш для всех сторон;
- сквозная безопасность - защита полного жизненного цикла данных, включая их трансформацию, передачу и уничтожение;
- видимость и прозрачность политики обработки персональных данных;

- уважение к конфиденциальности пользователей - ориентированность на пользователя.

Проведенный анализ позволяет обобщить некоторые положения имеющихся практик и сделать ряд значимых выводов:

- Необходимость обезличивания данных непосредственно связана с процессами обмена данных, возникающими в современных интеграционных процессах. Модели, ориентированные на статические ситуации хранения данных без учета обмена ими или публикации, не охватывают весь жизненный цикл;
- Проблематика обезличивания персональных данных тесно связана с общим управлением безопасностью персональными данными;
- Хотя данное направление обезличивания все еще находится в стадии развития (новые работы появляются регулярно), в отношении управления обезличиванием на основе рисков имеется сложившийся консенсус: все современные модели базируются на оценке рисков в процессе обезличивания;
- Важное значение для проведения процессов обезличивания имеют контекстные риски, связанные с организационными-техническими мероприятиями внутри организаций. Игнорирование таких мер ведет к большей доступности сторонних информационных ресурсов значительно увеличивающих вероятность атак с использованием дополнительных ресурсов и более жестким требованиям к обезличенным данным;
- Текущая отечественная система работы с персональными данными в части обезличивания не учитывает изменившийся информационный ландшафт, ориентирована на систему запретов и излишне государственно-центрирована. Игнорирование использования больших данных бизнесом и некоммерческими организациями только увеличивает риски утечек данных и, в конечном счете, приводит к технологическому отставанию.

2.2. Общий подход к вычислению риска

Данные обезличиваются с целью эффективного и безопасного обмена информацией между несколькими определенными сторонами или публикации их для неопределенного круга лиц в целях обеспечения прозрачности и доверия к принимаемым решениям, исследования наборов данных независимыми аналитиками данных и ряда других целей (рисунок 5). При этом безопасность данных (в части конфиденциальности персональных данных) и информационная эффективность (полезность) получаемых в результате обезличивания данных находятся в противоречии.

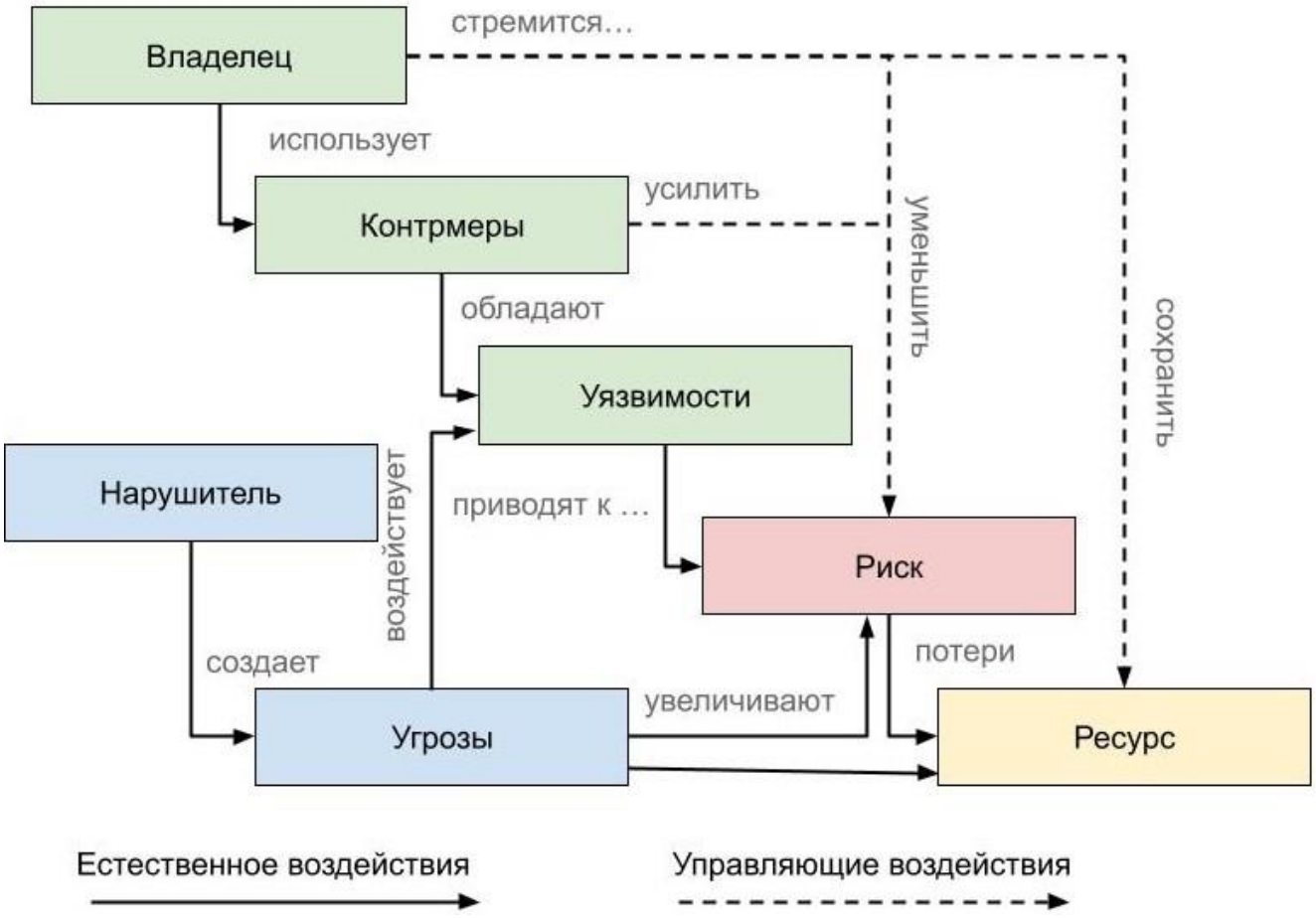

Рисунок 5 - Процесс оценки рисков, принятый в ИБ

Безопасность персональных данных отличается от классической концепции информационной безопасности [6, 45], фокус которой – несанкционированная

деятельность (собственно приводящая к потере конфиденциальности, целостности или доступности). При обработке персональных данных с целью обезличивания речь идет о плановой деятельности, которая, однако, может создать негативные последствия для частной жизни отдельных лиц. Угрозы конфиденциальности персональных данных возникают как в результате авторизованной обработки данных, так и в результате несанкционированного доступа:

- в результате несанкционированного доступа к данным – эмоциональные страдания лиц, экономические потери от кражи личных данных, а также физический или психологический вред из-за преследований;
- в результате снижения качества обработки информации в рамках плановой обработки – снижение качества оказываемых услуг, сбои в работе систем, неверно принятые или затянутые решения, которые могут также воздействовать на здоровье отдельных лиц или целые группы населения.

Ключевые положения предлагаемой модели риска процедуры обезличивания учитывают проблемы конфиденциальности, возникающие как вследствие несанкционированного доступа, так и вследствие плановой обработки данных:

- риск публикации обезличенных данных является мерой того, в какой степени субъекту угрожают потенциальные обстоятельства или события и является функцией вероятности P возникновения таких событий и неблагоприятных последствий нарушения конфиденциальности персональных данных – ущерба (I, Impact) [77]:

$$R = P * I \tag{1}$$

- построение полностью анонимизированных наборов данных, сохраняющих полезные сведения, является идеализированной гипотезой, не осуществимой в практических решениях. Предлагаемый подход вместо это базируется на вероятности проведения *атак повторной идентификации* и установке приемлемого риска;

- процесс, с помощью которого выявляются риски, согласно называется *оценкой риска*;
- для оценки рисков требуется модель риска, определяющая факторы риска и взаимосвязи между ними:
  - факторизация риска;
  - сценарии использования данных;
  - количественные пороги риска;
  - необходимый уровень полезности получаемых обезличенных данных;
  - порядок построения риск-модели для конкретной процедуры обезличивания, включая возможность переоценки риска при использовании различных методов обезличивания.

Ущерб может включать:

- расходы на соблюдение нормативных требований, связанных с проблемами конфиденциальности у физических лиц;
- прямые затраты на штрафы и судебные издержки;
- потери бизнеса, отказ от использования продуктов или услуг;
- репутационные издержки, ведущие к утрате доверия со стороны пользователей;
- снижение производительности или невозможность осуществления миссии организации.

Таким образом оценка риска, предлагаемая в данной модели, содержит ущерб в качестве параметра, в дальнейшем под оценкой риска понимается оценка вероятности наступления событий $P \in [0;1]$;

- общая формула факторизации вероятность риска повторной идентификации:

$$P_{re-id} = P_{context} * P_{data} \quad (2)$$

где $P_{re\text{-}id}$ – общий риск;

$P_{context}$ - контекстный риск, определяемый набором организационных и технических рисков;

P_{data} – риски данных.

- риски данных связаны с конкретным набором данных и должны вычисляться с учетом характеристик набора данных – в частности, выделением прямых идентификаторов и квази-идентификаторов. Прямые идентификаторы и их замена являются наименьшей проблемой решения задачи о рисках повторной идентификации. Квази-идентификаторы, их корректное выделение и применение методов обезличивания с учетом их структуры и типа – значительно более сложная задача, учитывающая вероятностный подход к оценке риска.

Проведение обезличивания с учетом риск-модели требует баланса между полезностью получаемых в результате обезличивания данных и приемлемой величиной риска. Пороговые значения риска устанавливаются в соответствии со сценариями использования. Общее представление о порядках риска можно извлечь из данных в таблице 1[5,18]:

Таблица 1 - Общие значения порога риска

Влияние на конфиденциальность ПДн	Приемлемый порок риска	Размер класса эквивалентности для агрегированных данных
Низкое	0,1	10
Среднее	0,075	15
Высокое	0,05	20

Приемлемый уровень риска определяется с учетом следующих параметров:

- вероятность осуществления угрозы конфиденциальным данным ($P_{context} * P_{data}$);
- потенциальный ущерб I (оцениваемый в терминах “низкий”, “средний”, “высокий” или через баллы) – см. [14, 33]. Ущерб также определяется с учетом объема публикуемых данных и чувствительности атрибутов;

- сценарий публикации – влияет на рекомендуемые классы эквивалентности и, следовательно, риск данных;
- бизнес-модели организации, публикующей данные. Этот параметр описывает толерантность к риску с учетом особенностей бизнеса (отрасль), стоимости активов (данных, содержащих приватные характеристики), предпочтений руководства.

В рамках рассматриваемой модели используется ряд оцениваемых количественно метрик:

- *Уровень риска.* Представляет собой произведение ущерба на вероятность риска повторной идентификации. Вероятность, в свою очередь, факторизуется на вероятность контекстных рисков и вероятность данных. Последняя представляет собой метрику, привязанную к данным – их структуре и содержанию, – и отражает способность внешнего агента (“злоумышленника”) восстановить связь между субъектом персональных данных и его характеристиками (атрибутами). Оценка вероятности риска повторной идентификации для данных зависит от того, известны ли злоумышленнику какие-либо данные о физических лицах, включенных в рассматриваемый набор;
- *Уровень полезности данных* (*data utility*). В процессе обезличивания неизбежно происходит потеря информации, связанная с применением методов обезличивания. Оценки для больших наборов удобно производить на основе внутренних метрик, привязанных к свойствам набора. Представляет из себя также оценку качества данных;
- *Обратимость* (*reversibility level*). Обратимость позволяет поддерживать связь исходного и обезличенного набора данных, может оцениваться в диапазоне [0-1]. Оценка обратимости может производиться, исходя из двух опций:
  - включение в обезличенный набор подготовленных однозначных идентификаторов данных (это могут быть специальные

идентификаторы, сохраненные в процессе преобразования, hash значения реальных идентификаторов или сформулированные специально псевдонимы). В этом случае

$$K_{revers} = 1 \quad (3)$$

– отсутствие в наборе атрибутов, однозначно связывающих исходный и обезличенный набор. В этом случае коэффициент обратимости равен вероятности повторной идентификации для набора данных:

$$K_{revers} = P_{data} \quad (4)$$

- *Вариативность метода обезличивания* (*variability*)**.** Ряд методов обезличивания допускают применение различных параметров, влияющих на способ формирования обезличенного набора:
  - в случае детерминированных методов обезличивания с конечным набором значений, принимаемых параметрами N (обобщение, подавление) коэффициент определяется:

$$K_{var} = \frac{N}{2} \quad (5)$$

  - в случае непрерывных изменений параметров коэффициент определяется дисперсией значений в результате преобразования:

$$K_{var} = \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2 \quad (6)$$

- *Изменяемость* (*flexibility*). Оценивает возможность внесения дополнений (искажений) в массив обезличенных данных. В силу некоторых причин уже обезличенный набор данных может быть подвергнут дополнительным изменениям. Поскольку методы обезличивания нацелены на отдельные атрибуты, применение повторных изменений к ранее измененным столбцам должно поддерживаться критериями, гарантирующими неизменность основных распределений по данному атрибуту:

– для контроля корректности проведенных изменений рекомендуется использовать проверки нулевой гипотезы (гипотеза о сходстве, равенство дисперсий):

$$H_0 = \sigma_1^2 - \sigma_2^2 \tag{7}$$

– в рамках этого подхода рекомендуется применять F-тест. Пусть выборка *X* из *m* случайных значений первоначально обезличенного набора сравнивается с выборкой *Y* из набора с n случайными значениями, претерпевшего повторное значение. Функция Фишера:

$$F = \frac{\sigma_X^2}{\sigma_Y^2} = \frac{\sqrt{\frac{1}{m-1}\sum_{i=1}^{m}(x_i - \bar{x})^2}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{8}$$

– при подобном преобразовании F ≈ 1. Для выполнения данного теста есть несколько ограничений: предполагается, что значения атрибутов в обоих наборах распределены нормально, выборка ограничена.

- *Стойкость обезличенного набора к атакам.* Определяется вероятностью успеха атак повторной идентификации;
- *Совместимость различных обезличенных наборов* (при сопоставлении атрибутов). Два набора могут быть сопоставлены при условии одинаковых классов эквивалентности для сравниваемых атрибутов, в данном случае рекомендуется проводить сравнение с учетом метрики ℓ - разнообразие;
- *Параметрический объем* определяется необходимым объемом дополнительной (служебной) информации для метода. В простейших случаях представляет собой число параметров для настройки методов обезличивания, а также вспомогательные данные журналирования

проведенных экспериментов, включая хранение промежуточных версий обезличенных наборов.

Выводы к главе 2

Проведение обезличивания с учетом риск-модели требует баланса между полезностью получаемых в результате обезличивания данных и приемлемой величиной риска. Пороговые значения риска устанавливаются в соответствии со сценариями использования.

В рамках рассматриваемой модели используется ряд оцениваемых количественно метрик:

- *Уровень риска.* Представляет собой произведение ущерба на вероятность риска повторной идентификации. Вероятность, в свою очередь, факторизуется на вероятность контекстных рисков и вероятность данных. Последняя представляет собой метрику, привязанную к данным – их структуре и содержанию, – и отражает способность внешнего агента (“злоумышленника”) восстановить связь между субъектом персональных данных и его характеристиками (атрибутами). Оценка вероятности риска повторной идентификации для данных зависит от того, известны ли злоумышленнику какие-либо данные о физических лицах, включенных в рассматриваемый набор;
- *Уровень полезности данных* (*data utility*)**.** В процессе обезличивания неизбежно происходит потеря информации, связанная с применением методов обезличивания. Оценки для больших наборов удобно производить на основе внутренних метрик, привязанных к свойствам набора. Представляет из себя также оценку качества данных;
- *Обратимость* (*reversibility level*). Обратимость позволяет поддерживать связь исходного и обезличенного набора данных.

С учетом вышесказанного, в следующей главе мы рассмотрим многоуровневый доступ к обработке транзакций.

ГЛАВА 3. МНОГОУРОВНЕВЫЙ ПОДХОД К БЕЗОПАСНОСТИ В РАСПРЕДЕЛЕННОМ РЕЕСТРЕ

3.1. Этапы формирования обработки транзакций на вертикальной масштабируемости

Хранение и использование данных на аппаратном уровне требует решения распределения данных по узлам связи с вертикальной масштабируемостью. Вертикальная масштабируемость на самом деле представляет собой возможность повысить эффективность используемого в настоящее время аппаратного или программного обеспечения за счет добавления к нему дополнительных ресурсов, путем добавления быстрых процессоров к серверу для увеличения его скорости [54]. Это означает добавление дополнительных аппаратных ресурсов к существующей машине за счет использования большего количества процессоров, памяти и т. д.

По сути, развертывание виртуального сервера является основополагающим условием для вертикального развертывания. Данная технология позволяет нам взаимно задействовать те же функции что и физические серверы. На самом деле, когда пользователь входит в консоль сервера (рисунок 6), фактически невозможно отличить физический сервер от виртуального до тех пор, пока вы не взглянете на драйверы [19, 20].

Рисунок 6 - Архитектура виртуального сервера

Различие между физическими серверами и виртуальными серверами заключается в том, что виртуальные серверы не устанавливаются на физическом оборудовании. Они скорее установлены на вещь, называемую гипервизором. Гипервизор позволяет запускать более одного виртуального сервера на одном физическом оборудовании. Использование виртуализации упрощает динамическое масштабирование приложения, поскольку клиенты взаимодействуют с виртуальным сервером на контроллере доставки приложений, а затем он взаимодействует с виртуальными серверами, расположенными на физических серверах внутри центра обработки данных [2, 3].

В грид-системе обеспечение взаимодействия между различными платформами, языками и программными средами является ключевым условием для эффективной работы. Для этого используются общие протоколы, которые регламентируют взаимодействие между элементами распределенной системы и структуру передаваемой информации [72,73].

Общая структура глобального грида описывается в виде стека протоколов. В этой модели каждый уровень стека предназначен для решения определенного набора задач и предоставляет услуги для более высоких уровней [1]. Верхние уровни стека ближе к пользователю и работают с абстрактными объектами, тогда как нижние уровни тесно связаны с физической реализацией грид-ресурсов. Эта структура стека протоколов аналогична модели OSI (Open Systems Interconnection Reference Model) для сетевых коммуникаций и разработки сетевых протоколов.

Стек грид-протоколов включает следующие уровни (рисунок 7):

1. *Аппаратный уровень (Fabric Layer):* содержит протоколы, которые используются соответствующими службами для работы с ресурсами непосредственно.
2. *Связывающий уровень (Connectivity Layer):* включает протоколы, обеспечивающие обмен данными между компонентами базового уровня, а также протоколы аутентификации.
3. *Ресурсный уровень (Resource Layer):* представляет собой ядро многоуровневой системы, где протоколы взаимодействуют с ресурсами через унифицированный интерфейс, не зависимо от архитектурных особенностей конкретного ресурса.
4. *Коллективный уровень (Collective Layer):* отвечает за координацию использования доступных ресурсов.
5. *Прикладной уровень (Application Layer):* описывает пользовательские приложения, работающие в среде виртуальной организации. Приложения используют протоколы, определенные на более низких уровнях стека.

Такая структура протоколов позволяет эффективно организовать взаимодействие и обеспечить совместимость между различными компонентами грид-системы, используя общие принципы и стандарты.

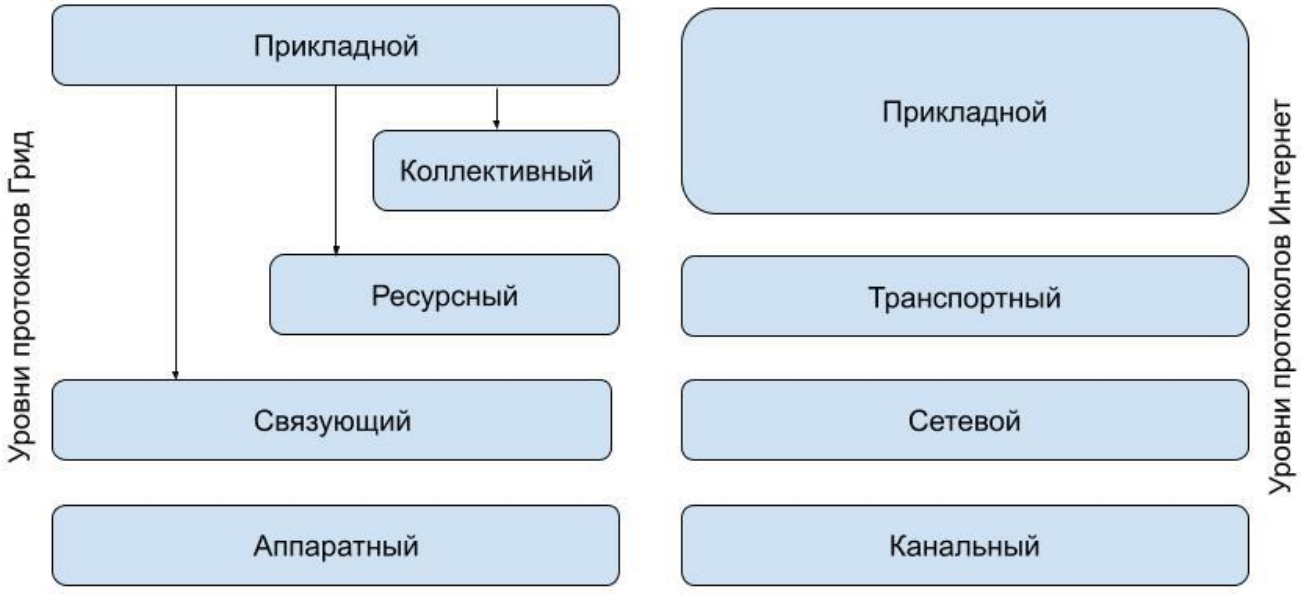

Рисунок 7 - Стеки протоколов грид-системы и сетевой модели

Количество серверов внутри центра обработки данных может быть изменено без каких-либо проблем с безопасностью, ускорением и доступностью соответствующего приложения, поскольку эти функции сосредоточены на контроллере доставки приложений, который может быть автоматизирован для добавления/удаления серверов внутри центра обработки данных без каких-либо проблем. По мере добавления каждого виртуального сервера, его программное обеспечение обращается к базовому оборудованию, сокращая время простоя [10].

Большинство алгоритмов оптимизации требуют синхронизацию локального одноранговогo доступа к информации глобальной сети. Это отдельная проблема, известная как проблема агрегации [64] и относящаяся к набору функций, обеспечивающих доступ к таким компонентам распределенной системы, как размер сети, средняя загрузка и время безотказной работы и так далее.

Решение задач устанавливалось на платформе DGT, которая рассматривает задачу отказоустойчивости с многоуровневым подходом обработки данных. Данный подход основан на развертывание виртуальной сети для решения задач или при разработке приложений. Сервера или узлы могут также быть расположены в различных физических сетях.

Кластеры узлов входят в более крупное деление сети – сегменты, которые могут быть двух видов: публичные и приватные (рисунок 8). В отдельной сети возможен только один публичный сегмент, присоединяющиеся узлы могут свободно входить во взаимодействие с другими узлами сегмента. Сеть может иметь несколько приватных сегментов, основное отличие которых от публичного сегмента состоит в контролируемой топологии.

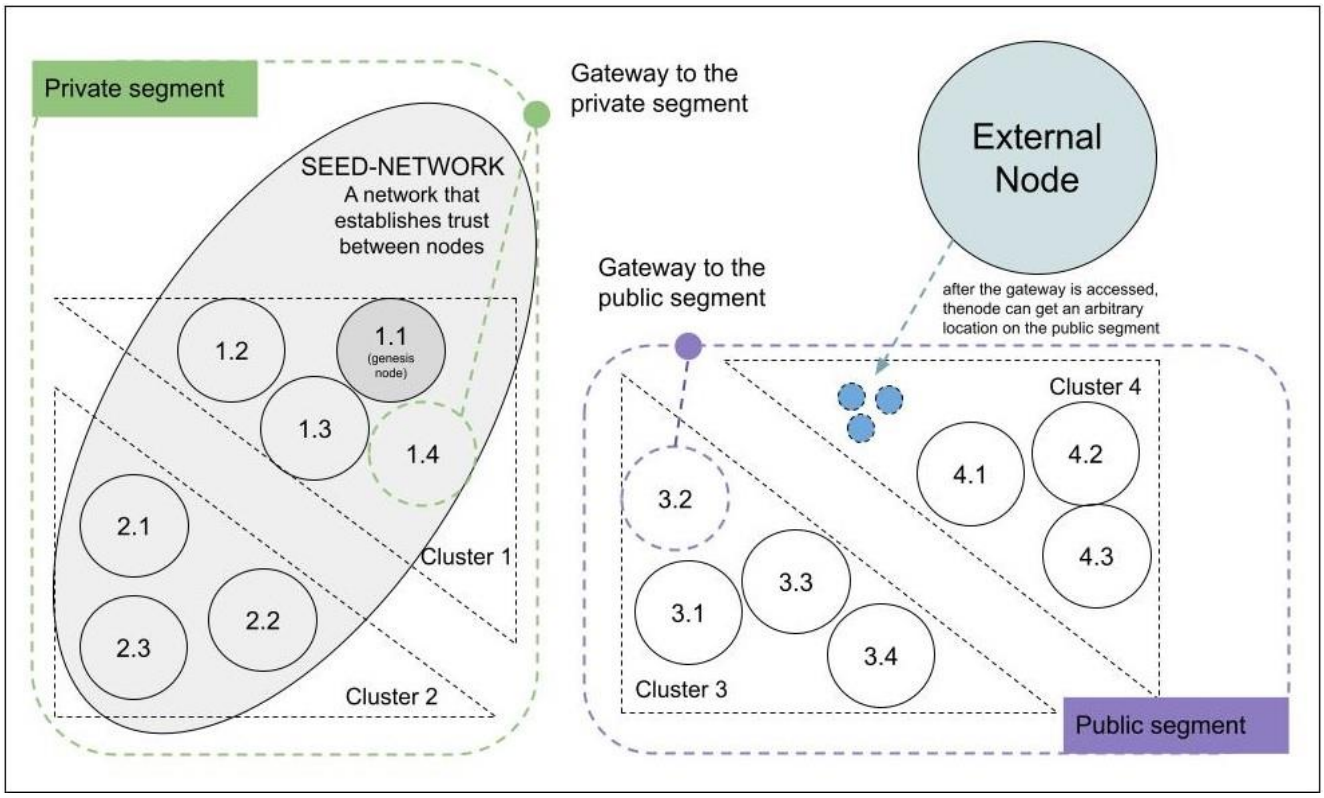

Рисунок 8 - DGT Network Topology and Node Attaching

Начальная реализация сети, называемая также “статическое ядро”, которое представляет собой группу узлов/кластеров, образующих специальные связи доверия (публичные ключи таких узлов заранее известны и прописываются в момент развертывания ядра). Присоединение других узлов требует обработки сертификатов узла для закрытых сегментов и/или динамического соединения в случае публичных сегментов.

Дополнительно можно рассмотреть сегментацию данных на платформе DGT с учетом кластерной топологии. Под кластером понимается группа узлов, осуществляющих первичный раунд голосования. Кластеры могут образовывать сложные структуры.

Размер кластера, присоединения к нему узлов определяется процессором топологии, представляющим собой отдельное семейство транзакций. Внутри каждого кластера определяется переменный лидер, который меняется после нескольких раундов голосования. Если при запросе лидера он не отвечает в течение определенного времени – происходит процедура выбора нового лидера.

Голосование первоначально осуществляется в кластере, затем по специальному алгоритму выбирается арбитр за пределами кластера. P-BFT препятствует атакам типа “двойной траты” за счет разницы времен “голосования” внутри кластера и характерного времени синхронизации DAG (state), которое осуществляется через пермалинки.

Кластеризация обеспечивает следующие преимущества:

- Формирование “топологически близких” групп узлов, что улучшает доверие между узлами (снижает риски атак) и повышает производительность;
- Позволяет обеспечить “sharding” сети, включая формирование приватных ветвей DAG’a;
- Повышает горизонтальную масштабируемость сети.

DGT позиционируется, как платформа для распределенных и децентрализованных вычислений, где система обрабатывает данные вне зависимости от конкретной прикладной задачи. Для решения конкретной задачи требуется настройка семейства транзакций, а также надстройка прикладной клиентской части [7, 19]. По сути, программное обеспечение DGT представляет из себя N-ное количество типовых узлов – Node, которые обеспечивают взаимодействие с другими узлами, проверку данных и вставку новых данных в хранилище (реестр), называемое также DAG или State. Оно нацелена на поддержку консорциум-базированных сетей. Это означает, что присоединение узла к сети возможно при выполнении некоторых условий. В простейших условиях это может быть проверка узла на наличие сертификата. В зависимости от реализации якорного механизма степень открытости сети различается – от полностью открытой (публичной), до полностью закрытой (приватной).

Узлы объединяются в группы и называются кластерами. Первоначальное взаимодействие осуществляется через связи между узлами с одним выделенным узлом в кластере – Leader. Лидер собирает данные от проверок транзакций в каждом узле. Такие проверки называются “голосованиями”. Если число

голосований превысило некоторый заданный порог, то транзакция считается одобренной в кластере и ожидает арбитраж, дополнительной проверки, выполняемый вне кластера. Внутри кластера узлы взаимодействуют между собой по выделенным каналам, называемым также пермалинками.

Вслед за Sawtooth, DGT является мульти-транзакционной системой, в которой возможно обращение нескольких семейств транзакций. Каждое семейство обрабатывается отдельным процессором транзакций. Семейства транзакций дополняют технологию смарт-контрактов, а также позволяют установить границы доступности разных типов транзакций для разных сегментов сети.

Но данных подход не может обеспечивать полную защиту системы, так как могут выйти из строя компоненты сервера или сам сервер.

Изначально был проведен набор экспериментов с двухуровневой системой обработки на платформе DGT. В ходе экспериментов была проверена отказоустойчивость системы, которая успешно справилась с более чем 1000 транзакций со стабильными 24 узлами. Результаты экспериментов показали, что пропускная способность составила в среднем 0,009 секунд за одну транзакцию (рисунок 9).

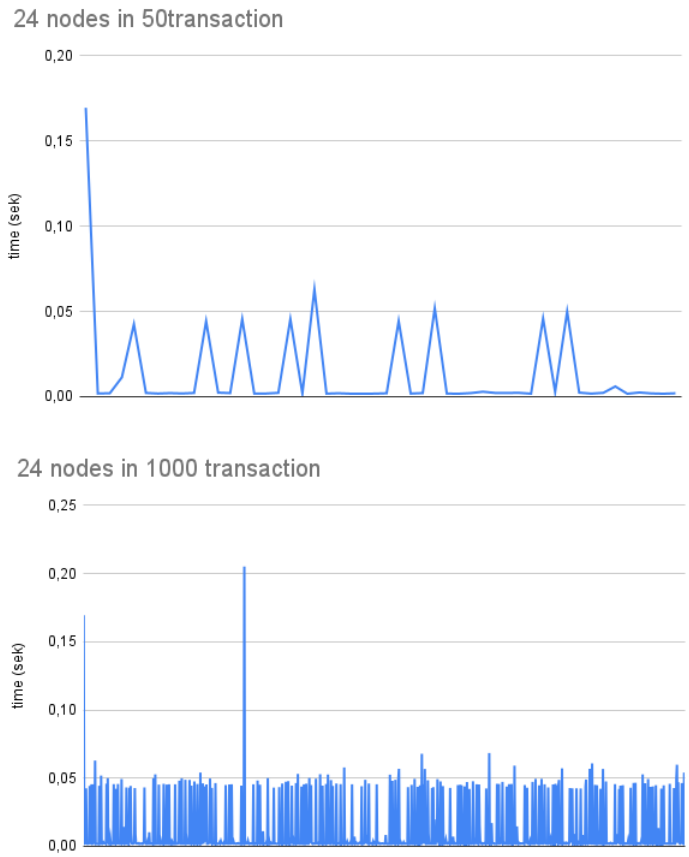

Рисунок 9 - График обработки трансакций при 24 развернутых узлах

На рисунках 10, 11 смоделирован принудительный сбой 6 узлов из 24 при обработке 1000 транзакций, в данном сценарии лидер узла начал обработку раундов голосования среди 18 узлов, и продолжил обрабатывать данные. В данном положении обработка заняла в среднем 0,0077 сек.

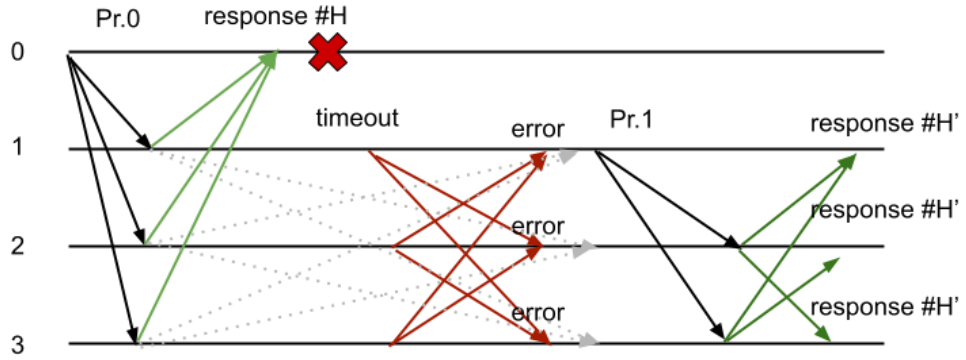

Рисунок 10 - При отказе в одном узле, переход на следующий для обработки транзакций

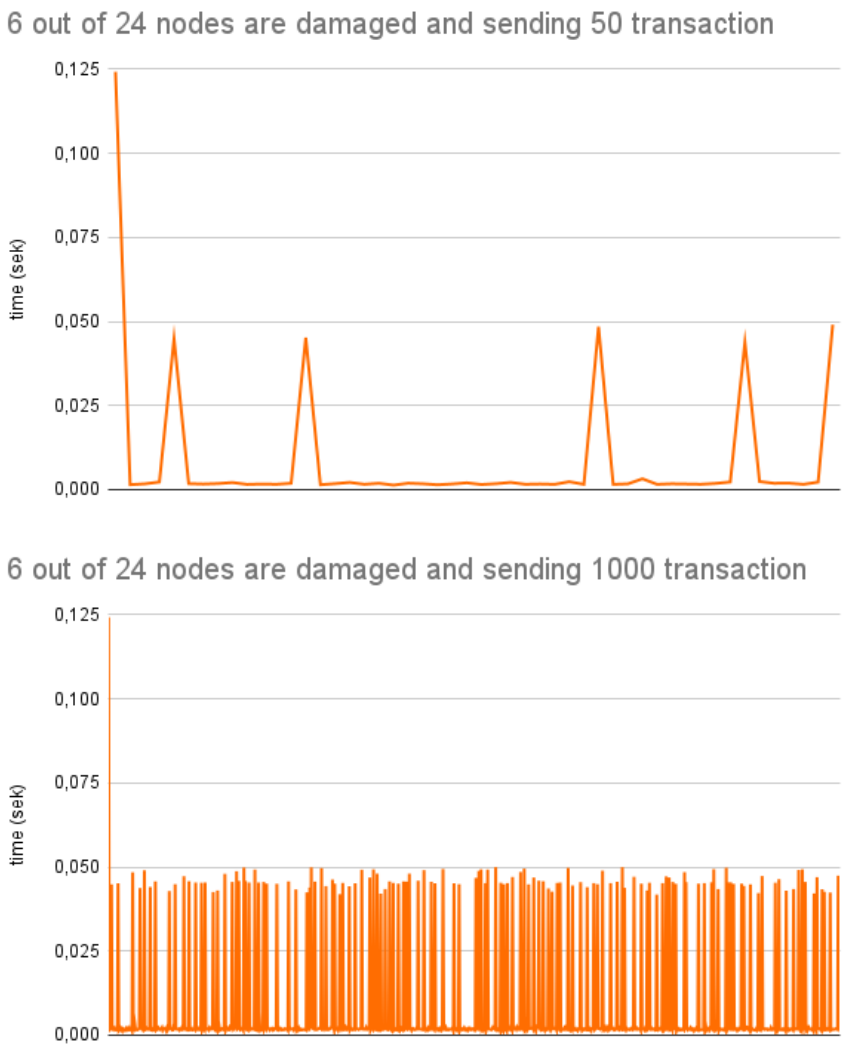

Рисунок 11 - График обработки транзакций при 24 развернутых узлах, при которой на 6 узлах принудительно реализован системный отказ

В третьем сценарии было протестировано решение отказоустойчивости на 12 узлах из общего числа в 24. Несмотря на такой масштабный сбой, алгоритм консенсуса PBFT продолжал работать стабильно, обрабатывая одну транзакцию в среднем за 0,0073 секунды (рисунок 12).

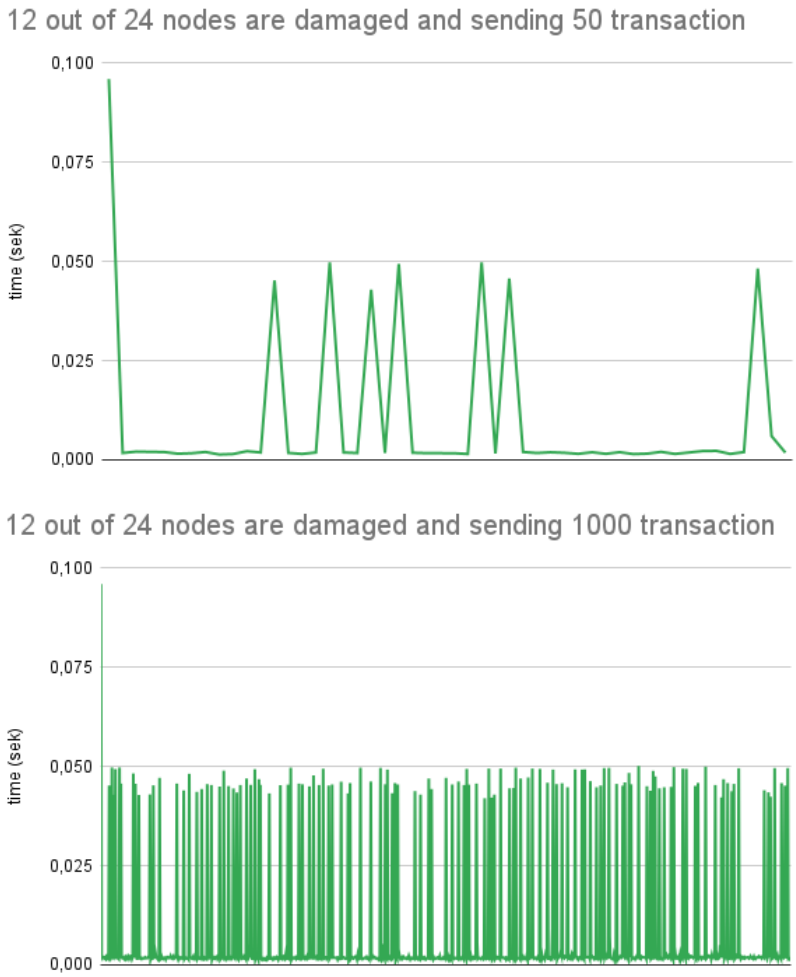

Рисунок 12 - График обработки транзакций при 24 развернутых узлах, при которой на 18 узлах принудительно реализован системный отказ

Каждый узел представляет собой набор сервисов, взаимодействующих между собой и отвечающих за организацию сети, хранение данных, обработку транзакций. Даже один узел поставляет значительный сервис, поддерживающий через API клиентские приложения. В то же время ряд сетевых возможностей платформы может быть востребован только при наличии нескольких узлов. Поэтому, благодаря данной технологии, мы нивелируем главный недостаток распределенной системы — потеря/сбой данных, что позволяет нам уменьшить риски, связанные с выходом из строя части системы.

Обратите внимание на рисунки 13, 14: время, необходимое для того, чтобы отправить событие фиксации блока из идентификатора, для абонента считается незначительным, поскольку оба они находятся на одной виртуальной машине и как таковые задержки в сети отсутствуют. Планировщик транзакций может быть либо

последовательный, или параллельный. При использовании последовательного выполнение транзакции блокируется до тех пор, пока не завершится выполнение предыдущей транзакции, при параллельном планировании транзакций последующая транзакция, не имеющая зависимостей от текущей транзакции, может выполняться параллельно с текущей транзакцией, отправленной на исполнение.

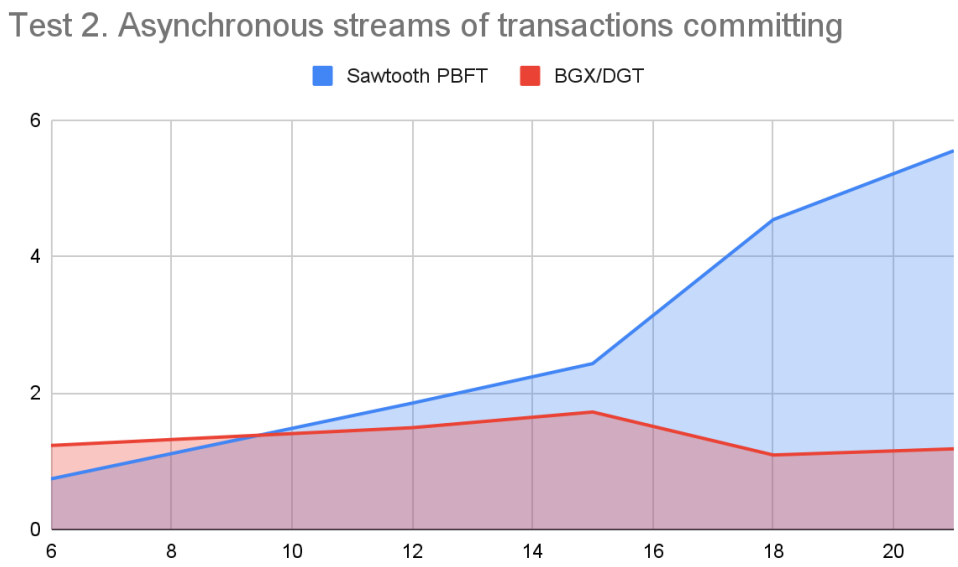

Рисунок 13 - Сравнения первоначальных транзакций BGX/DGT с платформой Sawtooth PBFT на пропускную способность с повышением узлов связи

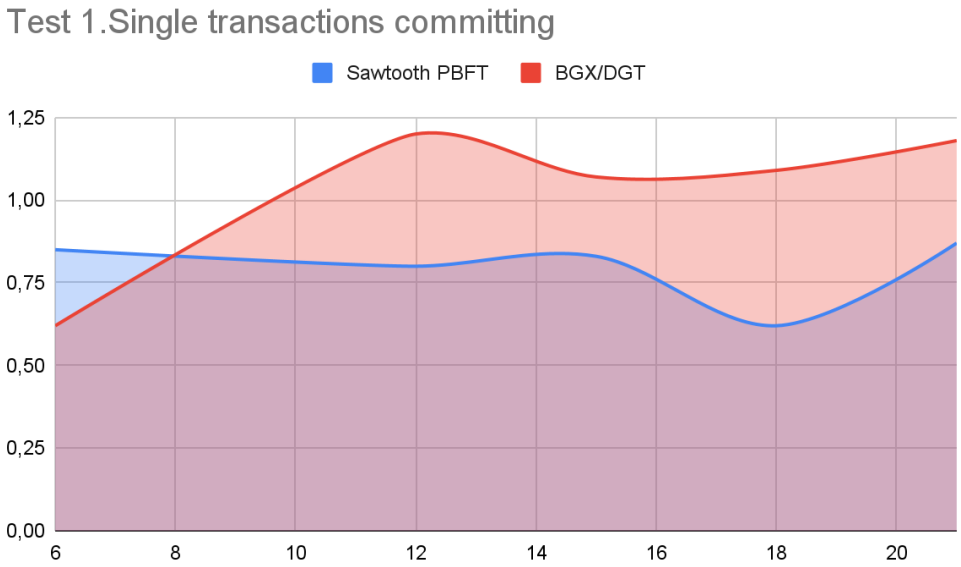

Рисунок 14 - Сравнения асинхронные потоков транзакций BGX/DGT с платформой Sawtooth PBFT на пропускную способность с повышением узлов связи

3.2. Формирование нового уровня проверки в консенсусе P-BFT, доступ разграничения транзакций

Консенсус PBFT (Practical Byzantine Fault Tolerance) может предоставить механизм для обмена репликами файлов друг с другом, чтобы каждая копия оставалась непротиворечивой даже в случае повреждения. В качестве альтернативы, в Биткойне упорядочение происходит посредством процесса, называемого майнингом, когда конкурирующие компьютеры соревнуются, чтобы решить криптографическую головоломку, определяющую порядок, на котором впоследствии строятся все процессы.

Каждый узел должен обмениваться данными с другими узлами для обеспечения безопасности сети, из чего следует: по мере увеличения числа узлов увеличиваются огромные накладные расходы на сеть. Консенсус PBFT отлично работает с небольшими группами. Модель PBFT уязвима для атак Sybil, в которых большое количество узлов может быть использовано одной стороной в сети, что ставит под угрозу безопасность. Угроза может быть уменьшена путем увеличения размера сети, но механизм PBFT не поддерживает большие сети по вышеуказанной причине. Таким образом, используя его в сочетании с другим механизмом консенсуса, его можно оптимизировать [75].

Инфраструктура открытых ключей используется для создания криптографических сертификатов, которые привязаны к организациям, сетевым компонентам и конечным пользователям или клиентским приложениям. В результате контролем доступа к данным можно управлять в более широкой сети и на уровне каналов. Этот подход помогает решать вопросы, в которых конфиденциальность имеет первостепенное значение.

Сеть блокчейн – это техническая инфраструктура, которая предоставляет приложениям услуги бухгалтерского учета и смарт-контрактов (которые упакованы как часть «chaincode»). В первую очередь, смарт-контракты используются для генерации транзакций, которые впоследствии распределяются между всеми одноранговыми узлами в сети, где они неизменно записываются в их

копию реестра. Пользователи приложений могут быть конечными пользователями, использующими клиентские приложения, или администраторами сети блокчейн.

Обработки транзакций изначально проходили 2 уровня, на первом уровне в кластере развертывались N узлов и проводилось голосования. По окончанию голосования результаты передавались лидеру узла на последующий, 2-этап голосования, на 3 уровне нотариальный узел с помощью алгоритма RAFT подтверждал транзакции и добавлял в новою цепочку блоков (рисунок 15).

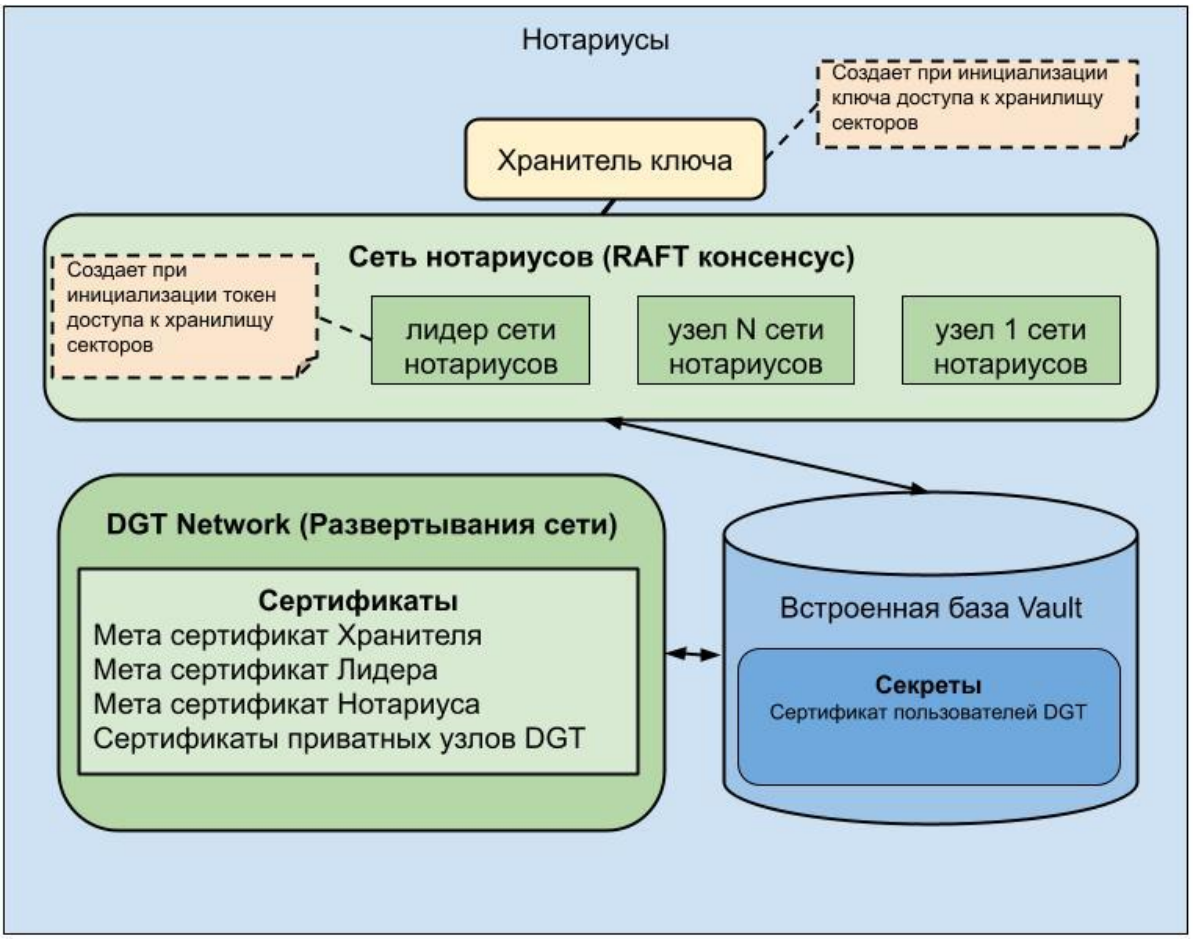

Рисунок 15 - Подтверждения данных с помощью нотариального узла

Алгоритм Raft состоит из двух фаз: выборы лидера и репликация журнала операций. В случае выхода лидера из строя выполнение алгоритма повторяется с первой фазы. В каждый момент времени узел находится в одном из трех состояний: лидер, кандидат или последователь.

Алгоритм Raft нумеруется последовательными целыми числами. Каждый этап начинается с выборов, на которых один или несколько кандидатов пытаются

стать лидером. Если кандидат побеждает на выборах, то он остается лидером до конца срока. Переход состояния показан на рисунке 15 [3]. Все узлы начинаются с состояния ведомого. Если последователь не получает известий от лидера в течение определенного периода времени, то он становится кандидатом. Затем кандидат запрашивает голоса других узлов, чтобы стать лидером. Другие узлы ответят на запрос голосования. Если кандидат получает голоса большинства узлов, он становится лидером. Этот процесс называется выборами лидера. В частности, если последователь получает сигнал пульса в течение минимального тайм-аута выборов от текущего лидера, он не отдает свой голос за кандидата. Это помогает максимизировать продолжительность работы лидера и избежать частых сбоев из-за некоторых изолированных узлов. При нормальной работе Raft есть ровно один лидер, а все остальные узлы являются последователями. Лидер периодически отправляет всем своих последователей, чтобы поддерживать авторитет. Все транзакции в течение этого срока проходят через лидера [76]. Каждая транзакция добавляется как запись в реестр узла.

В алгоритме Raft есть несколько настроек времени ожидания. Один из них контролирует избирательный процесс. Тайм-аут выборов — это количество времени, которое нужно подождать последователю, чтобы стать кандидатом. Последователь переходит в состояние кандидата, когда время выборов достигает нуля. Счетчик времени сбрасывается до случайного значения, когда ведомый получает пульс от лидера. Таймеры случайных выборов в Raft помогают снизить вероятность того, что несколько подписчиков одновременно перейдут к кандидатам.

Определим вероятность потери пакетов как p и предположим, что p является постоянным значением для данной сети. Обозначим значение тайм-аута для каждого раунда выборов как E_t, которое изначально равномерно выбирается из диапазона $[a, b]$. Интервал между двумя частотностями равен τ. Приняты дискретная и целочисленная шкалы времени. Таким образом, если ведомый не может последовательно получить $K = [E_t/\tau]$ тактов, то он предполагает, что жизнеспособного лидера нет, и переходит в состояние кандидата, чтобы начать

выборы. Заметим, что $K \in \{K_1,K_2,...,K_r\}$, K – равномерно выбранная из множества $\{K_1,K_2,...,K_r\}$, где $K_1 = [a/\tau]$ и $K_r = [b/\tau]$. В последующем анализе K обозначает максимальное количество тактовых импульсов для выборного счетчика до тайм-аута.

Пусть $g(n)$ — стохастический процесс, представляющий состояние стадии $\{1, 2,...,r\}$ данного узла в момент времени n. Пусть $b(n)$ — стохастический процесс, представляющий левые шаги счетчика времени выбора для узла в момент времени n. Как только предполагается независимость между $g(n)$ и $b(n)$, мы можем смоделировать это как двумерный процесс $\{g(n), b(n)\}$.

Примем краткое обозначение: $P\{i, k_i - 1|i, k_i\} = P\{g(n + 1) = i, b(n + 1) = k_i - 1|g(n) = i, b(n) = k_i\}$. В этой цепи Маркова единственными ненулевыми одношаговыми переходными вероятностями являются

$$P\{i, k_i - 1|i, k_i\} = p \tag{9}$$

$$P\{i, K_i|i, k_i\} = (1 - p)/r \tag{10}$$

$$P\{i, 0|i, 0\} = 1 \tag{11}$$

где $i, j = 1, 2, ... ,r$ и $k_i \in \{1,...,K_i\}$.

Уравнение (9) основано на том факте, что ведомый не получает тактов от текущего лидера, и его счетчик времени выбора уменьшается на 1. Уравнение (10) показывает тот факт, что ведомый получает пульс и сбрасывает счетчик времени выборов. Уравнение (11) показывает, что как только счетчик времени выбора достигает нуля, ведомый переходит в состояние кандидата.

Обозначим $\{i, 0\}$ как поглощающее состояние. Поскольку $i = 1, 2,...,r$, имеется r поглощающих состояний. Обозначим другие состояния, кроме состояния $\{i, 0\}$ в пространстве состояний $\{g(n), b(n)\}$, как переходные состояния. Имеется t переходных состояний, где $t = \sum_{i=1}^{r} K_i$. Упорядочим состояния так, чтобы первые t состояний были переходными, а последние r состояний поглощающими. Матрица перехода имеет следующий канонический вид:

$$P = \begin{pmatrix} Q & R \\ 0 & I \end{pmatrix} \tag{12}$$

где Q — матрица размера $t \times t$, R — ненулевая матрица размера $t \times r$, 0 — нулевая матрица размера $r \times t$, I — единичная матрица размера $r \times r$.

В частности, вход q_{ij} в Q определяется как вероятность перехода из переходного состояния S_i в переходное состояние S_j, а вход r_{mn} в R определяется как вероятность перехода из переходного состояния S_m в поглощающее состояние S_n.

Когда $K_r - K_1 < K_1$ или $b - a < \tau$, значение тайм-аута выбора имеет только одно значение. Таким образом, $r = 1$ и $t = K$, и тогда единственные ненулевые вероятности одношагового перехода в (9)–(11) можно упростить следующим образом:

$$P\{k - 1|\, k\} = p \tag{13}$$

$$P\{K|k\} = 1 - p \tag{14}$$

$$P\{0|0\} = 1 \tag{15}$$

где $k \in \{1,...,K\}$.

Таким образом, матрица перехода $\boldsymbol{P}$ становится матрицей $(K + 1) \times (K + 1)$ как

$$P = \begin{pmatrix} 1-p & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} \tag{16}$$

На n шаге матрица перехода есть P^n, а элемент $p_{ij}^{(n)}$матрицы P^n есть вероятность оказаться в состоянии S_j из состояния S_i.

Для простоты предположим, что счетчик времени ожидания выборов имеет фиксированное значение K в следующем анализе. Несложно распространить аналитические результаты на случай, когда тайм-аут выборов является случайным значением $r > 1$.

Выводы к главе 3

Многоуровневая сеть на основе консенсуса P-BFT (Practical Byzantine Fault Tolerance) и RAFT (Replicated State Machine Protocol) — это система, которая использует комбинацию двух протоколов консенсуса для обеспечения высокой отказоустойчивости и достоверности данных в распределенной среде. P-BFT - это протокол консенсуса, который обеспечивает отказоустойчивость в условиях наличия вредоносных атак на сеть. Он позволяет распределенным узлам достигнуть согласия относительно состояния системы, используя алгоритм, который обрабатывает отдельные блоки данных в сериях. Протокол P-BFT обеспечивает надежную доставку данных и защиту от дублирования блоков данных.

RAFT — это другой протокол консенсуса, который также используется для обеспечения достоверности данных в распределенной среде. Он основан на модели реплицированного состояния, где каждый узел в системе содержит полную копию данных. Реплики используют протокол голосования для определения правильного порядка выполнения операций и подтверждения достижения консенсуса. Многоуровневая сеть, использующая комбинацию протоколов P-BFT и RAFT, обычно состоит из нескольких уровней. На каждом уровне используются отдельные экземпляры протоколов консенсуса, которые обрабатывают блоки данных на этом уровне. Это позволяет достичь более высокой отказоустойчивости и надежности данных, так как система может обнаруживать и исправлять ошибки на более ранних стадиях.

Сети первого уровня могут использовать протокол P-BFT для обеспечения высокой скорости обработки транзакций, а второго уровня - протокол RAFT для обеспечения высокой надежности данных и защиты от вредоносных атак. Это позволяет достичь максимальной отказоустойчивости и надежности данных в распределенной среде.

ГЛАВА 4. РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

4.1. Анализ многоуровневого доступа обработки транзакций

Многоуровневая сеть на основе комбинации протоколов P-BFT и RAFT состоит из нескольких уровней. На верхнем уровне используется P-BFT для выбора лидера, который будет координировать работу сети на следующем уровне. На следующем уровне используется RAFT, который позволяет лидеру этого уровня координировать работу узлов на более низких уровнях. Каждый уровень может иметь своего лидер, который выбирается методом голосования. Таким образом, комбинация протоколов P-BFT и RAFT позволяет достичь высокой степени безопасности и отказоустойчивости в крупных сетях при более низкой стоимости коммуникации и обработки сообщений.

Многоуровневые сети на основе комбинации протоколов P-BFT и RAFT показали следующие значения по результатам:

1. Высокая производительность: многоуровневая сеть на основе протоколов P-BFT и RAFT показала высокую производительность в обработке транзакций и достижении консенсуса в распределенной среде. Исследователи отметили, что использование комбинации протоколов позволяет достичь высокой производительности и отказоустойчивости в различных условиях.
2. Надежность данных: многоуровневые сети на основе комбинации протоколов P-BFT и RAFT также обеспечивают высокую надежность данных в распределенной среде. Исследования показали, что системы, использующие эту комбинацию протоколов, могут обнаруживать и исправлять ошибки на более ранних этапах, что повышает надежность данных и уменьшает вероятность возникновения проблем.

3. Снижение нагрузки на сеть: использование комбинации протоколов P-BFT и RAFT может снизить нагрузку на сеть в распределенных системах. Это связано с тем, что протоколы могут использовать различные методы передачи данных и обработки транзакций, что позволяет снизить объем передаваемых данных и ускорить процесс обработки.
4. Устойчивость к вредоносным атакам: многоуровневые сети на основе комбинации протоколов P-BFT и RAFT обладают высокой устойчивостью к вредоносным атакам. Исследования показали, что системы, использующие эту комбинацию протоколов, могут эффективно защититься от атак типа Sybil и DDoS, что повышает безопасность данных в распределенных системах.

Общие результаты исследований показывают, что комбинация протоколов P-BFT и RAFT может быть эффективным инструментом для обеспечения высокой отказоустойчивости, надежности данных и безопасности в распределенных системах.

Численный метод многоуровневой сети на основе комбинации протоколов P-BFT и RAFT может быть реализован с помощью алгоритмов, обеспечивающих высокую скорость обработки транзакций и надежность данных в распределенной среде. Он выполняет следующие шаги:

1. Создание многоуровневой сети на основе комбинации протоколов P-BFT и RAFT с несколькими узлами и уровнями.
2. Разделение узлов на группы, которые могут работать независимо друг от друга и обрабатывать транзакции параллельно.
3. Реализация механизма консенсуса, основанного на протоколах P-BFT и RAFT, который позволяет достигать единства среди узлов и сохранять надежность данных.
4. Оптимизация процесса обработки транзакций, например, путем использования параллельных вычислений и распределения нагрузки между узлами

5. Реализация механизмов защиты от различных видов атак, включая атаки типа DoS и атаки типа "манипуляция с транзакциями", для обеспечения безопасности и надежности сети.
6. Реализация механизмов мониторинга и управления сетью, включая средства анализа и отслеживания производительности, механизмы обнаружения и исправления ошибок и проблем сети, а также механизмы обновления и масштабирования сети.
7. Обеспечение достаточного уровня масштабируемости и гибкости сети, чтобы она могла адаптироваться к изменяющимся потребностям пользователей и изменяющимся условиям работы в распределенной среде.
8. Реализация механизмов хранения и управления данными в сети, включая механизмы резервного копирования, восстановления и синхронизации данных между узлами.
9. Разработка и реализация интерфейсов и приложений, позволяющих пользователям взаимодействовать с сетью и использовать ее для обмена данными и проведения транзакций.
10. Проведение тестирования и оптимизации работы сети, включая тестирование на прочность, производительность, надежность и безопасность, а также поиск и устранение ошибок и уязвимостей.

Оптимизированный византийский алгоритм отказоустойчивости для блокчейн консорциума на основе комбинации протоколов P-BFT и RAFT представляет собой решение для обеспечения высокой отказоустойчивости блокчейн-системы в условиях распределенной среды. Он основан на сочетании протоколов P-BFT и RAFT, которые обеспечивают надежное достижение консенсуса и устойчивость к сбоям.

Алгоритм работает следующим образом: при возникновении нового блока в блокчейне, узлы сети отправляют свои голоса за данный блок. Если достигнут кворум по голосам, блок с наивысшим номером, полученный от лидера, выбирается как новый текущий блок. Если же кворум не был достигнут, то

применяется протокол RAFT для выбора нового лидера, который затем будет продолжать процесс выбора блоков.

Такой подход к отказоустойчивости блокчейн консорциума на основе комбинации протоколов P-BFT и RAFT позволяет обеспечить высокую скорость обработки транзакций и достижение консенсуса при соблюдении условий надежной работы сети. Он также обеспечивает отказоустойчивость блокчейн-системы в условиях распределенной среды и устойчивость к сбоям.

4.2. Комбинированный подход обработки транзакций

Математическая модель многоуровневой сети, построенной на комбинации протоколов P-BFT и RAFT, может быть представлена в виде системы уравнений, описывающих процессы обмена сообщениями между узлами сети и принятие решений о консенсусе.

Одна из возможных математических моделей такой сети может быть построена на основе следующих уравнений:

1. Уравнения для обмена сообщениями между узлами сети:

- Сообщение запроса блока от узла i к узлу j: $M_{ij} = (i, j, n, H)$, где M_{ij} - сообщение, i и j - идентификаторы узлов, n - номер блока, H - хеш блока.
- Сообщение ответа на запрос блока от узла j к узлу i: $M_{ji} = (j, i, n, H, D)$, где D - данные блока.
- Сообщение запроса голоса от узла i к узлу j: $M_{ij} = (i, j, n, V)$, где V - голос узла i за блок с номером n.
- Сообщение ответа на запрос голоса от узла j к узлу i: $M_{ji} = (j, i, n, V)$, где V - голос узла j за блок с номером n.
- Сообщение запроса консенсуса от узла $M_{ij} = (i, j_1, j_2, \dots, j_k, n)$, где n - номер блока.

- Сообщение ответа на запрос консенсуса от узлов $j_{1,}, j_2, \dots, j_{k,}$к узлу i i: $M_{ij} = (j_1, j_2, \dots, j_k, i, V)$, где V - решение об утверждении блока с номером n.

2. Уравнения для принятия решений о консенсусе:

- Условие достижения кворума по голосам узлов: $|V| > f$, где $|V|$ - число голосов, f - число ошибок, которые может допустить система.
- Условие достижения кворума по ответам на запросы блоков: $|D| > f$, где $|D|$ - число блоков, полученных от других узлов, f - число ошибок.
- Условие достижения кворума по ответам на запросы консенсуса: $|V| > f$, где $|V|$ - число голосов, полученных от других узлов, f - число ошибок;
- Функция выбора лидера, отвечающего за сбор и обработку голосов.

3. Алгоритм выбора лидера в протоколе RAFT:

- Каждый узел начинает в состоянии “подозревающего” (suspect);
- Узел может перейти в состояние "просматривающего" (viewing), если получает сообщение с более высоким номером предлагаемого лидера;
- Узел переходит в состояние "подтверждающего" (confirming), если он получает сообщения от большинства узлов в состоянии "просматривающего";
- Узел переходит в состояние "лидера" (leader), если он получает подтверждение от большинства узлов в состоянии "подтверждающего";
- Если узел не получает подтверждения в течение некоторого времени, он начинает новый раунд выборов, увеличивая номер предлагаемого лидера.

4. Алгоритм формирования блока:

- Узел собирает транзакции в пуле транзакций;
- Узел добавляет заголовок блока, включающий номер предыдущего блока и хеш предыдущего блока;
- Узел рассчитывает хеш текущего блока на основе его заголовка и транзакций;

- Узел добавляет хеш текущего блока в заголовок и формирует окончательный блок.

5. Алгоритм проверки блока:

- Узел проверяет, что номер предыдущего блока в заголовке соответствует номеру предыдущего блока в цепочке блоков;
- Узел проверяет, что хеш предыдущего блока в заголовке соответствует хешу предыдущего блока в цепочке блоков;
- Узел проверяет, что хеш текущего блока в заголовке соответствует хешу текущего блока, рассчитанному на основе заголовка и транзакций;
- Узел проверяет, что все транзакции в блоке корректны;
- Если проверка проходит успешно, узел принимает блок и добавляет его в цепочку блоков.

Функция выбора решения, основанная на комбинации протоколов P-BFT и RAFT: если был достигнут кворум по голосам, то выбирается блок с наивысшим номером, полученный от лидера; если кворум не был достигнут, то используется алгоритм RAFT для выбора лидера и последующего принятия решения.

Функция выбора решения, основанная на комбинации протоколов P-BFT и RAFT, имеет следующий алгоритм:

1. Узел-лидер предлагает новый блок для добавления в блокчейн.
2. Узлы в сети голосуют за блок: каждый узел отправляет запрос голоса другим узлам в сети.
3. Если был достигнут кворум по голосам, т. е. число голосов превышает число ошибок f, то выбирается блок с наивысшим номером, полученный от лидера. Если же кворум не был достигнут, то переходим к шагу 4.
4. Используется алгоритм RAFT для выбора нового лидера. Узлы проводят выборы лидера с помощью протокола RAFT.
5. Новый лидер предлагает блок для принятия решения.
6. Этот блок может быть принят в качестве правильного решения, если большинство участников его поддерживает.

Таким образом, функция выбора решения комбинирует преимущества двух протоколов согласования и позволяет выбирать правильное решение, даже если не все участники согласны с ним. Если кворум был достигнут, то выбирается блок с наивысшим номером, который был предложен лидером. Если же кворум не был достигнут, то выбирается новый лидер с помощью протокола RAFT, после чего этот лидер предлагает новый блок для принятия решения.

Эта математическая модель позволяет описать процессы, происходящие в многоуровневой сети, построенной на комбинации протоколов P-BFT и RAFT, и позволяет определить условия, необходимые для принятия решения о консенсусе. Она также демонстрирует, как эти протоколы могут работать вместе, чтобы обеспечить надежную и эффективную систему для достижения консенсуса в распределенной среде.

Важным аспектом многоуровневой сети на основе комбинации протоколов P-BFT и RAFT является поддержка сертификатов безопасности. Сертификаты безопасности используются для обеспечения безопасности в процессе выбора лидера и принятия решений. Сертификаты безопасности могут использоваться для аутентификации участников сети, обеспечивая тем самым защиту от возможных атак, связанных с подменой или подделкой сообщений. Кроме того, сертификаты могут быть использованы для обеспечения безопасности процесса выбора лидера, гарантируя, что только доверенные участники могут участвовать в процессе выбора лидера.

В многоуровневой сети на основе комбинации протоколов P-BFT и RAFT, сертификаты могут быть использованы для аутентификации участников каждого уровня сети, а также для обеспечения безопасности процесса выбора лидера и принятия решений на каждом уровне. Кроме того, сертификаты могут быть использованы для обеспечения конфиденциальности и целостности данных, передаваемых между участниками сети.

Поддержка сертификатов безопасности является важным аспектом многоуровневой сети на основе комбинации протоколов P-BFT и RAFT,

обеспечивая безопасность процесса выбора лидера и принятия решений на каждом уровне сети.

Выводы к главе 4

Многоуровневая сеть на основе комбинации протоколов P-BFT и RAFT может включать в себя следующие шаги:

1. Настройка многоуровневой сети: в зависимости от конкретной задачи необходимо выбрать оптимальную конфигурацию сети, включая количество уровней, количество узлов на каждом уровне и параметры протоколов P-BFT и RAFT.
2. Распределение узлов по уровням: узлы должны быть распределены по уровням таким образом, чтобы каждый уровень содержал определенное количество узлов, необходимое для достижения необходимой производительности и надежности.
3. Инициализация сети: при запуске сети необходимо произвести инициализацию каждого узла и установить соединение между узлами.
4. Обработка транзакций: для обработки транзакций используются протоколы P-BFT и RAFT. При этом каждый уровень обрабатывает транзакции с определенным уровнем сложности. Транзакции обрабатываются по протоколу P-BFT, который обеспечивает высокую скорость обработки и надежность транзакций.
5. Синхронизация данных: чтобы обеспечить согласованность данных на всех уровнях сети, используется протокол RAFT, который обеспечивает синхронизацию данных и обнаружение ошибок. При этом каждый уровень имеет свой независимый журнал транзакций, который синхронизируется с журналами на других уровнях.
6. Обработка ошибок: при возникновении ошибок в сети, таких как отказ узла или конфликты данных, протокол RAFT обеспечивает автоматическое восстановление сети и решение конфликтов.
7. Мониторинг и анализ: для оптимизации производительности и надежности сети необходимо постоянно мониторить ее состояние и проводить анализ результатов работы

ЗАКЛЮЧЕНИЕ

По результатам проведенных исследований можно сделать следующие выводы:

1. Создание виртуальных машин, реализующих операционное окружение системы блокчейн, является важным шагом в разработке гетерогенных программно-аппаратных комплексов.
2. Разработанная методология запуска приложений в многоуровневых виртуальных средах действительно позволяет повысить общую производительность таких комплексов.
3. Разработанный подход к построению операционного окружения пользовательской подсистемы обеспечивает безопасный доступ пользователей к ресурсоемким приложениям в гетерогенной распределенной облачной вычислительной среде.
4. Исследование методов повышения надежности аутентификации и авторизации и разработанная методика их применения в гетерогенной облачной среде являются важными компонентами для обеспечения безопасности работы таких комплексов.

Таким образом анализ и разработка блокчейн-систем существенно ускоряет процесс обработки данных.

Список литературы

1. Bogdanov, A. Risk model of application of lifting methods [Электронный ресурс] / A. Bogdanov, A. Degtyarev, N. Shchegoleva [et al.] // The 9th International Conference «Distributed Computing and Grid Technologies in Science and Education» on. — 2021. — Vol. 3041. P. 369–374. — Режим доступа: https://doi.org/10.54546/mlit.2021.77.69.001 (дата обращения: 24.06.2023).

2. Bogdanov, A. Solving the problems of byzantine generals [Электронный ресурс] / A. Bogdanov, A. Degtyarev, N. Shchegoleva [et al.] // The 9th International Conference «Distributed Computing and Grid Technologies in Science and Education» on. — 2021. — Vol. 3041. P. 573–578. — Режим доступа: https://doi.org/10.54546/mlit.2021.72.42.001 (дата обращения: 24.06.2023).

3. Bogdanov, A. A Multilayer Approach to the Security of Blockchain Networks of the Future Bogdanov / A. Bogdanov, A. Degtyarev, N. Shchegoleva [et al.] // ICCSA 2022 Workshops. — [s. l.] : Springer, 2022. — Vol. 13377. — P. 205–216.

4. Bogdanov, A. Protection of Personal Data Using Anonymization / A. Bogdanov, A. Degtyarev, N. Shchegoleva [et al.] // Computational Science and Its Applications. ICCSA 2021. — [s. l.] : Springer, 2021. — Vol. 12956. — P. 447–459.

5. Bogdanov, A. Testing and Comparative Analysis of the F-BFT-based DLT Solution / A. Bogdanov, N. Shchegoleva, V. Korkhov [et al.] // International Conference on Computational Science and Its Applications. — [s. l.] : Springer, 2021. — Vol. 12952. — P. 31–41.

6. Bogdanov, A. Digitalization of health care: what can be done now / A. Bogdanov, N. Zalutskaya, N. Schegoleva [et al.] // Information Society Journal. — 2022. — P. 58–70

7. Bogdanov, A. Comparative analysis and applicability determination for several dlt solutions [Электронный ресурс] /A. Bogdanov, V. Korkhov, N. Shchegoleva [et al.] // The 9th International Conference «Distributed Computing and Grid Technologies in Science and Education» on. — 2021. — Режим доступа: https://doi.org/10.54546/mlit.2021.13.56.001 (дата обращения: 24.06.2023).

8. Imoize, A. L. 6g enabled smart infrastructure for sustainable society: Opportunities, challenges, and research roadmap / A. L. Imoize, O. Adedeji, N. Tandiya, S. Shetty — Sensors, 2021. — Vol. 21. — I. 5.

9. Musleh, A. S. Blockchain Applications in Smart Grid–Review and Frameworks / A. S. Musleh, G. Yao, S. M. Muyeen // IEEE., — 2019. — Vol. 7. — P. 86746–86757.

10. Aijaz, A. Private 5G: The Future of Industrial Wireless / IEEE Industrial Electronics Magazine. — 2020. — Vol. 14. - №4 — P. 136–145.

11. Anshuman Kalla. A survey on the use of blockchain for future 6G: Technical aspects, use cases, challenges and research directions [Электронный ресурс] / Anshuman Kalla Chamitha de Alwis, Pawani Porambage, Gürkan Gür, Madhusanka Liyanage. — Journal of Industrial Information Integration. — 2022. — P. 100404. — Режим доступа: url: https://doi.org/10.1016/j.jii.2022.100404 (дата обращения: 24.06.2023).

12. Aazhang, B. Key Drivers and Research Challenges for 6G Ubiquitous Wireless Intelligence / Matti Latva-aho, Kari Leppänen (eds.) — Finland : University of Oulu, 2019. — 36 p.

13. Burtyka, Ph. Batch Symmetric Fully Homomorphic / “Proceedings of ISP RAS” journal. — 2014. — Vol. 26. — I. 5. — P. 99–116.

14. Graham, C. Anonymisation: managing data protection risk code of practice / C. Graham. — [s. l.] : Information Commissioner’s Office, 2012. — P. 108.

15. Benzaid, C. AI-driven Zero Touch Network and Service Management in 5G and Beyond: Challenges and Research Directions / C. Benzaid, T. Taleb. — IEEE Network. — 2020. —Vol. 34. — № 2 — P. 186–194.

16. Gaber, C. Liability-aware security management for 5G / C. Gaber, J. S. Vilchez, G. Gür [et al.] // 2020 IEEE 3rd 5G World Forum (5GWF). — IEEE, 2020. — P. 133–138.

17. Hu, C. A novel blockchain-based anonymous handover authentication scheme in mobile networks / C. Hu, D. Zheng, R. Guo, A. Wu, L. Wang, S. Gao. — International Journal of Network Security. — 2020. — Vol. 22. — №5 — P. 874–884.

18. Children’s Hospital of Eastern Ontario Research Institute Pan-Canadian De-Identification Guidelines for Personal Health Information. — Canada : Office of the Privacy Commissioner of Canada, 2007. — 87 p.

19. DGT ONE PAGER [Электронный ресурс] //DGT. — Режим доступа: https://dgt.world/docs/DGT_OnePager.pdf. (дата обращения: 10.05.2023)

20. DGT The Blockchain Handbook [Электронный ресурс] //DGT. —Режим доступа: https://dgt.world/docs/DGT_BLOCKCHAIN_ABC.pdf (дата обращения: 10.05.2023)

21. ENISA Data Pseudonymization: Advanced Techniques & Use Case / Athena Bourka (ENISA), Prokopios Drogkaris (ENISA) (eds.). — [s. l.] : ENISA, 2021. — 53 p.

22. Hu, F. Full spectrum sharing in cognitive radio networks toward 5G: A survey / F. Hu, B. Chen, K. Zhu // IEEE Access. — 2021. — V. 6. — P. 15754–15776.

23. Fraser, R. Tools for de-identification of personal health information / R. Fraser, D. Willison. — Canada : Pan Canadian Health Information Privacy (HIP), 2009. — 40 p.

24. Dik, G. Challenges of IoT Identification and Multi-Level Protection in Integrated Data Transmission Networks Based on 5G/6G Technologies / G. Dik, A. Bogdanov, N. Shchegoleva [et al.] — Computers. — 2022. — Vol. 11. — № 12. — 178 p.

25. Goverment of Canada Personal Information Protection and Electronic Documents Ac: [Нормативный акт, регулирующий обращение с конфиденциальной информацией в Канаде: принят Праламентом Канады 13 апреля 2000 г. по состоянию на 2023 г.] — Canada: Minister of Justice, 2023.

26. GSMA Securing the 5g era [Электронный ресурс] / GSMA. — Режим доступа: https://www.gsma.com/security/securing-the-5g-era/. (дата обращения: 03.05.2021)

27. Zhang, H. Blockchain-based trust management for internet of vehicles / H. Zhang, J. Liu, H. Zhao [et al.] // IEEE Transactions on Emerging Topics in Computing. — 2021.- Vol. 9. — № 3 — P. 1397–1409.

28. Haverinen, J. Extensible authentication protocol method for 3rd generation authentication and key agreement / J. Haverinen, H. Arkko. — USA : RFC Editor, 2006. — 79 p.

29. “Best Practice” Guidelines for Managing the Disclosure of De-Identified Health Information / Health System Use Technical Advisory Committee Data De-Identification Working Group. — Ottawa : Canadian Institute for Health Information, 2010. — 53 p.

30. Ahmad, I. Security for 5G and beyond / I. Ahmad, S. Shahabuddin, T. Kumar [et. al.] // IEEE Communications Surveys & Tutorials. —2019. — Vol. 21. — № 4. — P. 3682–3722.

31. Shayea, I. Key Challenges, Drivers and Solutions for Mobility Management in 5G Networks: A Survey / I. Shayea, M. Ergen, M. H. Azmi [et al] // IEEE Access. — 2020. — Vol. 8. — P. 172534–172552.

32. De-identification Guidelines for Structured Data / Information and Privacy Commissioner of Ontario. — Toronto, Ontario : [s.n.], 2016.

33. ISO 17432:2004 Health informatics. Messages and communication. Web access to DICOM persistent objects. - [s.l.] : Standardinform, 2010.

34. Demertzis, K. Anomaly detection via blockchained deep learning smart contracts in industry 4.0 / K. Demertzis, L. Iliadis, N. Tziritas [et al.]. — Neural Computing and Applications. — 2020. — Vol. 32. — P. 17361–17378.

35. Gai, K. Permissioned blockchain and edge computing empowered privacy-preserving smart grid networks / K. Gai, Y. Wu, L. Zhu [et al.] // IEEE Internet of Things Journal. —2019. — Vol. 6. — P. 7992–8004.

36. Leppänen, K. Key drivers and research challenges for 6G ubiquitous wireless intelligence / K. Leppänen, M. Latva-Aho. — Finland : Oulu, 2019.

37. Liang, Y.-C. Dynamic Spectrum Management / Y.-C. Liang. — [s .l.] : Springer Nature, 2020. — P. 21–27.

38. Crosby, M. Blockchain technology: Beyond bitcoin / M. Crosby, P. Pattanayak, S. Verma [et al.]. — [s.l.] : Applied Innovation, 2016. — Vol. 2. — 16 p.

39. Renzo, M. Di. Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead / M. Di Renzo, A. Zappone, M. Debbah [et al.] // IEEE Journal on Selected Areas in Communication. — 2020. — Vol. 38. — I. 11. — P. 2450–2525.

40. Chowdhury, M. Z. 6G wireless communication systems: Applications, requirements, technologies, challenges, and research directions / M. Z. Chowdhury, M. Shahjalal, S. Ahmed [et al.]. // IEEE Open Journal of the Communications Society — 2020. — Vol. 1. — P. 957–975.

41. Fredrikson, M. Privacy in Pharmacogenetics: An End-to-End Case Study of Personalized Warfarin Dosing / M. Fredrikson, E. Lantz, S. Jha // 23rd USENIX Security Symposium. — San Diego, CA : Security Symposium, 2014. — P. 17–32.

42. Weerasinghe, N. A novel blockchain-as-a-service (baas) platform for local 5g operators / N. Weerasinghe, T. Hewa, M. Liyanage [et al.] // IEEE Open Journal of the Communications Society. — 2021. — Vol. 2. — P. 576–601.

43. Zaynalov, N. Information Security Issues For Travel Companies / N. Zaynalov, A. Muhamadiev, U.bekburodov [et al.] // 2019 International Conference on Information Science and Communications Technologies (ICISCT). — [s.l.] : IEEE, 2019. — P. 1–4.

44. Zaynalov, N. Hiding short message text in the uzbek language / N. Zaynalov, U. Narzullayev, A. Muhamadiev [et al.] // 2020 International Conference on Information Science and Communications Technologies (ICISCT). — Tashkent, Uzbekistan : IEEE, 2020. — P. 1–6.

45. NISTIR 8062 An Introduction to Privacy Engineering and Risk Management / S. Brooks, M. Garcia, N. Lefkovitz [et al.]. — [s. l.] : U.S. Department of Commerce, 2017.

46. Privacy Enhancing Technologies – A Review of Tools and Techniques / Office of the Privacy Commissioner of Canada. — Canada : Office of the Privacy Commissioner of Canada, 2017. — 11 p.

47. Porambage, P. The Roadmap to 6G Security and Privacy / P. Porambage, G. Gur, D. P. M. Osorio [et al.] // IEEE Open Journal of the Communications Society. — 2021. — Vol. 2. — P. 1094–1122.

48. Porambage, P. 6G Security Challenges and Potential Solutions / P. Porambage, G. Gür, D. P. M. Osorio [et al.] // 2021 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit). — [s.l.] : IEEE, 2021. — P. 622–627.

49. Advisory Guidelines on enforcement of the data protection provisions / Personal Data Protect Comission, Singapore. — [s.l.] : PDPC, 2016. − 52 p.

50. General Data Protection Regulation: [regulation of the European Parliament and of the Council]. — [s. l.] : Intersoft Consulting, 2018. — P. 1–5.

51. Hu, S. Blockchain and artificial intelligence for dynamic resource sharing in 6g and beyond / S. Hu, Y.-C. Liang, Z. Xiong [et al.] // IEEE Wireless Communications. — 2021. — Vol. 28. — № 4 — P. 145–151.

52. Kim, S. A Survey of Scalability Solutions on Blockchain / S. Kim, Y. Kwon, S. Cho // 2018 International Conference on Information and Communication Technology Convergence (ICTC). — Jeju, Korea (South) : IEEE, 2018. — P. 1204–1207.

53. Kiyomoto, S. On blockchain-based authorization architecture for beyond-5G mobile services / S. Kiyomoto, A. Basu, M. S. Rahman [et al.] // 2017 12th International Conference for Internet Technology and Secured Transactions (ICITST). — Cambridge, UK : IEEE, 2018. — P. 136–141.

54. Tanwar, S. Machine learning adoption in blockchain-based smart applications: The challenges, and a way forward / S. Tanwar, Q. Bhatia, P. Patel [et al.] // IEEE Access. — 2019. — Vol. 8. — P. 474–488.

55. A survey on mobile edge networks: Convergence of computing, caching and communications / S. Wang, X. Zhang, Y. Zhang [et al.] // IEEE Access. — Vol. 5. — P. 6757–6779.

56. 6G The Next Hyper — Connected Experience for All / Samsung Research. — [s. l.] : Samsung Research, 2020. — 46 p.

57. Shchegoleva, N. New Technologies for Storing and Transferring Personal Data / N. Shchegoleva, N. Zalutskaya, A Dambaeva [et al.] // Computational Science and Its Applications — ICCSA 2022 Workshops. — 2022. — V. 13380. — P. 680–692.

58. Ariyarathna, T. Dynamic Spectrum Access via Smart Contracts on Blockchain / T. Ariyarathna, P. Harankahadeniya, S. Isthikar [et al.] // WCNC. — Marrakesh, Morocco : 2019 IEEE Wireless Communications and Networking Conference (WCNC), 2019. — P. 1–6.

59. Higashino, T. Context Recognition of Humans and Objects by Distributed Zero-Energy IoT Devices / T. Higashino, A. Uchiyama, S. Saruwatari [et al.] // 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS). — Dallas, TX, USA : IEEE, 2019. — P. 1787–1796.

60. Maksymyuk, T. Blockchain-empowered framework for decentralized network management in 6g / T. Maksymyuk, J. Gazda, M. Volosin [et al.] // IEEE Communications Magazine. — 2020. — Vol. 58. — № 9 — P. 86–92.

61. Privacy-aware blockchain innovation for 6g: Challenges and opportunities / T. Nguyen, N. Tran, L. Loven [et al.] // IEEE. — 2020. — P. 1–5.

62. Miao, W. Unlocking the potential of 5g and beyond networks to support massive access of ground and air devices / W. Miao, C. Luo, G. Min [et al.] // IEEE Transactions on Network Science and Engineering. — 2021. — Vol. 8. — № 4 — P. 2825–2836.

63. Saad, W. A vision of 6G wireless systems:Applications, trends, technologies, and open research problems / W. Saad, M. Bennis, M. Chen // IEEE Network. — 2020. — Vol. 34. — № 3 — P. 134–142.

64. Yang, W. A Survey on Blockchain-Based Internet Service Architecture: Requirements, Challenges, Trends, and Future / W. Yang, E. Aghasian, S. Garg [et al.] // IEEE Access. — 2019. — Vol. 7. — P. 75845–75872.

65. Li, X. Network Slicing for 5G: Challenges and Opportunities / X. Li, M. Samaka, H. A. Chan [et al.] // IEEE Internet Computing. — 2017. — Vol. 21. — P. 20–27.

66. Liang, X. Integrating blockchain for data sharing and collaboration in mobile healthcare applications / X. Liang, J. Zhao, S. Shetty [et al.] // 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC). — [s.l.] : IEEE, 2017.

67. Ling, X. Blockchain radio access network (B-RAN): Towards decentralized secure radio access paradigm / X. Ling, J. Wang, T. Bouchoucha [et al.] // IEEE Access. — 2019. — Vol.7. - P. 9714–9723.

68. . Liu Y. Blockchain and Machine Learning for Communications and Networking Systems /F. R. Yu, X. Li [et al.] // IEEE Communications Surveys & Tutorials. — 2020. — Vol. 22. — № 2 — P. 1392–1431.

69. Haddad, Z. Blockchain-based Authentication for 5G Networks / Z. Haddad, M. M. Fouda, M. Mahmoud, M. Abdallah // 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT). — [s.l.] : IEEE, 2020. — P. 189–194.

70. Zhang, Z. 6G wireless networks: Vision, requirements, architecture, and key technologies / Y. Xiao, Z. Ma [et al.] // IEEE Vehicular Technology Magazine. — 2019. — Vol. 14. — № 3 — P. 28–41.

71. Towards artificial intelligence enabled 6G: State of the art, challenges, and opportunities / D. Zhu, S. Zhang // Computer Networks. — 2020. — Vol. 183. — P. 107556.

72. Naing, Y. M. Development of a system for launching resource-intensive applications in a cloudy heterogeneous environment : abstract of Ph.D. diss.... cand. tech. : Sciences: 05.13. 15 / Y. M. Naing. — ETY “LETI” : 2013. — 125 p.

73. Demichev, A. P. Introduction to Grid Technology / A. P. Demichev, V. A. Ilyin, A. P. Kryukov // Preprint NIINP MSU-2007-11/832. — 2007. — 87 p.

74. Paraskevov, A. V. Comparative analysis of legal regulation of personal data protection in russia and abroad [Электронный ресурс]/ A. V. Paraskevov, A. V. Levchenko, Y. A. Cuhol // Scientific journal KubSAU. — 2015. — № 110. — Режим доступа: https://cyberleninka.ru/article/n/sravnitelnyy-analiz-pravovogo-regulirovaniya-zaschity-personalnyh-dannyh-v-rossii-i-za-rubezhom (дата обращения: 23.06.2023).

75. Romanenko, K. S. Features of the development of information systems on the hyperledger blockchain platform / K. S. Romanenko, E. A. Ishukova // Perspektiva— 2021. — 2022. — P. 51–55.

76. Huang, D. Performance analysis of the raft consensus algorithm for private blockchains / D. Huang, X. Ma, S. Zhang //IEEE Transactions on Systems, Man, and Cybernetics: Systems. — 2019. — Vol. 50. — №. 1. — P. 172–181.

77. Свидетельство о государственной регистрации программы для ЭВМ № 2023618607 Российская Федерация. Data Privacy Framework (Система поддержки конфиденциальных данных): № 2023617657 : заявл. 26.04.2023: опубл. 26.04.2023 / А. Г. Дик, Ж. У. Киямов, В. В. Хватов, Н. Л. Щеголева; заявитель Общество с ограниченной ответственностью "АССОЦИАЦИЯ ЦИФРОВЫХ ТЕХНОЛОГИЙ И ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ". – EDN CKPZHW.

ПРИЛОЖЕНИЕ А

Подход определения квази-идентификаторов по индексу верхнего уровня доступа

Платформа больших данных, как правило, состоит из инфраструктурной платформы, платформы хранения структурированных и неструктурированных данных и платформы обработки данных. Поэтому обеспечение защиты платформы больших данных — весьма трудоемкий процесс: необходимо обеспечить безопасность обработки в распределенных программных системах, защиту информации в базах средствами различных СУБД; должны быть защищены данные и журналы транзакций; для контроля доступа и отслеживания ключей нужно предусмотреть управление ключами. Кроме того, для обеспечения надлежащего контекста безопасности и функционирования данных на каждом этапе важно гарантировать легитимность происхождения данных, а для обеспечения их доступности требуется предусмотреть меры противодействия *DoS*-атакам.

Защита средств управления системой больших данных

Средства управления системой больших данных предоставляют широкие возможности для внедрения механизмов безопасности, позволяющих осуществлять в режиме реального времени мониторинг состояния компонентов, управление правилами разграничения доступа, идентификацию источников данных и др. Однако требуются дополнительные меры по защите самих средств управления такой системой — именно они представляют особую ценность для нарушителей.

Международные стандарты и подходы к обезличиванию

Европейский Союз

В Европейском Союзе базовым документом, регулирующим отношения, связанные с ПДн, является Общий регламент защиты персональных данных (GDPR) [50].

GDPR использует термины “псевдонимизация” и “анонимные данные”. Их различие раскрывается посредством положений п. 26 Преамбулы и п. 5 ст. 4. Поскольку “подвергнутые псевдонимизации персональные данные” могут быть “соотнесены с физическим лицом посредством использования дополнительной информации”, то псевдонимизированные данные рассматриваются в качестве ПДн. Анонимные же данные к числу ПДн не относятся, и на них не распространяются соответствующие требования, т. к. они “не относятся к идентифицированному или идентифицируемому физическому лицу” или предоставляются таким образом, что “субъект данных не идентифицируется”. Но для дополнительного различения псевдономизированных и анонимных данных GDPR предписывает “обратить внимание на все объективные факторы, такие как расходы на идентификацию и количество времени, необходимого для идентификации, с учетом имеющихся на момент обработки технологий и технологических разработок”.

Целями применения псевдонимизации могут быть “снижение рисков для субъектов данных” и обеспечение “защиты данных”, наряду с другими техническими и организационными мерами защиты данных. При этом применяемые меры должны соотноситься с оценкой степени риска.

При оценке степени риска необходимо принимать во внимание уровень техники, затраты на реализацию и характер, объем, контекст и цели обработки, а также риск различной вероятности и серьезности для прав и свобод физических лиц. С учетом степени риска контролер и обработчик должны реализовывать соответствующие технические и организационные меры для обеспечения уровня безопасности, соответствующего риску, включая, помимо прочего, при необходимости:

- псевдонимизацию и шифрование личных данных;

- способность обеспечивать постоянную конфиденциальность, целостность, доступность и отказоустойчивость систем обработки и услуг;
- возможность своевременно восстановить доступность и доступ к персональным данным в случае физического или технического инцидента;
- процесс регулярного тестирования, оценки эффективности технических и организационных мер по обеспечению безопасности обработки.

Оценка соответствующего уровня безопасности также должна учитывать возможность случайного или незаконного уничтожения, потери, изменения, несанкционированного раскрытия или доступа к персональным данным, передаваемым, хранящимся или обрабатываемым иным образом.

Развитием положений о псевдонимизации занимается Агентство Европейского союза по кибербезопасности (ENISA). Так, одним из новейших актов ENISA является “Псевдонимизация данных ENISA: Передовые методы и примеры использования. Технический анализ мер кибербезопасности в области защиты данных и конфиденциальности” (январь 2021) [21].

К а н а д а

Канада имеет широкую нормативную базу, касающуюся обезличивания данных, особенно в сфере медицины. Основополагающий акт Канады о ПДн - Закон о защите личной информации и электронных документов (Personal Information Protection and Electronic Documents Act (PIPEDA), 2000 [25]). К числу специальных актов можно отнести:

- общеканадское руководство по обезличиванию личной медицинской информации (Pan-Canadian De-Identification Guidelines for Personal Health Information, 2007) [18],
- инструменты для обезличивания личной медицинской информации (Tools for De-Identification of Personal Health Information (Authored by: Ross Fraser and Don Willison, 2009)) [23],

- руководство по лучшим практикам для управления раскрытием обезличенной медицинской информации (Подготовлено: Рабочая группа по обезличиванию данных Технического консультативного комитета по использованию систем здравоохранения) (‘Best Practice’ Guidelines for Managing the Disclosure of De-Identified Health Information (Prepared by the: Health System Use Technical Advisory Committee Data De-Identification Working Group, October 2010) [29],
- руководящий документ по публичному раскрытию клинической информации (Guidance document on Public Release of Clinical Information (March 12, 2019)) [18],
- руководство по обезличиванию структурированных данных (De-identification Guidelines for Structured Data (June 2016)) [32].

В указанных документах, несмотря на некоторое различие в терминологии, связанное в т. ч. и с разным временем принятия актов, процесс обезличивания рассматривается в связи с двумя аспектами - снижением риска повторной идентификации, с одной стороны, и сохранением полезности данных, с другой, - в целях поиска баланса между ними. Поэтому задачей обезличивания является снижение уровня риска повторной идентификации до минимально возможного уровня, для чего используется процедура оценки степени риска повторной идентификации для каждого конкретного случая.

Предполагается, что владелец массива данных не будет раскрывать массив данных, если уровень риска повторной идентификации выше, чем установленное пороговое значение. Пороговое значение, в свою очередь, обусловливается тремя параметрами:

1) вероятностью попыток повторной идентификации;
2) последствиями успешной повторной идентификации;
3) влиянием обезличивания на полезные свойства данных.

Особое внимание в процессе расчета рисков в отношении общедоступных источников данных уделяется квази-идентификаторам, наличие которых в массиве

данных представляет относительно низкий уровень риска повторной идентификации:

- регион проживания (без присутствия других квази-идентификаторов);
- пол (без присутствия других квази-идентификаторов);
- год рождения (без присутствия других квази-идентификаторов);
- комбинация пола и региона проживания.

Остальные квази-идентификаторы и их комбинации рассматриваются как имеющие достаточно высокий уровень риска деидентификации.

Концептуальная схема обезличивания имеет следующий вид:

1. Определение способа распространения обезличенных данных, круга лиц, имеющих доступ к обезличенным данным, наиболее вероятных лиц, которые попытаются совершить повторное обезличивание;
2. Определение квази-идентификаторов в обезличенном массиве данных и определение наличия в открытых источниках данных, которые вместе с имеющимися в обезличенном массиве квази-идентификаторами могут привести к повторной идентификации;
3. Оценка риска повторной идентификации при помощи применения эвристического подхода к имеющимся квази-идентификаторам и последующее применение к этим квази-идентификаторам различных методов анонимизации;
4. В зависимости от результатов применения вышеперечисленных шагов последними шагами процесса будут являться (a) определение порогового значения уровня риска повторной идентификации и (b) очистка данных, т. е. выделение персональных идентификаторов и их кодирование, удаление или рандомизация.

Пороговое значение допустимого уровня риска определяет минимальный уровень обезличивания, которому должен быть подвергнут массив ПД для того, чтобы полученный массив мог быть признан обезличенным.

Оценка риска может осуществляться на основе количественного (проценты, числовое значение от нуля до одного) или качественного («низкий», «средний», «высокий») подходов. Количественный подход основан на эмпирическом измерении и, следовательно, является более точным, менее субъективным и, как правило, позволяет в большей степени сохранить полезные свойства данных. Так, поскольку прямые идентификаторы существенно влияют на возможность повторной идентификации, то в их отношении риск оценивается как 100%, или 1.0. Прямые идентификаторы должны быть подвержены обезличиванию в обязательном порядке, чтобы уровень риска повторной идентификации был ниже порогового значения.

Риск повторной идентификации для косвенных идентификаторов оценивается на уровне субъектов данных. Например, путем вычисления размеров ячеек, который определяется числом субъектов данных в массиве, обладающих одинаковыми значениями косвенных идентификаторов. Рекомендованное пороговое значение допустимого уровня риска повторной идентификации, равное 0,09, достигается при размере ячейки, равном 11.

В Руководстве по обезличиванию структурированных данных [32] приводится методика подсчета контекстных рисков.

Необходимо определить вероятность трех различных атак или угроз повторной идентификации:

- преднамеренная внутренняя атака,
- непреднамеренная реидентификация индивидуума в наборе данных знакомым,
- утечка данных.

При оценке контекстного риска следует использовать наибольшую из этих вероятностей.

Вероятность того, что получатель данных попытается провести де-обезличивание зависит от двух факторов:

- степень контроля, установленного в соглашении об обмене данными в отношении конфиденциальности и безопасности данных,
- мотивы и возможности получателя в отношении проведения атаки на повторную идентификацию.

Оба эти фактора подразумевают качественную оценку, в результате чего значения находятся в диапазоне "низкий", "средний" или "высокий".

Для оценки первого фактора — степени контроля, установленного в соглашении, - необходимо сравнить принимаемые меры с мерами, предлагаемыми нормативными актами и рекомендациями.

Оценка второго фактора — мотивы получателя — может быть осуществлена с учетом следующих моментов:

- наличие инцидентов при работе получателя;
- причины данных инцидентов;
- наличие у получателя технических знаний и/или финансовых ресурсов для попытки повторной идентификации;
- наличие у получателя доступа к другим частным базам данных или наборам данных, которые могут быть связаны с данными для повторной идентификации;
- уровень контроля конфиденциальности и безопасности в соглашении об обмене данными.

В рассматриваемом руководстве предлагается следующая таблица для оценки вероятности атаки на наборы данных:

Таблица А.1 - Граничные значения вероятностей атак повторной идентификации

Контроль конфиденциальности и безопасности	Мотивы и возможности получателя	Вероятность атаки повторной идентификации
Высокая	Низкая	0,05
	Средняя	0,1

	Высокая	0,2
Средняя	Низкая	0,2
	Средняя	0,3
	Высокая	0,4
Низкая	Низкая	0,4
	Средняя	0,5
	Высокая	0,6

Помимо преднамеренной попытки атаки, получатель данных может также случайно повторно идентифицировать одно или несколько лиц. Вероятность возникновения такой "атаки" равна вероятности того, что случайный получатель знает кого-то из набора данных. Для ее расчета можно использовать следующее уравнение:

$$P = 1 - (1 - p)^m \tag{17}$$

здесь P — это процент людей в населении, имеющих состояние или характеристику, о которых идет речь в наборе данных, а m - количество людей, которых в среднем знает человек. Значение p должно быть определено на основе последних статистических данных о населении. С другой стороны, значение m может варьироваться в зависимости от того, какого рода отношения с человеком требуются для того, чтобы иметь о нем сведения относительно условия или характеристики, обсуждаемой в наборе данных.

Третья атака, которую следует рассмотреть, — это утечка данных. Вероятность такой атаки равна вероятности нарушения безопасности данных на объектах получателя. Для расчета этого значения следует использовать

общедоступные данные о распространенности случаев нарушения данных в соответствующей отрасли получателя.

Общее значение риска повторной идентификации рассчитывается по формуле:

$$P_{general} = P_{data\ risk} * P_{context\ risk} \quad (18)$$

В зависимости от вычисленного показателя риска и необходимости сохранения полезности массива данных применяются различные методы обезличивания и их комбинации.

Российский подход к обезличиванию данных

Основы текущего регулирования обработки персональных данных определяются Федеральным законом от 27 июля 2006 г. N 152-ФЗ "О персональных данных" [4]. Данный закон содержит основные положения о понятии и видах ПД, принципах и условиях обработки ПДн, правах субъектов ПД и обязанностях операторов ПД, мерах обеспечения безопасности ПД, а также закладывает основы подзаконного регулирования данной сферы.

Последние обновление ФЗ о ПД произошло в марте 2021 г., когда вступил в силу Федеральный закон от 30 декабря 2020 г. N 519-ФЗ "О внесении изменений в Федеральный закон "О персональных данных". Главным нововведением стало принятие новой категории ПДн - “персональные данные, разрешенные субъектом персональных данных для распространения”.

Однако ФЗ о ПД практически не затрагивает вопроса об обезличивании данных, ограничиваясь указанием определения данного термина и случаев использования процедуры обезличивания:

- обезличивание обрабатываемых ПДн по достижении целей обработки или в случае утраты необходимости в достижении этих целей, если иное не предусмотрено федеральным законом;
- для обработки персональных данных в статистических или иных исследовательских целях;

- в целях Федерального закона "О проведении эксперимента по установлению специального регулирования в целях создания необходимых условий для разработки и внедрения технологий искусственного интеллекта в субъекте Российской Федерации - городе федерального значения Москве и внесении изменений в статьи 6 и 10 Федерального закона "О персональных данных" .

В отношении непосредственного регулирования процедур обезличивания ПДн можно выделить следующие нормативно-правовые акты:

1. Постановление Правительства РФ от 21 марта 2012 г. N 211 "Об утверждении перечня мер, направленных на обеспечение выполнения обязанностей, предусмотренных Федеральным законом "О персональных данных" и принятыми в соответствии с ним нормативными правовыми актами, операторами, являющимися государственными или муниципальными органами" [49].

   Среди подобных мер, касающихся обезличивания, акт предписывает принимать в государственных и муниципальных органах следующие меры:

   – утвердить правила работы с обезличенными данными в случае обезличивания персональных данных;
   – утвердить перечень должностей служащих государственного или муниципального органа, ответственных за проведение мероприятий по обезличиванию обрабатываемых персональных данных, в случае обезличивания персональных данных;
   – в случаях, установленных нормативными правовыми актами РФ, в соответствии с требованиями и методами, установленными уполномоченным органом по защите прав субъектов персональных данных, осуществляют обезличивание персональных данных, обрабатываемых в информационных системах персональных

данных, в том числе созданных и функционирующих в рамках реализации федеральных целевых программ.

2. Приказ РКН от 5 сентября 2013 г. N 996 "Об утверждении требований и методов по обезличиванию персональных данных".

   Приказ принят в соответствии с подпунктом “3” пункта 1 Постановления Правительства РФ от 21 марта 2012 г. N 211 "Об утверждении перечня мер, направленных на обеспечение выполнения обязанностей, предусмотренных Федеральным законом "О персональных данных" [49] и принятыми в соответствии с ним нормативными правовыми актами, операторами, являющимися государственными или муниципальными органами" и содержит описание:

   - свойств обезличенных данных;
   - требований к свойствам получаемых обезличенных данных;
   - характеристик методов обезличивания ПДн;
   - требований к методам обезличивания;
   - требований к свойствам метода обезличивания;
   - наиболее перспективных и удобных для практического применения методов обезличивания.

Сравнение подходов к обезличиванию

Первоначально на проблему защиты персональных данных на международном уровне обратила внимание Организация по экономическому сотрудничеству и развитию (ОЭСР), принявшая в 1980 г. Директиву о защите неприкосновенности частной жизни и международных обменов персональными данными. В дальнейшем эти принципы были детализированы в Конвенции Совета Европы «Об охране личности в отношении автоматизированной обработки персональных данных» (1981 г.), в Директиве Европейского сообщества о защите граждан в плане обработки информации личного характера от 27 июля 1990 г., в Директиве Европейского Союза и Парламента 95/46/ЕС от 24 октября 1995 г. о защите прав частных лиц применительно к обработке персональных данных и

свободном движении таких данных и Директиве 97/66/ЕС от 15.12.1997 по обработке персональных данных, защите, конфиденциальности в телекоммуникационном секторе.

В отмеченных директивных документах были определены основные принципы организации обработки данных личного характера и обеспечения права граждан на защиту персональных данных:

- данные персонального характера должны быть собраны только для определенных целей и в строгом соответствии с законом;
- данные должны соответствовать требованиям, быть точными, полными и вовремя обновленными;
- цели, для достижения которых собираются и обрабатываются персональные данные, должны быть определены и утверждены до начала деятельности и использоваться только в этих целях;
- в системах учета персональных данных должны быть внедрены механизмы, предотвращающие потери или неправильное использование персональных данных;
- деятельность организаций (как государственных, так и частных), имеющих базы данных, содержащих персональные данные, должна быть открытой;
- держатели данных должны быть подконтрольными для обеспечения соблюдения настоящих принципов, для этих целей должно быть предусмотрено создание независимого контролирующего органа как важного элемента защиты личности при автоматизированной обработке информации личного характера.

Рассмотрим регулирование правил защиты персональных данных на федеральном и региональном уровне в различных странах и наличие органа власти по контролю за соблюдением требований по защите персональных данных. Существует три типа систем правового регулирования: децентрализованная, централизованная и смешанная.

1. Децентрализованная система:
   - отсутствие единого подхода к защите персональных данных в рамках отраслевого законодательства;
   - регламентация защиты персональных данных осуществляется посредством профильных нормативных актов комплексных отраслей законодательства или на разных уровнях власти;
   - акты рекомендательного характера играют значительную роль;
   - отсутствие единого надзорного органа.
2. Централизованная система:
   - прямое действие международных норм, гармонизирующих национальные законодательства государств (Конвенция о защите физических лиц при автоматизированной обработке персональных данных, Директива 95/46/ЕС, Директива 2002/58/ЕС);
   - наличие национальных отраслевых законов, содержащих общеобязательные нормы в отношении защиты персональных данных (например, в Германии Bundesdatenschutzgesetz (BDSG));
   - регулирование обработки персональных данных посредством учреждения единого надзорного ведомства («мегарегулятора»). Примеры: Страны ЕС, Израиль, Мексика, Гонконг, Швейцария, Сингапур.
3. Смешанная система правового регулирования подразумевает наличие одного или нескольких признаков, позволяющих отнести систему правового регулирования защиты персональных данных государства к централизованной или децентрализованной системе.

С практической точки зрения большой интерес вызывает Канадский Закон об охране персональной информации, который предусматривает реальные механизмы защиты персональных данных и реализации права на доступ к сведениям о себе.

Основные выводы в результате сравнительного анализа:

- рассмотренные системы в целом похожи в части адресации вопросов защиты персональных данных. Российский подход носит более консервативный характер, широко толкуя состав персональных данных и оставляя без внимания ряд действий по обеспечению безопасности;
- все рассмотренные подходы поддерживают права субъекта персональных данных о согласии на обработку персональных данных, а также право на забвение;
- состав персональных данных по-разному толкуется в разных законодательных системах. Так, технические параметры (онлайн-идентификаторы), такие как IP или cookies относятся к персональным данным в GDPR и не относятся к таковым в Российской Федерации;
- в GDPR и других западных законах отдельно выделен вопрос об утечке данных: организации обязаны объявлять о таких случаях, в российском законодательстве таких требований нет;
- в отличии от российской законодательной системы, большинство западных систем имеет возможность переноса/передачи данных и ответственности, что делает более адресуемыми вопросы публикации и связанные с этим действия по обезличиванию;
- ответственность за нарушения по законодательству РФ все еще существенно ниже, чем в GDPR или PIPEDA (Канада). Западные аналоги требуют значительно более высоких штрафов, кроме того, имеют экстра-территориальный характер;
- использование таких методов, как псевдонимизация, при котором ряд прямых идентификаторов заменяется абстрактным идентификатором, в российском правовом поле остается все еще не ясным. В требованиях GDPR персональные данные с использованием псевдонима, который можно приписать физическому лицу после использования дополнительной информации, необходимо рассматривать как информацию о физическом лице, которую можно идентифицировать;

- в Российской Федерации, в отличии от сравниваемых подходов, отсутствует категорирование по уровню чувствительности данных, что делает затруднительным оценку понесенного ущерба физическим лицам.

ПРИЛОЖЕНИЕ Б

РОССИЙСКАЯ ФЕДЕРАЦИЯ

СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2023618607

Data Privacy Framework (Система поддержки конфиденциальных данных)

Правообладатель: ***Общество с ограниченной ответственностью "АССОЦИАЦИЯ ЦИФРОВЫХ ТЕХНОЛОГИЙ И ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ" (RU)***

Авторы: ***Дик Александр Геннадьевич (RU), Киямов Жасур Уткирович (UZ), Хватов Валерий Владимирович (RU), Щеголева Надежда Львовна (RU)***

Заявка № **2023617657**
Дата поступления **26 апреля 2023 г.**
Дата государственной регистрации
в Реестре программ для ЭВМ ***26 апреля 2023 г.***

Руководитель Федеральной службы по интеллектуальной собственности

ДОКУМЕНТ ПОДПИСАН ЭЛЕКТРОННОЙ ПОДПИСЬЮ
Сертификат 68b80077e14e40f0a94edbd24145d5c7
Владелец Зубов Юрий Сергеевич
Действителен с 2.03.2022 по 26.05.2023

Ю.С. Зубов

РОССИЙСКАЯ ФЕДЕРАЦИЯ

ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

RU2023618607

ГОСУДАРСТВЕННАЯ РЕГИСТРАЦИЯ ПРОГРАММЫ ДЛЯ ЭВМ

Номер регистрации (свидетельства):
2023618607
Дата регистрации: **26.04.2023**
Номер и дата поступления заявки:
2023617657 26.04.2023
Дата публикации и номер бюллетеня:
26.04.2023 Бюл. № 5
Контактные реквизиты:
n.shchegoleva@spbu.ru

Автор(ы):
Дик Александр Геннадьевич (RU),
Киямов Жасур Уткирович (UZ),
Хватов Валерий Владимирович (RU),
Щеголева Надежда Львовна (RU)
Правообладатель(и):
Общество с ограниченной ответственностью "АССОЦИАЦИЯ ЦИФРОВЫХ ТЕХНОЛОГИЙ И ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ" (RU)

Название программы для ЭВМ:
Data Privacy Framework (Система поддержки конфиденциальных данных)

Реферат:
Программный комплекс предназначен для анализа защиты персональных данных. Для оценки эффективности защиты обезличенного набора данных реализованы следующие программные модули: основанные на деобезличивании хеш-функцией с солью по заданным начальным условиям (атака методом перебора или со словарем на применяемые хеш-функции), применении метрики k-Anonymity для оценки эффективности обезличенного набора данных (поиск минимального значения количества повторяющихся строк атрибутов), использовании метода Мондриана для определения степени приватности в обезличенном наборе данных (разбиение поступающей информации на прямоугольные области с одинаковым значением метрики анонимизации и последующей обработкой до достижения заданного уровня обезличивания в системах обработки больших данных).

Язык программирования: Python

Объем программы для ЭВМ: 1,8 Мб