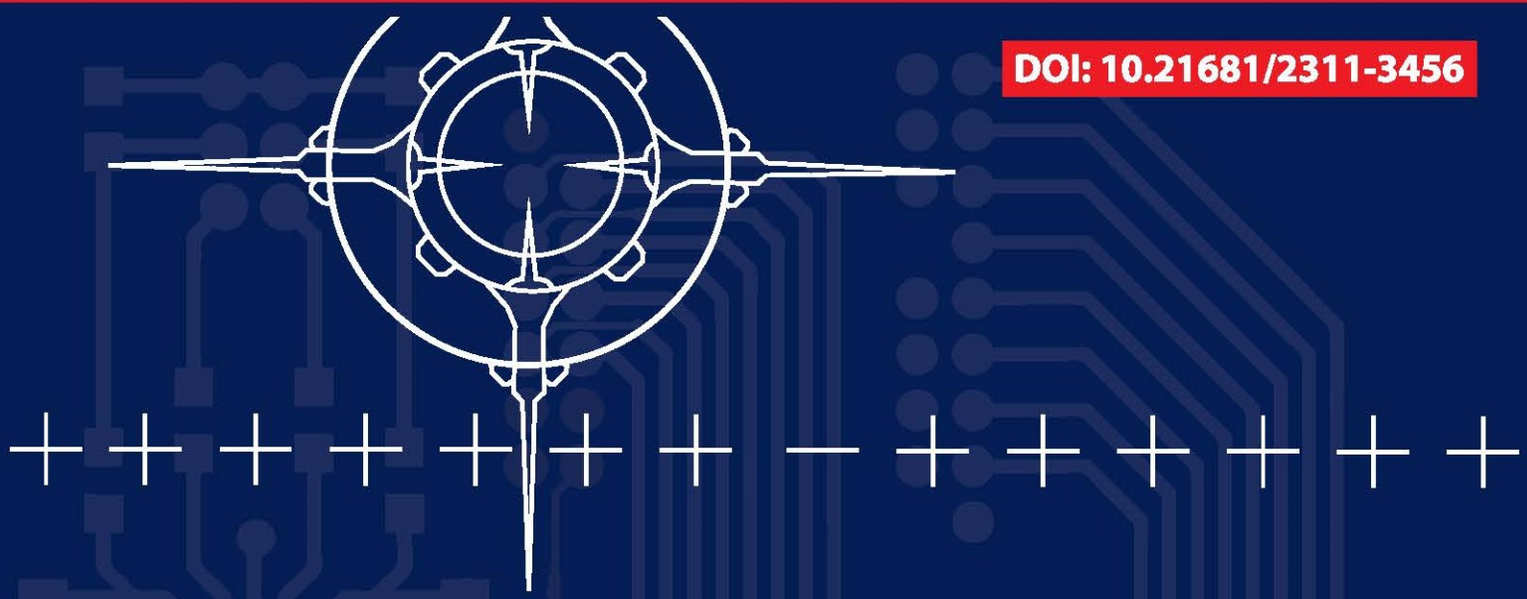
ВОПРОСЫ КИБЕРБЕЗОПАСНОСТИ

№5 2025 (69)

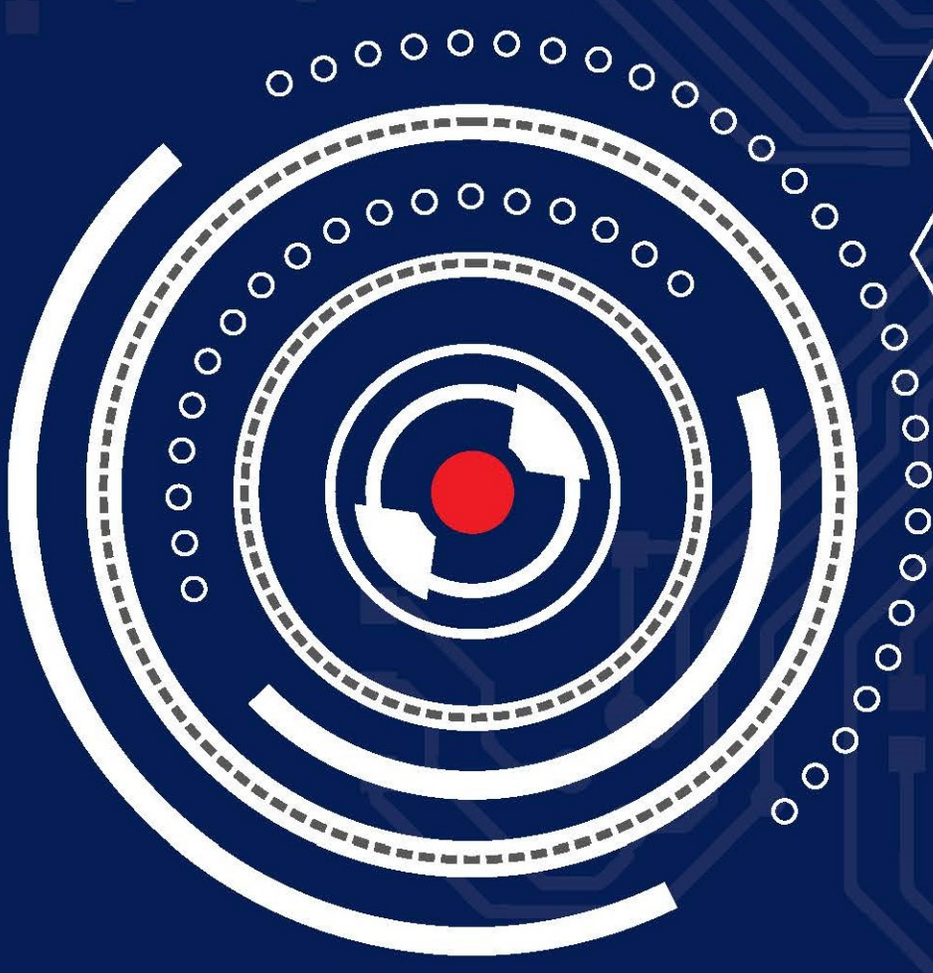
DOI: 10.21681/2311-3456



Совершенствование систем защиты коммерческой тайны

Систематизация средств информационной безопасности

Оценка защищенности систем управления большими данными



BIT INTERNATIONAL SCIENTIFIC AND TECHNICAL CONFERENCE SECURE INFORMATION TECHNOLOGIES

31 октября 2025
МГТУ имени Н.Э.Баумана

XIV Всероссийская научно-техническая конференция с международным участием «Безопасные информационные технологии» (БИТ 2025)

XIV Всероссийская научно-техническая конференция с международным участием «Безопасные информационные технологии» (БИТ 2025). Традиционно проводится при организационной поддержке кафедры «Информационная безопасность» (ИУ-8) МГТУ им. Н.Э.Баумана.

Конференция проходит при поддержке Комитета ТПП РФ по безопасности предпринимательской деятельности.

Направления и секции конференции:

- Криптографические методы защиты информации;
- Методы и средства защиты инфокоммуникационных и биометрических систем;
- Правовые и организационно-технические меры информационной безопасности;
- Методы анализа защищенности информационных ресурсов;
- Общие проблемы информационной безопасности и подготовки специалистов.

https://baumanist.ru/

МИФИ Национальный исследовательский ядерный университет

КИБ-2025 КИБЕРНЕТИКА И ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

3 и 4 декабря 2025
НИЯУ «МИФИ»

Третья Всероссийская научно-техническая конференция «Кибернетика и информационная безопасность» (КИБ-2025)

Третья Всероссийская научно-техническая конференция «Кибернетика и информационная безопасность» (КИБ-2025) посвящена обсуждению актуальных вопросов обеспечения информационной безопасности, выработки эффективных подходов к решению задач по защите информации, обмена практическим опытом построения информационных систем и интеллектуальных систем управления в защищенном исполнении.

Конференция проходит при поддержке Комитета ТПП РФ по безопасности предпринимательской деятельности.

Тематические направления:

- Доверенная электронная компонентная база и ПАК для критической информационной инфраструктуры;
- Защищенные компьютерные системы и технологии;
- Интеллектуальное управление сетевой безопасностью;
- Информационная безопасность социотехнических систем;
- Разработка безопасного программного обеспечения;
- Теоретическая и практическая криптография;
- Информационно-аналитические системы безопасности;
- Проблемы информационной безопасности в системе Высшей школы.

https://kib.mephi.ru/

ВОПРОСЫ КИБЕРБЕЗОПАСНОСТИ

НАУЧНЫЙ РЕЦЕНЗИРУЕМЫЙ ЖУРНАЛ

№ 5 (69) 2025 г.

Выходит 6 раз в год

Журнал выходит с 2013 г. (Свидетельство о регистрации ПИ № ФС77-75239). Перерегистрировано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций 07.03.2019.

Журнал входит в рейтинг научных изданий ВАК в категории К1, индексируется в RSCI, публикует статьи по специальностям 1.2.4 и 2.3.6 – физ-мат. науки; 2.2.15, 2.3.1, 2.3.5, 2.3.6 – техн. науки

Учредитель и издатель

АО «Научно-производственное объединение «Эшелон»

Над номером работали:

Г. И. Макаренко – *шеф-редактор*, И. Ю. Шахалов – *отв. секретарь*, С. С. Игнатов – *верстка*, Ю. С. Логинова – *зам. главного редактора*

Подписано к печати 20.10.2025 г.

Общий тираж 120 экз. Цена свободная

Адрес: 107023, Москва, ул. Электрозаводская, д. 24, стр. 1.
E-mail: editor@cyberrus.info, тел.: +7 (985) 939-75-01.
Требования, предъявляемые к рукописям, размещены на сайте: **https://cyberrus.info/**

СОДЕРЖАНИЕ

Подписка на журнал осуществляется в почтовых отделениях по каталогу «Пресса России». Подписной индекс 40707

СОВЕРШЕНСТВОВАНИЕ СИСТЕМЫ ЗАЩИТЫ КОММЕРЧЕСКОЙ ТАЙНЫ: ПРИНЦИПЫ, КЛАССИФИКАЦИЯ, МЕТОДЫ И ТЕХНОЛОГИИ

Минзов А. С.[1], Невский А. Ю.[2], Минзов С. А.[3]

DOI: 10.21681/2311-3456-2025-5-2-14

Цель исследования: *обоснование системы защиты коммерческой тайны в различных формах ее представления на основе классификации, обоснования принципов, методов и технологий защиты.*

Методы исследования: *ретроспективный анализ требований к защите коммерческой тайны в России и за рубежом, системный анализ при обосновании классификации коммерческой тайны, концептуальное моделирование системы её защиты на основе концепции «нулевого доверия» (Zero Trust), синтез системы защиты коммерческой тайны на всех этапах её жизненного цикла.*

Результаты исследования: *полученные результаты не противоречат существующим нормативным документам по защите коммерческой тайны и могут быть использованы для усиления защитных свойств различных объектов, где возникает необходимость защиты коммерческой тайны в России и за рубежом.*

Научная новизна: *в статье предложены новые подходы к классификации коммерческой тайны с позиций её защиты от разглашения (утечки), принципы защиты коммерческой тайны на основе концепции «нулевого доверия» и система управления защитой коммерческой тайны в виде циклического управляемого защищаемого процесса от создания инновационной идеи, проектирования, внедрения и её эксплуатации.*

Практическая значимость: *предложенные авторами решения и подходы к защите коммерческой тайны позволят повысить уровень защищенности хозяйствующих субъектов, где возникает необходимость её защиты, увеличить инновационную активность в рыночных отношениях и противодействовать промышленному и экономическому шпионажу.*

Ключевые слова: *trade secret, система защиты информации, режим коммерческой тайны, zero trust, нулевое доверие.*

Введение

Понятие «коммерческая тайна» (*trade secret*) впервые появилось в странах с рыночной экономикой в середине 18-го века и законодательно оформлена в современной трактовке в США и европейских странах только в конце 20-го века. Многие исследователи связывают интенсивный экономический рост европейских государств и США с созданием правового института защиты коммерческой тайны [1]. Современное представление о коммерческой тайне во многом основано на принятом в США в 1979 г. законе «О коммерческой тайне» (Uniform Trade Secret Act) [2]. В этом законе коммерческая тайна (КТ) определяется как «***информация*** (включая формулы, модели, программы, механизмы, способы, технологии) или ***технология***, обладающая самостоятельной экономической ценностью (действительной или потенциальной) и недоступна для других лиц, которые могли бы извлечь экономическую выгоду из ее использования или разглашения, и в отношении которой приняты меры по защите ее секретности» [3]. В этом определении есть некоторые противоречия, связанные с классификацией КТ. Например, что включает в себя информационная технология, если в понятие «информация» включены программы и модели? О каких технологиях идет речь в классе «технологии»? Почему данные (маркетинговые, социологические и другие исследования) не могут быть отнесены к КТ в определении КТ?

Эти и другие вопросы не позволяют в судебной практике США четко определить отношения объекта права к КТ. Поэтому было разработано дополнение и определены вопросы, по которым необходимо провести исследование для определения возможного отнесения объекта права к КТ и обоснования судебных решений по КТ, которое включает [4,5]:

1. Экономическую ценность информации для владельца КТ и его конкурентов.
2. Степень известности информации о КТ за пределами бизнеса, использующего КТ.
3. Степень известности этой информации сотрудникам и другим лицам, участвующим в этом бизнесе.
4. Уровень мер, принимаемых владельцем бизнеса по охране конфиденциальности информации о КТ.

1 Минзов Анатолий Степанович, доктор технических наук, профессор, Национальный исследовательский университет МЭИ, г. Москва, Россия. E-mail: MinzovAS@mpei.ru

2 Невский Александр Юрьевич, кандидат технических наук, доцент, Национальный исследовательский университет МЭИ, г. Москва, Россия. E-mail: NevskiyAY@mpei.ru

3 Минзов Степан Анатольевич, заместитель начальника отдела АСУ(БД) АКБ «Фора-Банк», г. Москва, Россия. E-mail: minzov@forabank.ru

5. Количество усилий или средств, затраченных на разработку КТ.
6. Легкость или сложность, с которой информация может быть надлежащим образом получена или воспроизведена другими.

К сожалению, значение критериев для этих вопросов не были конкретно определены, поэтому при решении задач защиты КТ в судах США возникают проблемы отнесения объекта права к КТ. И, тем не менее, приведенное выше определение КТ и классификация вопросов, по которым принимается судебное решение, имеет весьма глубокий смысл, который в современном представлении включает в себя условия, требования и механизмы защиты КТ. Сформулируем эти условия с точки зрения защиты информации.

Первым условием отнесения информации к КТ является ее экономическая ценность. Это означает, что владелец тайны должен уметь *доказать (продемонстрировать) её экономическую эффективность (выгоду)*. Из этого следует, что коммерческая тайна будет защищаться государством до тех пор, пока её владелец сможет доказать её экономическую ценность. Это очень важный момент прекращения действия режима защиты КТ со стороны государства. Отсюда вывод: если КТ не имеет экономической ценности, то нет смысла ее защищать.

Второе условие заключается в том, что КТ должна быть недоступна для лиц за пределами бизнеса, которые могут её использовать. Здесь следует отметить, что любая тайна создается не мгновенно, а путём интеллектуальной целенаправленной деятельности её владельца. Отсюда возникает необходимость защиты КТ на всех этапах её проектирования и внедрения, а не только на этапе её применения. Очень важное следствие из анализа этого условия заключается в том, что КТ основывается на инновационных решениях, которые могут принести экономическую выгоду. Современная методология создания инновационных проектов предусматривает набор процессов от генерации инновационных идей, до разработки инновационного проекта, его внедрения и создания механизмов его защиты. Следовательно, все эти процессы должны быть защищены.

Третье условие заключается в создании таких требований к разработке КТ, ее внедрению и эксплуатации, при которых распространение КТ среди персонала организации является *минимально необходимым и контролируемым*. Следует отметить, что такая форма интерпретации этого условия с позиций защиты КТ в зарубежной печати отсутствует.

Четвертое условие заключается в обеспечении владельцем КТ разумных и достаточных мер её защиты. Критерии «разумности» и достаточности» защиты в законодательства США и Европы четко не определены и, обычно, выясняются судом присяжных в судебном процессе путем оценки разумности мероприятий при организации защиты КТ. Следует отметить, что выполнение этого условия практически не регулируется и носит общий характер, который можно сформулировать в форме следующих рекомендаций [5]:

1) предупреждение сотрудников и третьих сторон о конфиденциальном характере информации посредством соглашений о конфиденциальности, указаний на конфиденциальность в документах;
2) реализация программ профессиональной подготовки для сотрудников или в инструкциях по работе с КТ для сотрудников;
3) защита паролей и межсетевых экранов;
4) физическая блокировка конфиденциальной информации;
5) ограничение доступа к физическим и электронным архивам, где хранятся коммерческие секреты;
6) минимизация количества сотрудников, допущенных к КТ.

При этом, совершенно открытым остается вопрос: а этих мер достаточно для защиты КТ?

Пятое условие связано с первым и используется для оценки значимости КТ. Оно используется для того, чтобы можно было обосновать некоторую модель ответственности[4] за разглашение КТ. Кроме того, этот фактор связан с экономической ценностью КТ для определения максимального размера разумных затрат на систему защиты КТ.

Шестое условие, так же, как и пятое, связано непосредственно с моделью ответственности за разглашение (утечку) информации как со стороны владельца коммерческой тайны, так и со стороны сотрудника, который её разгласил. С учетом увеличивающейся ценности коммерческих секретов и сложности их защиты в США в 1996 г. был принят Акт об экономическом шпионаже, согласно которому кража коммерческих секретов приравнена к федеральному уголовному преступлению, с административным наказанием в виде штрафа до 10 млн долл. США и уголовным сроком до 15 лет [3].

В зарубежных научных обзорах рассматриваются вопросы разработки параллельных проектов КТ в разных организациях. С точки зрения зарубежного законодательства считается вполне допустимым, если результаты достигаются с использованием различных технологий, материалов, методов, условий и других факторов. Такое отношение к параллельным

4 Под термином «модель ответственности» мы понимаем условия, при которых либо применяется законодательство с обоснованием определенных мер ответственности за разглашение (утечку) КТ, либо не применяется законодательство при невыполнении условий защиты или отсутствии доказательства разглашения или утечки КТ.

проектам КТ вполне логично. В мире известно много параллельных научных достижений, выполненных разными учеными в одно время. Например, Александр Попов и Гульельмо Маркони – изобретатели радио, Дмитрий Менделеев и Лотар Мейер создатели периодической системы элементов, позволяющей предсказывать наличие новых элементов в этой системе и другие подобные примеры. Такие коллизии могут быть следствием параллельной разработки известных проблем из открытых источников. Это подтверждает наш тезис о том, что КТ должна защищаться на этапе постановки задачи проекта, относящегося к КТ.

Не менее важным для обсуждения остается вопрос отношений КТ, патента и полезной модели. Патент и коммерческая тайна (ноу-хау) – это два разных способа защиты интеллектуальной собственности, но у них есть и общие черты. Патент предоставляет исключительные права на изобретение на определенный срок (обычно 20 лет), в обмен на публичное раскрытие информации о нем и охраняется государством. Информация о патенте приводится на уровне понимания его сущности, с приведением технического её описания и доказательства новизны. КТ, напротив, предполагает сохранение этой информации в тайне, и ее защита может длиться неограниченно долго, пока информация остается секретной. Есть разница и между понятиями ноу-хау и КТ. КТ – это более общее понятие и главное ее свойство – это секрет, который дает преимущество в рыночных отношениях, а термин ноу-хау относится к тайне производства, которая также имеет экономическую ценность.

Анализ вопросов защиты коммерческой тайны был бы неполным, если бы не были рассмотрены законодательства КТ в других странах. Среди них наиболее интересными в области защиты КТ является законодательства Китая и Японии.

В КНР защита КТ регулируется законом о противодействии недобросовестной конкуренции (Anti-Unfair Competition Law [6]). Этот закон во многом повторяет законодательство США, также определяет коммерческую тайну, устанавливает требования к владельцам КТ и правила ее защиты. Основные требования к владельцам коммерческой тайны в КНР можно сформулировать в следующем виде:

1. Доказательство статуса коммерческой тайны:
   a) Информация должна быть секретной и не общедоступной.
   b) Информация должна иметь экономическую ценность.
   c) Владельцы должны принимать разумные меры для защиты этой информации.
2. Меры по защите КТ:
   a) Владельцы должны внедрять внутренние процедуры и политику для защиты информации, такие как соглашения о конфиденциальности и ограничение доступа.
3. Доказательства утечки КТ:
   a) Владельцы должны иметь возможность продемонстрировать (доказать), как и кем была разглашена КТ.

Очевидно, что существенным различием этих требований с законодательством США в сфере КТ является требование к владельцу КТ по созданию системы контроля использования КТ и *определения источника ее разглашения*. Это требование создает повышенные сложности в создании системы защиты КТ для ее владельцев.

Защита коммерческой тайны Японии регулируется законом о предотвращении утечки коммерческой тайны (Act on the Prevention of Unauthorized Use of a Trade Secret [7]) и практически не отличается от требований сформулированных в законодательстве КНР.

В заключение этого раздела анализа концепции защиты коммерческой тайны в зарубежном законодательстве остановимся на следующих особенностях:

1. Практически во всех законодательных актах зарубежных государств коммерческая тайна рассматривается как очень важная инновационная деятельность (Ноу-хау), направленная на развитие экономического, технологического, производственного и научного суверенитета государства в рыночных отношениях. Несовершенство законодательных актов в сфере защиты КТ приводит к повышению уровня промышленного шпионажа и недобросовестной конкуренции, усложняет проведение расследований преступлений по разглашению КТ и замедляет развитие экономической деятельности хозяйствующих субъектов.
2. Следует отметить, что в настоящее время не существует совершенного законодательства по защите КТ и это связано с различными формами её представления и классификациями, неопределёнными критериями «разумности и достаточности» её защиты, требованиями по предоставлению системы доказательств разглашения (утечки) информации. Это усложняет процедуры защиты КТ в судах.
3. Слабые стороны режимов защиты коммерческой тайны предприятий, которые обеспечивают очень узкий набор требований к защитным мерам КТ и не гарантируют доверия к системе ее защиты. Это являются одной из основных проблем в области управления коммерческой тайной.

4. Весьма поверхностно проработаны вопросы применения систем искусственного интеллекта при проектировании КТ (модели GPT, промпты, Data-Sets и результаты решения задач). Как в этом случае идентифицировать утечку КТ ?

Состояние вопроса по рассматриваемой проблеме в России

В Российской империи в начале 20 века юрист В. В. Розенберг предложил ввести термин «промысловая тайна», однако, этот термин не прижился и вместо него окончательно утвердился термин «коммерческая тайна», объединяющий тайну любой деятельности, имеющей целью извлечение прибыли [8].

После победы Великой Октябрьской социалистической революции уже 27 ноября 1917 г. положением о рабочем контроле, принятым ВЦИК и СНК РСФСР коммерческая тайна была упразднена. В 30-х годах институт коммерческой тайны был заменен государственной и военной тайной. Инновационная деятельность населения в этот период времени поощрялась в форме рационализаторских предложений, трудовых движений, рекордов и других форм инновационной активности населения. Но главное, результаты этой деятельности свободно распространялись в стране. На наш взгляд, именно такая форма инновационной деятельности общества сыграла значительную роль в экономическом развитии СССР в период 20-х – 40-х годов прошлого столетия.

Возрождение во второй половине 80-х гг. предпринимательской деятельности и переход страны к рыночным отношениям повлекли за собой разработку нормативных документов, в том числе касающихся коммерческой тайны. В первую очередь потребовалось сформулировать определение коммерческой тайны. Такое определение было дано в Законе СССР «О предприятиях в СССР» от 4 июня 1990 г. В нем сказано: «Под коммерческой тайной предприятия понимаются не являющиеся государственными секретами сведения, связанные с производством, технологической информацией, управлением, финансами и другой деятельностью предприятия, разглашение (передача, утечка) которых может нанести ущерб его интересам». Руководителю предприятия предоставлялось право определять состав, объем и порядок защиты сведений, составляющих коммерческую тайну.

В отечественном законодательстве учитывается опыт зарубежных правовых механизмов защиты коммерческой тайны. Тем не менее, развитие системы защиты КТ в нашей стране имеет свои особенности, главная из которых заключается в несколько упрощенной форме её защиты путем создания только режима конфиденциальности (ограничения доступа) к коммерческой тайне, что обеспечивается созданием механизма ответственности за ее разглашение и ряда других организационных мер. Такой подход был заимствован из зарубежных законодательных актов и, как нами было рассмотрено ранее, является поверхностным по отношению к защите КТ. Современные условия требуют более совершенных механизмов защиты КТ особенно, если это касается торговых отношений с другими государствами. На международной конференции по комплексной защите информации было высказано мнение о том, что «ущерб от разглашения коммерческой тайны часто выше, чем от разглашения государственной тайны, как бы кощунственно это не звучало[5]».

Сегодня остаются открытыми несколько вопросов, в том числе: достаточно ли этих мер для защиты КТ, необходимо ли усиливать роль государства в защите КТ, обеспечивает ли режим защиту от недобросовестной конкуренции и другие. Все это требует научного анализа построения системы защиты КТ для различных условий и форм ее представления.

Но начнем анализ с определения понятия «коммерческая тайна» в отечественном законодательстве. В настоящее время защита КТ определяется Федеральным законом № 98[6]. В этом документе коммерческая тайна рассматривается как «*сведения любого характера (производственные, технические, экономические, организационные и другие), в том числе о результатах интеллектуальной деятельности в научно-технической сфере, а также сведения о способах осуществления профессиональной деятельности, которые имеют действительную или потенциальную коммерческую ценность в силу неизвестности их третьим лицам, к которым у третьих лиц нет свободного доступа на законном основании и в отношении которых обладателем таких сведений введен режим коммерческой тайны*».

Столь сложная трактовка этого определения, по существу, относит к коммерческой тайне любую деятельность за исключением сведений, рассматриваемых в статье № 5 Федерально закона «О коммерческой тайне», не относящиеся к другим видам тайн и в других ФЗ, определяющих сведения, доступные для всех. На наш взгляд, в этом законе не могут присутствовать слова с неопределённым смыслом, такие как, «*и другие*», «*действительная или потенциальная коммерческая ценность*». Есть и нарушения логики, когда коммерческая тайна определяются сначала как «сведения любого характера» и приводятся примеры этой тайны, а затем следует фраза «*и другие*» (сведения), что требует уточнения формулировки понятия «коммерческая тайна», а также введение четкой классификации видов и форм её представления. Это вполне очевидно, так как невозможно

5 Рособоронэкспорт озаботился защитой информации / 7-я Международная конференция «Комплексная защита информации», 25–27 февраля 2025 г. Минск. URL: https://www.cnews.ru/articles/rosoboroneksport_ozabotilsya_zashchitoj (дата обращения: 01.09.2025).

6 ФЗ №98 «О коммерческой тайне», 2006 г.

построить одинаковую защиту информации, если она представлена в разных формах. Например, защищенный бумажный документооборот отличается от электронного документооборота и способы защиты информации совершенно отличаются друг от друга.

В составе коммерческой тайны некоторые специалисты выделяют две категории сведений: информация являющаяся результатом интеллектуальной деятельности и другие сведения, которые также относятся к коммерческой тайне [9]. Вторая категория коммерческой тайны является весьма субъективной и может представлять собой регистры внутреннего бухгалтерского учёта, досье на конкурентов, списки клиентов, результаты деловой разведки и другую подобную информацию. Определить экономическую ценность этой информации практически невозможно. Точно также невозможно и определить степень ущерба, который может быть нанесён организации, если эта информация получит огласку. Надо ли в этом случае вводить эту информацию в статус коммерческой тайны и привлекать государственные институты для решения проблем с ответственностью при её разглашении или утечке? Этот вопрос для нас остается открытым. Зарубежное законодательство сфокусировано на первой категории КТ.

Защита КТ по ФЗ № 98 осуществляется путем создания и введения правового режима коммерческой тайны в организации, который включает:

- ограничение доступа к информации;
- обозначение носителей информации грифом «Коммерческая тайна»;
- ознакомление работников с правилами обращения с конфиденциальной информацией;
- заключение с работниками соглашений о неразглашении;
- определение ответственных лиц за соблюдение режима коммерческой тайны.

Следует отметить, что этих мер во многих случаях недостаточно и требуется уточнение необходимых мер для различных форм представления КТ. Требования к обладателям КТ в этом ФЗ конкретно не определены в части: разумности и достаточности принятых ими защитных мер, возможности предоставления доказательств об экономической ценности КТ и о каналах утечки данных (разглашении) о КТ.

Более полную ясность в понятие «коммерческая тайна» вносит методический документ[7]. В нём под коммерческой тайной понимается режим конфиденциальности информации (ограничения доступа), позволяющий ее обладателю при существующих или возможных обстоятельствах увеличить доходы, избежать неоправданных расходов, сохранить положение на рынке товаров, работ, услуг или получить иную коммерческую выгоду.

Здесь требуется сделать отступление и вернуться к понятию «коммерция». Этот термин в современном представлении имеет более широкий смысл, чем это было несколько десятков лет тому назад. Коммерция (от лат. commercium – торговля, купля), предпринимательская деятельность экономических агентов – государства, компаний, домохозяйств, нацеленная на получение прибыли (производство товаров, оказание платных услуг, проведение обменных операций, осуществление инвестиций на финансовых рынках и т. д.); в узком смысле – это торговля[8]. Есть ещё одна интересная деталь. На западе и в США коммерческая тайна рассматривается как ***Trade Secret*** [2], что имеет смысл как «секрет рыночных отношений». В такой формулировке вложен более глубокий смысл, чем термин «коммерческая тайна». На наш взгляд, более точное понимание смысла коммерческой тайны заключается в том, что ***это результат интеллектуальной деятельности человека или системы искусственного интеллекта, направленный на совершенствование товаров или услуг в системе рыночных отношений и получения экономической выгоды***. Использование искусственного интеллекта для решения инновационных задач с одной стороны позволяет расширить возможности субъектов рыночных отношений в совершенствовании товаров и услуг, а с другой стороны вызывает необходимость усиленной защиты систем искусственного интеллекта и результатов их деятельности.

Решение ***первой задачи*** может быть основано на методах генеративного искусственного интеллекта [10–13] или путем поиска решений на основе технологий творческого проектирования (Метод А. Половинкина) [14] и теории решения изобретательских задач (ТРИЗ) [15].

Вторая задача связана с новым направлениям кибербезопасности систем искусственного интеллекта и обеспечения защиты от угроз [16–18]. Однозначного решения защиты систем искусственного интеллекта, генерирующего инновационные решения, относящиеся к коммерческой тайне от угроз сегодня не существует.

По мнению авторов статьи [19] искусственный интеллект поднимает множество и других сложных вопросов для законодательства:

- Защита технологий ИИ как коммерческой тайны.
- Защита результатов работы систем ИИ как коммерческой тайны.

7 Разъяснение Президиума ФАС России от 21.02.2018 N 13 «Об информации, составляющей коммерческую тайну, в рамках рассмотрения дела о нарушении антимонопольного законодательства, проведении проверок соблюдения антимонопольного законодательства, осуществлении государственного контроля за экономической концентрацией» (утв. протоколом Президиума ФАС России от 21.02.2018 N 2).

8 Большая Российская энциклопедия, 2022 г.

- Риски для коммерческой тайны, когда модель генеративного ИИ не может дать точного ответа и дает искажённый результат (bias) или галлюцинирует [20].
- Определение разницы между ИИ с закрытым и открытым исходным кодом и оценка их последствий для коммерческой тайны.

На наш взгляд, защите КТ с использованием ИИ также должны подвергаться и сценарии работы с ChatGPT (prompt и pipeline), которые и определяют результаты работы ИИ.

Очень важно, что современное понятие «коммерческая тайна» распространяется на широкий круг субъектов торговых отношений. Среди них сегодня выделяется крупные компании и государственные корпорации такие как РосАтом, Газпром, Роснефть, ОАО «ФСК ЕЭС», Рособоронэкспорт, а также НИИ, ВУЗы, многие производственные и другие организации. Практически во всем мире сложилась ситуация, когда каждое государство не только выполняет свои обязательства по защите коммерческой тайны, но и обеспечивает развитие научно-технического потенциала субъектов рыночных отношений за счёт выполнения ими инновационных проектов, относящихся к коммерческой тайне и имеющих преимущество на рынке товаров и услуг. К сожалению, эта сторона коммерческой тайны, как направление управления развитием научно-технического и технологического потенциала страны, в нашей научной среде сегодня практически не обсуждается, хотя такая потребность существует. Этот тезис подтверждается и в зарубежных исследованиях [1], но даже в тех странах, где коммерческая тайна защищается государством более 200 лет (США и европейские страны) вопрос ставится только об изучении влияния коммерческой тайны на развитие инновационного потенциала. Вопросы управления этим потенциалом не рассматриваются.

Таким образом, создание только режима защиты коммерческой тайны по Российскому законодательства в условиях применение систем ИИ для нахождения инновационных решений в системе рыночных отношений явно недостаточно.

Утверждение этого тезиса мы находим и в других зарубежных аналитических исследованиях по проблемам коммерческой тайны: «слабые стороны режимов защиты коммерческой тайны предприятий, низкий уровень деловой осведомленности, ограничение мобильности сотрудников, кибербезопасность, слабые стороны идентификации и защиты коммерческой тайны являются одними из основных политических проблем в области управления коммерческой тайной сегодня» [4].

Концепция и принципы защиты коммерческой тайны

Существующая концепция защиты коммерческой тайны сегодня определена Федеральным законом № 98 и заключается в создании правового режима коммерческой тайны. Мы уже отмечали ранее, что этот режим не обеспечивает достаточную защиту коммерческой тайны и требует совершенствования. Возникает вопрос, а что в этом случае можно применить? Сегодня в РФ существует ряд нормативно-методических документов в форме постановлений Правительства РФ, приказов ФСТЭК и методических документов[9], определяющих требования по защите конфиденциальной информации, относящиеся к персональным данным, государственным учреждениям, банковской тайне и значимым объектом критической информационной инфраструктуры. Для информации, не относящиеся к конфиденциальной, применяются государственные стандарты серии ГОСТ Р ИСО/МЭК 27000, которые являются эквивалентами международных стандартов ISO/IEC 27000. Концепции защиты информации в этих двух группах нормативно-методических документов, действующих на территории РФ, существенно отличаются. Если группа отечественных нормативно-методических документов меры по защите информации определяет в зависимости от класса или уровня защищенности информационной системы, то группа международных стандартов рекомендует использовать меры в зависимости от уровня рисков безопасности информации. Оценка возможности применения этих концепций защиты коммерческой тайны показывает, что ни одна из них не может быть использована в полной мере. Основная причина заключается в том, что система защиты КТ *работает только до первой реализации угрозы утечки (разглашения) КТ*. После этого нет необходимости в ее защите, так как дальнейшее ее применение уже не даст экономических выгод и, следовательно, не целесообразно. Это требует другой концепции создания архитектуры информационной безопасности, основанной на более высоком уровне защищенности КТ и доверия к ней.

Как в настоящее время создаются системы с заданным, повышенным или измеряемым уровнями доверия к ним?

Доверие к ИТ-проектам в концепции стандарта ГОСТ 15408[10] это «основа для уверенности в том, что продукт ИТ отвечает целям безопасности».

В этом стандарте определен и механизм обеспечения доверия к системе информационной безопасности, как «бездоказательное утверждение,

9 Вся система нормативно-методических документов приведена сайте ФСТЭК https://fstec.ru.

10 ГОСТ Р ИСО/МЭК 15408-3-2013 «Информационная технология. Методы и средства обеспечения безопасности. Критерии оценки безопасности информационных технологий. Часть 3. Компоненты доверия к безопасности».

предшествующий аналогичный или специфический опыт», а также с использованием активного исследования ИТ-продукта для определения его свойств безопасности. Требования доверия представляются в виде структуры: класс–семейство–компонент–элемент. Основные принципы этого стандарта состоят в том, что следует четко сформулировать угрозы безопасности, положения политики безопасности организации и продемонстрировать достаточность предложенных мер безопасности.

Основной способ достижения доверия к системе информационной безопасности основан на проведении его оценки (всего 6 уровней оценки доверия[11]). Методы оценки основаны на анализе процессов, требований к ним, верификации доказательств, независимом функциональном тестировании, анализе уязвимостей и тестировании на проникновение.

Доверие к техническим средствам ИТ-проектов обеспечивается транзитивно путем применения сертифицированных технических средств, удостоверяющих центров и других средств, имеющих сертификаты соответствия. Кроме того, доверие к системе информационной безопасности может быть обеспечено и другими средствами, методами и технологиями:

1. Аттестацией объектов информатизации.
2. Оценкой соответствия требований к системе управления ИБ через её аудит (ГОСТ Р ИСО/МЭК 27002-21 г.).
3. Применением абстрактных формальных моделей доступа, целостности и доступности (Модель Белла–ЛаПадулы, Биба, Clark–Wilson, Take–Grant и др.).
4. Тестированием системы ИБ на этапе проектирования ИТ-продукта[12].
5. Применением механизма доказательств доверия на основе языка событий Event-B и платформы Rodin[13]. Это совместный проект различных команд. Наибольший вклад в его разработку вносят Саутгемптонский университет, компания Systerel и Дюссельдорфский университет.

В настоящее время ни один из них не создает достаточную убедительную систему доказательств доверия к системе ИБ. Последний из рассмотренных средств, методов и технологий (Event-B) уже используется на практике, однако существует ряд проблем его применения[14]:

- Высокий уровень трудозатрат на проведение формальной верификации.
- Ограниченная поддержка командной разработки.
- Отсутствие поддержки выделения часто используемых выражений в отдельные сущности с последующим доступом к ним по ссылке.
- Возможность проявления субъективных оценок и ошибок при написании кода.
- Эта технология не прошла сертификацию ФСТЭК на НДВ и УД.

Но есть и другой подход к созданию системы доверия. С этой точки зрения более интересной является концепция Zero Trust (ZT) – нулевого доверия. Это парадигма кибербезопасности, в которой ни один источник информации и процесс в информационной системе не считаются доверенным без подтверждения [22–24]. Архитектура такой системы информационной безопасности построена на принципе постоянного и полного контроля достоверности источников информации, всех субъектов доступа (пользователи, приложения, устройства) и объектов доступа (корпоративная сеть, интернет, приложения, объекты ввода-вывода информации и другие компоненты информационных систем). Она не исключает использование существующей концепции транзитивного доверия третьей стороны (сертификаты на технические средства защиты информации, SSL, аттестованные объекты информатизации, удостоверяющие центры). Сложность подобно организованных информационных систем в несколько раз превышает сложность обычных систем. Основываясь на этой концепции, можно гарантировать безопасность, в основном, за счёт четырёх аспектов: динамической аутентификации, контроля доступа, непрерывного мониторинга и оценки состояния системы. Эти компоненты тесно сочетаются для реализации системы безопасности.

В настоящее время концепция ZT широко обсуждается в научном сообществе. Практическая реализация этой концепции начата в США в соответствии с указом президента 14028[15] («О повышении кибербезопасности нации»), в котором Федеральному правительству США необходимо начать переход к архитектурам с нулевым доверием в своих инфраструктурах SaaS, PaaS и IaaS. К сожалению, нормативных документов РФ по архитектурам информационных систем с нулевым доверием пока нет, но у нас, как и в мире, эта модель безопасности является активно развивающейся концепцией. В настоящее время широко используется несколько концепций, в том числе NIST [23] и Forrester [24].

11 Приказ ФСТЭК России от 02.06.2020 № 76 «О требованиях по безопасности информации».

12 ГОСТ Р 56939-2024 «Разработка безопасного программного обеспечения».

13 Илья Щепетков, Rodin – платформа для разработки и верификации моделей на Event-B, URL: https://www.ispras.ru/upload/iblock/5e5/5e5ac3663 3ead83d10476199d697be85.pdf (дата обращения: 01.09.2025).

14 Хорошилов А. В., Щепетков И. В. ADV_SPM – Формальные модели политики безопасности на практике. Труды Института системного программирования РАН. 2017;29(3):43–56.

15 US executive order 14028, Improving the Nation's Cyber Security, URL https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity (дата обращения: 01.09.2025).

Они практически отличаются только терминологией и включают следующие основные принципы[16] построения архитектуры ZT [22]:

1. *Безопасность источника данных*. Этот принцип предполагает оценку доверия к источнику информации и его контроль.
2. *Безопасность связи* – связь (канал) должна быть защищена.
3. *Безопасность сеанса* — доступ к ресурсам предоставляется на основе сеанса, а аутентификация и авторизация для одного ресурса не могут предоставлять привилегии другим.
4. *Контроль доступа* — доступ к ресурсам определяется динамической политикой, включая наблюдаемое состояние идентификационных данных клиента, приложения и запрашивающего актива.
5. *Максимальный уровень безопасности* — организация обеспечивает, чтобы все собственные и связанные с ним устройства находились в максимально безопасном состоянии, и отслеживает активы для обеспечения этого.
6. *Непрерывная аутентификация* — все процессы авторизации и аутентификации ресурсов динамические и строго соблюдаются. Организация, планирующая внедрение ZT, может использовать систему управления идентификацией, учётными данными и доступом, а также применять многофакторную аутентификацию для повышения безопасности.
7. *Регистрация информации* — организация собирает как можно больше информации о текущем состоянии сетевой инфраструктуры и коммуникаций и использует эту информацию для повышения уровня своей безопасности.

Анализ возможности применения этих принципов для создания архитектуры ZT коммерческой тайны с учётом требований к её определению и защите, которые были рассмотрены нами ранее в законодательстве США, а также разнообразие форм представления КТ показал, что этих принципов недостаточно и необходимо ввести дополнительно следующие:

1. *Периодическая оценка ценности КТ* на всех этапах её создания, внедрения и применения. Это обеспечивает контроль статуса информации как КТ.
2. *Выделение процессов, в которых используется КТ и персонал*, принимающий в этом участие. Этот принцип позволяет разделить технологические процессы на отдельные выделенные зоны и, тем самым, обеспечить их изолированность и снижение количества сотрудников, включённых в режим КТ.
3. *Обеспечение доказательства утечки КТ*. Реализация только принципа №7 не позволяет полностью реализовать этот принцип и потребует принятия дополнительных технических решений для сбора доказательной базы утечки КТ.
4. *Транзитивность доверия в отношении применяемых сертифицированных средств и технологий*.

16 Знание некоторых принципов освобождает от необходимости знания многих фактов. Клод Адриан Гельвеций.

Классификация форм представления коммерческой тайны и методы их защиты

В настоящее время не существует чёткой классификации форм представления коммерческой тайны. Это является одной из причин отсутствия чётких методических рекомендаций по созданию системы защиты КТ и объясняет появление в законодательстве отдельных стран требований по обеспечению правового режима КТ. Все остальные меры рекомендуется принимать на основе понятий разумности и достаточности. Однако без чёткой классификации построить систему защиты КТ практически невозможно, так как разная форма представления КТ определяет разные механизмы её защиты. Однако выход из этого положения существует, если применить другую парадигму классификации КТ, в которой положена общность механизмов её защиты. Основанная на таком принципе классификация представлена на рис. 1.

В основе любой КТ лежит инновационная идея и это является главным элементом её защиты. Отсюда возникает необходимость защиты инновационных идей на этапе их разработки иначе это может привести к параллельному проектированию КТ, что и происходит во многих случаях. По этой классификации инновационные идеи могут существовать в двух формах: информации и технологий. Носители информации также классифицируются исходя из общих подходов к их защите.

Прежде всего это персонал, обладающий доступом к КТ. Персонал является основным источником утечки информации и разглашения КТ. Технологии организации работы с персоналом по защите секретов в настоящее время достаточно хорошо развиты в отечественных и зарубежных нормативных документах по информационной безопасности [24–25]. Эти технологии основаны на создании: режима конфиденциальности КТ, профессиональной подготовки и обучении, осведомлении, управлении развитием профессиональной этики и поведения персонала при работе с КТ, созданием системы контроля и ответственности за её разглашение.

Класс «информация» разделяется ещё на три класса: материальные носители информации (документы и изделия, содержащие КТ) и электронные носители информации съемные и несъемные. Если материальные носители информации потребуют

Рис. 1. Классификация форм представления коммерческой тайны с позиций общих механизмов её защиты

обычных методов защиты связанные с организацией хранения секретных документов, их учётом и контролем, то электронные носители могут иметь дополнительные технические устройства, обеспечивающие доступ к информации и использование криптографических методов защиты КТ.

Вторая группа классификации «технологии» наиболее сложная и представляет собой несколько классов. Среди них по критерию общих механизмов защиты можно выделить следующие:

1. Информационные технологии, модели и алгоритмы, содержащие КТ. Они должны быть защищены как в процессах разработки, так и в процессах эксплуатации, а также при выводе их из условий эксплуатации. Это обеспечивается применением принципов и механизмов защиты в концепции ZT. В том случае, если эта технология передается по соглашению о нераспространении в другую организацию, необходимо предусмотреть разработку механизмов контроля за её распространением и определения источника несанкционированного распространения КТ. Это может достигаться использованием водяных знаков и скрытой маркировки технологии с использованием стеганографии.
2. Производственные технологии, содержащие КТ. Технологическая линия, в которой используется КТ разделяется на зоны, в которых используется КТ с разной степенью возможного её раскрытия. Проводится анализ каналов утечки информации и проектируются защитные мероприятия по предотвращению утечки производственной информации, составляющей КТ. Одновременно решаются вопросы по созданию режима КТ в отношении персонала, допущенного к КТ. В том случае, если используются информационные технологии, то выполняются меры по защите КТ, рассмотренные в п. 1 этой классификации.
3. При работе с рецептурами и штаммами, отнесенными к КТ необходимо разделить технологию производства на зоны их приготовления и использования. В зонах приготовления рецептур необходимо обеспечить защиту КТ после анализа каналов утечки информации. Кроме того, необходимо обеспечить контроль за распространением и учетом рецептур и штаммов. Принять меры защиты по отношению к персоналу (см. п. 1).

4. Отдельная категория КТ – это совокупность методов и способов обработки данных и специализированных технологий для принятия решений, оптимизации процессов и создания новых продуктов и услуг (BigData). Защита этой категории КТ обеспечивается применением принципов и методов защиты информации в концепции ZT.
5. Категория «искусственный интеллект», как коммерческая тайна, представляет собой технологию виде платформ, обычно на основе нейронных сетей и интерфейсов взаимодействия с ними в форме больших лингвистических моделей. Эти платформы используют для решения задач специальным образом организованные базы данных (DataSet) и Интернет. Вопросы обеспечения безопасности по критериям конфиденциальности, целостности и доступности представляют собой для них серьезную проблему. Если конфиденциальность и доступность этих технологий можно обеспечить, то вопросы их целостности остаются открытыми из-за динамичности изменения состояния баз данных. Это связано и с достоверностью результатов, генерируемых системами ИИ.
6. Организационные технологии также могут быть отнесены к коммерческой тайне. Они представляются обычно в форме концепций, положений, политик, процедур, инструкций, планов, и других документов, применение которых может привести к значительным экономическим эффектам. Электронные формы этих документов защищаются обычными методами с применением криптографических средств и разделением доступа персонала к ним. В качестве примера эффективного применения организационных мер можно привести решение Г. Форда по применению заводского конвейера для сборки автомобилей, что позволило сократить время на производство одного автомобиля с 12 до 1,5 часов[17].

Этапы создания, внедрения и применения коммерческой тайны в организации

Анализ зарубежных и отечественных концепций защиты коммерческой тайны привёл нас к интересному выводу: практически во всех правовых и методических документах КТ рассматривается как некоторое уже готовое решение. Но реально КТ создается не мгновенно, а путем определенной интеллектуальной деятельности, где в основе её всегда лежит некоторая инновационная идея, которая на последующих этапах разработки КТ и её внедрения преобразуются уже в готовое решение. Если эта идея становится общеизвестной, то в этом случае разработка КТ может проходить параллельно в других организациях, что может снизить ожидаемый экономический эффект, а ее правовой статус «коммерческая тайна» будет утерян. Следовательно, защиту КТ необходимо создавать на всех этапах разработки, внедрения и её

17 Форд Г. Генри Форд. Моя жизнь. Мои достижения. – Litres, 2017.

Рис. 2. Полный цикл управления системой защиты коммерческой тайны от её создания до вывода из эксплуатации в концепции ZT

применения. В отдельных случаях необходимо создавать систему защиты информации и при выводе её из эксплуатации. Полный цикл всех этапов разработки КТ от создания инновационных идей до вывода КТ из эксплуатации представлен на рис. 2. Практически не все этапы могут быть пройдены. Это определяется сложностью КТ, научным уровнем ее разработки, трудоемкостью решений и масштабами распространения. В любом случае два этапа «Внедрение КТ» и «Эксплуатация КТ» будут обязательны.

В каждый этап включено моделирование сценариев угроз разглашения (утечки) КТ. Сформулируем условия, при которых выполняются требования концепции ZT.

Пусть p –это процесс связанный с КТ и $p \in P$, t – угроза разглашения КТ и $t \in T$, $d(t)$ – функция доверия к созданию системы защиты от этой угрозы путем принятия мер $m(t)$, которая изменяется в пределах от «0» (отсутствие мер безопасности и полное недоверие) к «1» (полное доверие за счет принятия мер безопасности). Тогда условие реализации ZT для архитектуры ZT будет в следующем виде:

$$\forall p \in P(\neg\exists T(T \in ProcessSets(p) \wedge \\ \wedge (\forall t \in T,\ d(t) = 0),\ m(t) \notin \varnothing))). \quad (1)$$

Если $z(kt)$ – коммерческая ценность продукции, полученной с использованием КТ, а z – коммерческая ценность продукции без применения КТ, тогда условие отнесения ее к коммерческой тайне будет

$$z(kt) >> z\,. \quad (2)$$

При этом должны быть выполнены 11 рассмотренных нами ранее принципов разработки архитектуры безопасности в концепции ZT. На каждом этапе управления КТ моделируются сценарии разглашения (утечки) КТ. В основе моделирования положен анализ процессов работы с КТ и условий их реализации, при которых возможно её разглашение или утечка. Например, при разработке инновационной идеи необходимо определить возможность и необходимость ее обсуждения, формы и технологии её документирования, хранения и распространения. Это позволяет определить каналы возможного распространения КТ и принять необходимые организационные и технические решения, рекомендованные ФСТЭК. Аналогичные подходы к моделированию сценариев разглашения (утечки) КТ применяются и на других этапах управления системой защиты КТ, вплоть до вывода КТ из эксплуатации. Это необходимо в том случае, когда инновационные идеи КТ могут найти новые приложения для практического применения, либо информация является актуальной для дальнейшего использования.

Заключение

Рассмотренные результаты анализа систем защиты информации о коммерческой тайне за рубежом и в России позволяют сделать следующие выводы:

1. Совершенствование защиты коммерческой тайны в России является сегодня одним из важнейших направлений развития научно-технического, технологического, производственного, экономического потенциала страны и укрепление её суверенитета. Эта роль коммерческой тайны в нашей научной среде практически сегодня не обсуждается, хотя такая потребность есть.
2. В настоящее время в мире не существует совершенного законодательства по защите КТ и это связано с разнообразными формами её представления и классификациями, неопределёнными критериями «разумности и достаточности» к её защите. Это является одной из основных проблем в системе управления коммерческой тайной.
3. Современное понятие «коммерческая тайна» распространяется на широкий круг субъектов международных торговых отношений, где в поставляемых нашей страной товарах и изделиях также могут быть тайны, требующие защиты. В нашей научной среде этот вопрос практически не обсуждается, хотя такая потребность также существует.
4. Весьма поверхностно проработаны вопросы применения систем искусственного интеллекта при проектировании КТ [24]. Также требуется совершенствование модели ответственности за разглашение (утечку) КТ. В основу этой модели должен быть положен ущерб, который может понести владелец КТ за её разглашение или утечку.
5. Авторами предложены новые подходы, часть из которых (5a,5b) используется в зарубежных правовых актах, включающие:
   a. Обоснование коммерческой ценности КТ её владельцем, как необходимое условие для защиты КТ государством.
   b. Обязанность владельца КТ по обоснованию разумных и достаточных мер по её технической и организационной защите и доказательства разглашения или утечки КТ.
   c. Классификация КТ с точки зрения общих подходов к её защите.
   d. Обоснование и принципы защиты КТ на основе концепции «нулевого доверия».
   e. Представление КТ в форме циклического управляемого защищаемого процесса от создания инновационной идеи до проектирования, внедрения и эксплуатации КТ. В отдельных случаях обеспечивается защита КТ при выводе её из эксплуатации.

Литература

1. Nashkova S. Defining Trade Secrets in the United States: Past and Present Challenges–A Way Forward? // IIC-International Review of Intellectual Property and Competition Law. – 2023. – Т. 54. – №. 5. – С. 634–672.
2. Desaunettes-Barbero L. Trade Secrets Legal Protection // Munich Studies on. – 2023.
3. Kapczynski A. The public history of trade secrets //UC Davis L. Rev. – 2021. – Т. 55. – С. 1367.
4. O. Ozcan, D. Pickernell and P. Trott, A Trade Secrets Framework and Strategic Approaches, in IEEE Transactions on Engineering Management, vol. 71, pp. 10200–10216, 2024. DOI: 10.1109/TEM.2023.3285292.
5. Kim Y. et al. The effect of trade secrets law on stock price synchronicity: Evidence from the inevitable disclosure doctrine // The Accounting Review. – 2021. – Т. 96. – №. 1. – С. 325–348.
6. Anti-Unfair Competition Law. URL: http://en.npc.gov.cn.cdurl.cn/laws.html (дата обращения: 01.09.2025).
7. Act on Investment Trusts and Investment Corporations https://www.japaneselawtranslation.go.jp/en/laws/view/3605 (дата обращения: 01.09.2025).
8. Федорова Д. А., Котельникова М. А., Старченко А. С. Развитие законодательства Российской Федерации о коммерческой тайне. Порядок возникновения и прекращения права на коммерческую тайну // Международный журнал гуманитарных и естественных наук. – 2023. – № 5 3 (80). – С. 127–131.
9. Балычев А. П. Коммерческая тайна как вид конфиденциальной информации: правовое регулирование в Российской Федерации // Вестник науки. – 2024. – Т. 2. – №. 4(73). – С. 208–218.
10. Федоров П. Г. Формы проявления коммерческой тайны в цифровой экономике // Актуальные проблемы российского права. – 2025. – №. 1(170). – С. 86–97.
11. D. S. Generative artificial intelligence and trade secrecy // J. Free Speech L. – 2023. – Т. 3. – С. 559.
12. Слицкая А. Е. Использование генеративного искусственного интеллекта в SEO для электронной коммерции // Инновации и инвестиции. – 2023. – №. 11. – С. 326–329.
13. Столяров А. Д., Абрамов В. И., Абрамов А. В. Генеративный искусственный интеллект для инноваций бизнес-моделей: возможности и ограничения // Beneficium. – 2024. – № 3 (52). – С. 43–51.
14. Половинкин А. И. Основы инженерного творчества / А. И. Половинкин; Издательство: Лань. Серия. Техника. ТехниЛань; науки в целом. 2022. – 360 с.
15. Рубин М. С. Основы ТРИЗ для предприятий. М.: КТК «Галактика». 2022. – 354 с.
16. Rajendran, S., & Shankar, K. Artificial Intelligence Techniques for Cybersecurity. Security and Privacy, 2021. 4(1), e122.
17. Брабанд Й., Шебе Х. Оценка безопасности искусственного интеллекта // Надежность. – 2020. – Т. 20. – №. 4. – С. 25–34.
18. Артамонов В. А., Артамонова Е. В., Сафонов А. Е. Безопасность искусственного интеллекта // Защита информации. Инсайд. – 2022. – №. 6(108). – С. 8.
19. Hrdy, Camilla Alexandra, Trade Secrets and Artificial Intelligence (July 14, 2025). Rutgers Law School Research Paper, Trade Secrets and Artificial Intelligence Forthcoming in Elgar Concise Encyclopedia of Artificial Intelligence and the Law (Edward Elgar, eds. Ryan Abbott, Elizabeth Rothman, forthcoming, 2026), Available at SSRN: https://ssrn.com/abstract=5350892 or http://dx.doi.org/10.2139/ssrn.5350892 (дата обращения: 01.09.2025).
20. Ротман Дэнис. RAG и генеративный ИИ. Создаем собственные RAG-пайплайны с помощью LlamaIndex, Deep Lake и Pinecon. – Астана: «Спринт Бук», 2025. – 320 с.: ил. ISBN 978-601-12-3149-7.
21. Theory and Application of Zero Trust Security: A Brief Survey by Hongzhaoning Kang 1ORCID, Gang Liu 1, Quan Wang, Lei Meng and Jing Liu – November 2023 https://www.mdpi.com/1099-4300/25/12/1595 (дата обращения: 01.09.2025).
22. Seefeldt J. what's new in nist zero trust architecture // NIST Special Publication. – 2021. – Т. 800. – С. 207.
23. Gangina P. Demystifying Zero-Trust Architecture for Cloud Applications // Journal of Computer Science and Technology Studies. – 2025. – Т. 7. – №. 9. – С. 542–548.
24. Oforleta, Chibuzor, Reassessing Trade Secret Protections in the Era of AI: A Comparative Perspective on Legal and Ethical Challenges (February 18, 2025). Available at SSRN: https://ssrn.com/abstract=5143701 or http://dx.doi.org/10.2139/ssrn.5143701 (дата обращения: 01.09.2025).

IMPROVING THE TRADE SECRET PROTECTION SYSTEM: PRINCIPLES, CLASSIFICATION, METHODS, AND TECHNOLOGIES

Minzov A. S.[18], Nevsky A. Yu.[19], Minzov S. A.[20]

Keywords: *trade secret, information security system, trade secret regime, zero trust.*

Study objective: *to substantiate a system for protecting trade secrets in various forms based on classification, principles, methods, and technologies.*

Research methods: *a retrospective analysis of trade secret protection requirements in Russia and abroad; a systems analysis to substantiate trade secret classification; conceptual modeling of a trade secret protection system based on the Zero Trust concept; and a synthesis of a trade secret protection system at all stages of its life cycle.*

Study results: *the obtained results do not contradict existing regulatory documents on trade secret protection and can be used to enhance the protective properties of various objects where the need to protect trade secrets arises in Russia and abroad.*

18 Anatoly S. Minzov, Dr.Sc. of Technical Sciences, Professor, National Research University MPEI, Moscow, Russia. E-mail: MinzovAS@mpei.ru
19 Alexander Yu. Nevsky, Ph.D. of Technical Sciences, Associate Professor, National Research University MPEI, Moscow, Russia. E-mail: NevskiyAY@mpei.ru
20 Stepan A. Minzov, Deputy Head of the ACS (DB) Department, JSCB «ForaBank», Moscow, Russia. E-mail: minzov@forabank.ru

Scientific novelty: *the article proposes new approaches to classifying trade secrets from the perspective of protecting them from disclosure (leakage), principles for protecting trade secrets based on the "zero trust" concept, and a trade secret protection management system as a cyclical, controlled, and protected process from the creation of an innovative idea, through design, implementation, and operation.*

Practical relevance: *the authors' proposed solutions and approaches to protecting trade secrets will improve the level of security for economic entities where protection is necessary, increase innovative activity in market relations, and counter industrial and economic espiona.*

References

1. Nashkova S. Defining Trade Secrets in the United States: Past and Present Challenges–A Way Forward? // IIC-International Review of Intellectual Property and Competition Law. – 2023. – T. 54. – №. 5. – S. 634–672.
2. Desaunettes-Barbero L. Trade Secrets Legal Protection // Munich Studies on. – 2023.
3. Kapczynski A. The public history of trade secrets // UC Davis L. Rev. – 2021. – T. 55. – S. 1367.
4. O. Ozcan, D. Pickernell and P. Trott, A Trade Secrets Framework and Strategic Approaches, in IEEE Transactions on Engineering Management, vol. 71, pp. 10200–10216, 2024. DOI: 10.1109/TEM.2023.3285292.
5. Kim Y. et al. The effect of trade secrets law on stock price synchronicity: Evidence from the inevitable disclosure doctrine // The Accounting Review. – 2021. – T. 96. – №. 1. – S. 325–348.
6. Anti-Unfair Competition Law. URL: http://en.npc.gov.cn.cdurl.cn/laws.html (data obrashhenija: 01.09.2025).
7. Act on Investment Trusts and Investment Corporations https://www.japaneselawtranslation.go.jp/en/laws/view/3605 (data obrashhenija: 01.09.2025).
8. Fedorova D. A., Kotel'nikova M. A., Starchenko A. S. Razvitie zakonodatel'stva Rossijskoj Federacii o kommercheskoj tajne. Porjadok vozniknovenija i prekrashhenija prava na kommercheskuju tajnu // Mezhdunarodnyj zhurnal gumanitarnyh i estestvennyh nauk. – 2023. – № 53(80). – S. 127–131.
9. Balychev A. P. Kommercheskaja tajna kak vid konfidencial'noj informacii: pravovoe regulirovanie v Rossijskoj Federacii // Vestnik nauki. – 2024. – T. 2. – №. 4(73). – S. 208–218.
10. Fedorov P. G. Formy projavlenija kommercheskoj tajny v cifrovoj jekonomike // Aktual'nye problemy rossijskogo prava. – 2025. – №. 1 (170). – S. 86–97.
11. D. S. Generative artificial intelligence and trade secrecy // J. Free Speech L. – 2023. – T. 3. – S. 559.
12. Slickaja A. E. Ispol'zovanie generativnogo iskusstvennogo intellekta v SEO dlja jelektronnoj kommercii // Innovacii i investicii. – 2023. – №. 11. – S. 326–329.
13. Stoljarov A. D., Abramov V. I., Abramov A. V. Generativnyj iskusstvennyj intellekt dlja innovacij biznes-modelej: vozmozhnosti i ogranichenija // Beneficium. – 2024. – № 3 (52). – S. 43–51.
14. Polovinkin A. I. Osnovy inzhenernogo tvorchestva / A. I. Polovinkin; Izdatel'stvo: Lan'. Serija. Tehnika. TehniLan'; nauki v celom. 2022. – 360 s.
15. Rubin M. S. Osnovy TRIZ dlja predprijatij. M.: KTK «Galaktika». 2022. – 354 s.
16. Rajendran, S., & Shankar, K. Artificial Intelligence Techniques for Cybersecurity. Security and Privacy, 2021. 4(1), e122.
17. Braband J., Shebe H. Ocenka bezopasnosti iskusstvennogo intellekta // Nadezhnost'. – 2020. – T. 20. – №. 4. – S. 25–34.
18. Artamonov V. A., Artamonova E. V., Safonov A. E. Bezopasnost' iskusstvennogo intellekta // Zashhita informacii. Insajd. – 2022. – №. 6(108). – S. 8.
19. Hrdy, Camilla Alexandra, Trade Secrets and Artificial Intelligence (July 14, 2025). Rutgers Law School Research Paper, Trade Secrets and Artificial Intelligence Forthcoming in Elgar Concise Encyclopedia of Artificial Intelligence and the Law (Edward Elgar, eds. Ryan Abbott, Elizabeth Rothman, forthcoming, 2026), Available at SSRN: https://ssrn.com/abstract=5350892 or http://dx.doi.org/10.2139/ssrn.5350892 (data obrashhenija: 01.09.2025).
20. Rotman Djenis. RAG i generativnyĭ II. Sozdaem sobstvennye RAG-paĭplaĭny s pomoshh'ju LlamaIndex, Deep Lake i Pinecon. Astana: Sprint Buk. 2025. 320 s.: il. ISBN 978-601-12-3149 7.
21. Theory and Application of Zero Trust Security: A Brief Survey by Hongzhaoning Kang 1ORCID, Gang Liu 1, Quan Wang, Lei Meng and Jing Liu – November 2023 https://www.mdpi.com/1099-4300/25/12/1595 (data obrashhenija: 01.09.2025).
22. Seefeldt J. what's new in nist zero trust architecture // NIST Special Publication. – 2021. – T. 800. – S. 207.
23. Gangina P. Demystifying Zero-Trust Architecture for Cloud Applications // Journal of Computer Science and Technology Studies. – 2025. – T. 7. – №. 9. – S. 542–548.
24. Oforleta, Chibuzor, Reassessing Trade Secret Protections in the Era of AI: A Comparative Perspective on Legal and Ethical Challenges (February 18, 2025). Available at SSRN: https://ssrn.com/abstract=5143701 or http://dx.doi.org/10.2139/ssrn.5143701 (data obrashhenija: 01.09.2025).

ИНТЕРОПЕРАБЕЛЬНОСТЬ КАК ОСНОВА ДЛЯ СИСТЕМАТИЗАЦИИ МЕТОДОВ И СРЕДСТВ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

Гришенцев А. Ю.[1], Коровкин Н. В.[2], Коробейников А. Г.[3]

DOI: 10.21681/2311-3456-2025-5-15-27

Цель исследования: *развитие теоретических основ информационной безопасности за счёт обоснованной систематизации, методов и средств информационной безопасности на основе понятия интероперабельность.*

Методы исследования: *анализ информационного взаимодействия и угроз при информационном взаимодействии на основе стандартизированной эталонной модели интероперабельности и синтез систематической структурированной модели методов и средств информационной безопасности в контексте понятия интероперабельность.*

Результаты исследования: *на основе анализа научных направлений интероперабельность и информационная безопасность, предлагается дополнить область интересов информационной безопасности семантическим уровнем, в соответствии с эталонной моделью интероперабельности. Выполнен анализ угроз информационной безопасности объекту защиты реализуемых на семантическом уровне информационного взаимодействия. В ходе исследований доказана необходимость информационной безопасности на семантическом уровне для обеспечения полноты защиты информационного взаимодействия и удовлетворения интересов объекта информационной защиты. Предложена информационная модель разработки и реализации методов информационной безопасности способствующая достижению целей объекта защиты. Предложена модель информационной безопасности на основе эталонной модели интероперабельности. Предложена модель аудита и оценки рисков информационной безопасности на основе эталонной модели интероперабельности.*

Научная новизна: *заключается в новом подходе к систематизации методов, средств и увеличении сферы интересов информационной безопасности на основе современных научных представлений о уровнях информационного взаимодействия в соответствии с понятием интероперабельность.*

Ключевые слова: *защита информации, информационное взаимодействие, открытые системы, модели, стандарты.*

Введение

Отечественный ГОСТ по основным терминам и определениям интероперабельности с 2012 года обновился дважды, первое издание 2012 года ГОСТ Р 55062–2012[4], второе, обновлённое, дополненное и переработанное, 2021 года ГОСТ Р 55062–2021[5].

В соответствии с ГОСТ Р 55062–2021, интероперабельность – способность двух или более информационных систем (ИС) или компонентов к обмену информацией и к использованию информации, полученной в результате обмена.

Значительным фактором эволюции научного знания являются благоприятные условия внешней среды, стимулирующие развитие. Для информационной безопасности благоприятными факторами внешней среды, являются: интерес различных лиц и организаций к её проблемам и высокий спрос на специалистов. Но, пожалуй, основным фактором является постановка приоритетов развития и обозначение имеющихся преград и препятствий на государственном уровне. Так на заседании дискуссионного клуба «Валдай» [1] которое состоялось 7 ноября 2024 в Сочи, в речи президента РФ В. В. Путина впервые на высшем уровне озвучен факт не суверенного положения РФ, обозначен приоритет борьбы за суверенитет и отмечена ключевая роль и комплексность понятия безопасности.

По мнению авторов, актуальность исследований поддержана создавшимися политическими и экономическими условиями, сложившимися в настоящее время внутри Российской Федерации и за её рубежами.

1 Гришенцев Алексей Юрьевич, доктор технических наук, доцент, член-корреспондент Академии электротехнических наук Российской Федерации, доцент Федерального государственного автономного образовательного учреждения высшего образования Национальный исследовательский университет ИТМО. Санкт Петербург, Россия. E-mail: AGrishentsev@yandex.ru

2 Коровкин Николай Владимирович, доктор технических наук, профессор, действительный член Академии Электротехнических Наук Российской Федерации, профессор Федерального государственного автономного образовательного учреждения высшего образования Санкт Петербургский политехнический университет Петра Великого. Санкт-Петербург, Россия. E-mail: Nikolay.Korovkin@gmail.com

3 Коробейников Анатолий Григорьевич, доктор технических наук, профессор, заместитель директора по науке Санкт-Петербургского филиала Федерального государственного бюджетного учреждения науки Института земного магнетизма, ионосферы и распространения радиоволн им. Н. В. Пушкова Российской академии наук, профессор Федерального государственного автономного образовательного учреждения высшего образования Национальный исследовательский университет ИТМО. Санкт-Петербург, Россия. E-mail: Korobeynikov_A_G@mail.ru

4 ГОСТ Р 55062–2012 Информационные технологии. Системы промышленной автоматизации и их интеграция. Интероперабельность. Основные положения. М.: Стандартинформ, 2012. – 11 с.

5 ГОСТ Р 55062–2021 Информационные технологии. Интероперабельность. Основные положения. М.: Стандартинформ, 2021. – 11 с.

В частности вектором развития, о котором идёт немало дискуссий, в том числе в научных работах и верхних эшелонах власти, направленным на технологический и производственный суверенитет [2] в условиях санкционного давления и мирового системного кризиса, охватывающего все сферы интересов человеческой цивилизации от политического и экономического до культурного, демографического и миграционного [3]. Таким образом, актуальным является выявление угроз и обеспечение информационной безопасности на всех уровнях информационного взаимодействия информационных систем и/или их компонентов.

Состояние исследований по интероперабельности и предпосылки к постановке задачи

Впервые необходимость систематизации информационного взаимодействия в рамках понятия интероперабельность сформировалась в недрах военных ведомств[6] США и промышленных гигантов[7] в области информационных технологий. В частности, одно из первых определений интероперабельности дано министерством обороны США[8] (англ. Department of Defense, US DOD). В этом же документе, со ссылкой на (Electronic Warfare. Joint Pub. 3-13.1) даётся определение понятия информации. Информация – 1) факты, данные или инструкции в любом виде; 2) значение, которое человек придает данным с помощью известных соглашений, используемых при их представлении. Приведём точные формулировки на английском. Information – 1) Facts, data, or instructions in any medium or form. 2) The meaning that a human assigns to data by means of the known conventions used in their representation (JP 3-13.1). В части 2 определения информации подчёркивается значимость семантического уровня информационных взаимодействий, т.е. известных соглашений (о смыслах), и разделение собственно данных и информации на две различных понятийных категории. При этом в первом определении информации факты, данные или инструкции, являются её синонимами. По мнению авторов, второе определение понятия информации является более полным, т.к. ключевым аспектом информационной интерпретации тех или иных данных, фактов, инструкций является их смысловое наполнение, что достигается за счёт соглашений о смыслах.

Имеется немало отечественных публикаций о функции информационного взаимодействия в процессе управления и эволюции информационных систем [4–6], а также о вопросах безопасности информационного взаимодействия и противоборства [7–9]. При этом отсутствуют работы в явном виде, связывающие модель информационного взаимодействия на основе понятия интероперабельность и информационную безопасность.

Из определения интероперабельности ясно, что интероперабельность занимается систематизацией и исследованием способности к информационным отношениям между системами или их элементами техническими [10, 11], а также и/или биологическими, например, в виде человеко-машинных интерфейсов в промышленности и науке [12, 13] в медицине [14]. В свою очередь, информационная безопасность, как сказано в ряде различных трудов [15] и нормативных документов, например: ГОСТ Р 53114-2008, ГОСТ Р 50922-2006, занимается защитой интересов объектов (и субъектов) информационного взаимодействия, которые так же являются информационными системами. Следовательно, уровни информационного взаимодействия интероперабельности, по мнению авторов, могут и должны являться основой для систематизации методов и средств информационной безопасности и защиты информации. Это обоснованно тем, что информационная безопасность защищает интересы некоторой стороны информационного взаимодействия, а само информационное взаимодействие, как доказывает теория и практика развития интероперабельности происходит на уровнях, называемых уровни интероперабельности [16].

В рамках данной работы предлагается использовать интероперабельность как основу для систематизации методов и средств информационной безопасности.

Ещё одной предпосылкой к данному исследованию является то, что некоторые специалисты по информационной безопасности позиционируют информационную безопасность как методы защиты информации безразличные к её сущности и содержанию. На некоторых уровнях информационной безопасности такой подход приемлем и даже необходим, но не достаточен для реализации информационной безопасности как надёжного инструмента, обеспечивающего защиту интересов объектов информационного взаимодействия от отдельного гражданина до государства и цивилизации в целом.

Модель интероперабельности и информационная безопасность

Эталонная модель интероперабельности (ГОСТ Р 55062–2021), приведенная на рисунке 1, представляет собой развитие семиуровневой базовой эталонной модели взаимосвязи открытых систем (ВОС) согласно ГОСТ Р ИСО/МЭК 7498-1, и образована тремя уровнями: техническим, семантическим и организационным.

6 Department of Defense Dictionary of Military and Associated Terms. Joint Pub. 1-02. 1994. – 633 p.

7 Handley M., Schulzrinne H., Schooler E., Rosenberg J. SIP: Session Initiation Protocol. Network Working Group. 1999. RFC 2453. URL: https://www.ietf.org/rfc/rfc2543.txt (date of request: 16.05.2025).

8 Department of Defense Dictionary of Military and Associated Terms. Joint Pub. 1-02. 1994, as amended through 10 january 2000 URL: https://www.bits.de/NRANEU/others/jp-doctrine/jp1_02(00).pdf (date of request: 16.05.2025).

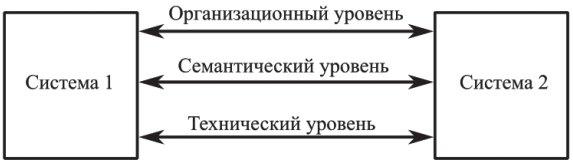


Рис. 1. Эталонная модель интероперабельности

Технический уровень – описывает синтаксис или форматы передаваемой информации, заостряя внимание на том, как представлена информация в коммуникационной среде. Технический уровень включает такие ключевые аспекты, как открытые интерфейсы, службы связи, интеграция данных и промежуточный слой программного обеспечения (Middleware), представление и обмен данными, службы доступности и защиты информации. Техническая интероперабельность достигается главным образом за счет использования стандартных протоколов связи типа TCP/IP.

Семантический уровень – описывает семантические аспекты взаимодействия, т. е. содержательную сторону информационного обмена. Семантическая интероперабельность позволяет системам комбинировать полученную информацию с другими информационными ресурсами и обрабатывать ее смысловое содержание. Семантическая интероперабельность достигается за счет применения стандартов типа XML (XSD, RDF, OWL).

Организационный уровень – акцентирует внимание на прагматических аспектах взаимодействия (деловых или политических). На этом уровне согласуются бизнес-цели и достигаются соглашения о сотрудничестве между административными органами, которые хотят обмениваться информацией, хотя имеют отличающиеся внутреннюю структуру и процессы. Организационная интероперабельность имеет своей целью удовлетворить требования сообщества пользователей: службы должны стать доступными, легко идентифицироваться, и быть ориентированными на пользователя. Организационная интероперабельность достигается не за счет применения стандартов (нормативно-технических документов), а за счет применения нормативно-правовых документов (соглашений, конвенций, договоров о сотрудничестве).

По отношению к объекту информационного взаимодействия различают внешнюю и внутреннюю интероперабельность. В ГОСТ Р 55062–2021 даны следующее определения:

- внешняя интероперабельность предприятия (external enterprise interoperability) – интероперабельность, которая определяет взаимодействие предприятия с другими предприятиями и конкурентоспособность предприятия на рынке;
- внутренняя интероперабельность предприятия (internal enterprise interoperability) – интероперабельность внутренней инфраструктуры (корпоративной системы) предприятия.

И дополнительно отметим, как в ГОСТ Р 55062–2021 определена интероперабельность предприятия (enterprise interoperability) – способность предприятий или находящихся в них сущностей (объектов) осуществлять эффективную связь и взаимодействие.

Эталонная модель интероперабельности предлагает разделение эффективного информационного взаимодействия на три уровня (рис. 1). Исследуем следующий вопрос: все ли уровни этого взаимодействия для реализации защиты интересов объектов и субъектов информационных отношений охватывает информационная безопасность, в том виде, в каком она позиционируется некоторыми экспертами и рядом нормативных документов?

Для поиска ответа на поставленный вопрос произведём сопоставление уровней определённых для интероперабельности с видами защиты информации определёнными в ГОСТ Р 50922-2006.

Правовая защита информации – защита информации правовыми методами, включающая в себя разработку законодательных и нормативных правовых документов (актов), регулирующих отношения субъектов по защите информации, применение этих документов (актов), а также надзор и контроль за их исполнением.

Техническая защита информации (ТЗИ) – защита информации, заключающаяся в обеспечении не криптографическими методами безопасности информации (данных), подлежащей (подлежащих) защите в соответствии с действующим законодательством, с применением технических, программных и программно-технических средств.

Криптографическая защита информации – защита информации с помощью ее криптографического преобразования.

Физическая защита информации – защита информации путем применения организационных мероприятий и совокупности средств, создающих препятствия для проникновения или доступа неуполномоченных физических лиц к объекту защиты.

Анализ показывает, виды информационной безопасности не включают семантический уровень информационного взаимодействия. Техническая, криптографическая и физическая защита информации относится к техническому уровню информационного взаимодействия, работа на техническом уровне информационной безопасности, во многих, но не во всех случаях, может вестись без учёта смысловой нагрузки защищаемой информации. Правовая защита информации относится к организационному

уровню, т.к. организационный уровень интероперабельности акцентирует внимание на прагматических аспектах взаимодействия (деловых или политических, и в том числе правовых). И надо отметить, что правовой уровень защиты информации не обходится без семантического, т.к. сама функция права не может быть реализована без вложения в право смысла. Поэтому, в неявном виде семантический уровень информационной безопасности имеется, но в явном виде, в виде, определяющем область научных интересов информационной безопасности – отсутствует.

Семантический уровень

Семантический уровень иначе уровень смыслов, в общем случае является ключевым для возможности реализации и обеспечения интересов и защиты сторон при информационном взаимодействии. Уровень смыслов необходим не только для внешнего межсистемного взаимодействия, но и для информационного взаимодействия внутри системы. Уровень смыслов определяет трактовку тех или иных частных и общих договорённостей, сообщений, соглашений. Самая значительная возможность семантического уровня это синтез новых смыслов и критический аудит имеющихся с целью повышения содержательной сути, а значит повышения качества информационного взаимодействия. В информационном межсистемном взаимодействии, побеждают более сильные идеи и смыслы, которые и определяют ход дальнейшей эволюции. Поэтому столь значима борьба за доминирование одних идей над другими. Так при колонизации метрополии навязывают свои смыслы и идеи колониям, как в период, предшествующий колонизации, так и после, для удержания статуса метрополии. Доминирующие на рынке компании определяют ценовую политику и способны регулировать спрос за счёт социальной инженерии, рекламы, лоббирования своих интересов в государственных органах управления и пр., тем самым навязывая свои идеи и смыслы социальным системам (рынкам) различного масштаба.

Один из способов оценки эффективности функционирования системы в условиях агрессивной внешней среды, основан на показателях скорости и точности достижения поставленной цели. Для социальных систем, цели, т.е. смыслы существования мотивы к развитию, обычно трансформируются в идеологию, задающую общий вектор развития и поддержанный отдельными частными целями и способствующими их достижению задачами. При отсутствии явно сформулированной цели, определяющей вектор движения системы, невозможно оценить эффективность движения, или наоборот можно дать любую оценку произвольному движению или топтанию на месте. А если такая цель не поставлена явно или сформулирована расплывчато, то означает ли это, что такой цели нет? Нет, не означает. Во-первых, у субъекта хозяйствования может иметься цель, но она не формулируется явно внутренними агентами или подменяется иными целями, не соответствующими реальным. Подмена цели обычно является следствием того, что фактическая цель является плохим мотиватором для элементов системы и потому не формулируется явно внутренними агентами системы. Во-вторых, цель для объекта, не имеющего собственной цели, будет поставлена из внешнего пространства. Следует отметить, что наличие собственной цели не гарантирует защиту от постановки внешней цели. Особенно, если в качестве объектов рассматривать достаточно крупные объекты хозяйственно-экономической деятельности как внутри отдельной страны, так и на международной арене. Понятно, что такие объекты, не имеющие собственной цели или имеющие слабую цель, в условиях борьбы за различные ресурсы не останутся без внимания других объектов и субъектов хозяйствования. Поэтому имеются основания полагать, что цель для объекта, не имеющего собственной цели или имеющего слабую цель, будет поставлена из внешнего пространства внешними субъектами, имеющими более сильную цель и располагающими ресурсами для её реализации. Так же имеются предпосылки полагать, что такая, внешняя, цель не всегда будет воспринята как действующий эффективный мотиватор для участников рассматриваемой системы и потому не формулируется явно внутренними управляющими агентами системы. Постановка цели внешними объектами и субъектами информационного взаимодействия и принуждение к её достижению может быть реализовано с помощью различных инструментов, в том числе инструментов манипуляции и принуждения (экономических, социо-психологических, политических, правовых, военных и др.), и так или иначе принуждая объект, не имеющий собственной цели или имеющий собственную слабую цель, действовать в соответствии с целями чужими.

Целевым концентратом, т.е. обобщением частных целей и задач всех государственных и действующих в правовом поле государства объектов можно назвать идеологию. Здесь уместно отметить Статью 13, часть 2 Конституции РФ «Никакая идеология не может устанавливаться в качестве государственной или обязательной» с учётом речи В. В. Путина на заседании дискуссионного клуба «Валдай» [1] в которой явного сказано о борьбе за суверенитет России. Отсутствие государственной идеологии есть прямое указание к отсутствию собственной цели государства.

Постановка цели внешними субъектами информационного взаимодействия далеко не всегда является деструктивным фактором для системы, в той или

иной степени находящейся под внешним управлением. Например, объекты хозяйственно-экономической деятельности, действующие в правовом поле на территории определённого государства, вынуждены соблюдать установленные внутри этого государства законы, а отдельное предприятие, ведущее свою деятельность в составе корпорации, выполняет задачи, поставленные корпорацией. В общем случае при информационном взаимодействии все вовлечённые объекты и субъекты в той или иной степени ограничены в реализации собственных целей, в предельном случае таким ограничителем являются законы природы (включая законы физические, экономические, социальные и пр). Если объекты претендуют на равноправие в построении информационных, экономических и прочих отношений, то цели, требующие общего участия сторон, должны быть выработаны как обоюдовыгодный компромисс. Причём степень уступок в компромиссе и определяется степенью принуждения со стороны бенефициара данных уступок. Эволюция цивилизации показывает, что процесс глобализации является неизбежным, и вероятно является законом развития сложных социальных систем, в условиях, глобализации смысл понятия «суверенитет» значительно отличается от смысла «суверенитета» в цивилизационный период до глобализации. Вот, например, президент «мирового гегемона» США и вероятно некоторая значительная часть его избирателей, считают, что США не имеют независимости и суверенитета, о чем можно сделать вывод из инаугурационной речи Д. Трампа: «Наш суверенитет будет восстановлен. Наша безопасность будет восстановлена… С этого дня Соединенные Штаты Америки будут свободной, суверенной и независимой нацией.» [17]. Вероятно, одним из самых значимых факторов суверенитета в современном мире, идущем по пути глобализации, является производственно-технологический суверенитет, о чём в той же речи говорит Д. Трамп: «Америка снова станет страной-производителем, и у нас есть то, чего никогда не будет ни у одной другой страны-производителя, – крупнейшие в мире запасы нефти и газа, и мы собираемся использовать их. Мы будем их использовать.» [17].

Определение уровня уступок и компромисса в условиях неизбежной глобализации является отдельной проблемой, которую необходимо исследовать и решать, в том числе на основе информационной безопасности с учётом всех уровней информационного взаимодействия.

Если говорить о человеческой цивилизации, в современном её состоянии, то распределение смыслов по значимости можно представить следующей моделью: смыслы, генерируемые цивилизацией в целом, т.е. совокупностью всех образующих цивилизацию объектов и субъектов, такие смыслы по модели В. И. Вернадского можно ассоциировать с некоторым планетарным явлением – ноосфера [18]; далее надгосударственные структуры, не отвечающие по обязательствам государств, но способные влиять на постановку целей государствами, например, Федеральная резервная система [19, 20]; отдельные государства; далее структуры в составе государства, коллективы и общественные организации; далее семья и отдельный человек. Нельзя сказать, что с точки зрения отдельного человека уровень смыслов, не имеет никакого значения, напротив, для развитого человека, реализующего свой творческий потенциал, цель жизни может быть основным мотиватором и двигателем его созидательного или разрушительного начала. Примером, формирования, развития смыслов и идей может быть жизнь и труд авиаконструктора Александра Сергеевича Яковлева [21], предпринимателя Генри Форда [22], учёного-электротехника Владимира Фёдоровича Мицкевича [23, 24] и др.; надо сказать – примеров постановки цели личностью и решительного стремления к ней немало в истории человечества. И не всегда эти цели были созидательные. Но раз в самом низу пирамиды целей и смыслов цель имеет столь великую силу, то ещё большую силу может иметь цель коллективная «Идеи становятся материальной силой, когда они овладевают массами» (К. Маркс). Таким образом, для всех перечисленных выше системных уровней цивилизации необходимы смыслы и цель, в первую очередь для эффективного и безопасного информационного взаимодействия, при котором на необходимый и достаточный уровень информационной безопасности могут рассчитывать все участники. При такой постановке вопроса информационная безопасность в соответствии с эталонной моделью интероперабельности можно так же подразделить на: техническую, семантическую и организационную.

Исследования в области интероперабельности, убедительно показывают, что информационное взаимодействие включает уровень смыслов, т.е. семантический уровень. Информационная безопасность, будучи не только техническим инструментом, но и научным направлением (о чем, например, говорят две научные специальности по номенклатуре ВАК: 2.3.6 Методы и системы защиты информации, информационная безопасность и 1.2.4 Кибербезопасность), должна систематически исследовать вопрос обеспечения интересов защищаемой стороны информационного взаимодействия, что означает исследовать угрозы информационного взаимодействия на всех уровнях: техническом, семантическом, организационном. Вероятно, не каждый специалист по информационной безопасности должен непосредственно заниматься совершенствованием

семантического и/или правого уровня информационной безопасности, но, по мнению авторов, каждый специалист по информационной безопасности, должен знать и понимать, что современная теория информационного взаимодействия основана на модели интероперабельности. Это, например, определяет особенности построения учебных программ для студентов и аспирантов соответствующих специальностей.

Информационная модель обеспечения безопасности

На основе приведённых рассуждений сформируем концепт обеспечения информационной безопасности объекта защиты. По мнению авторов, наиболее значимой задачей информационной безопасности является формирование и аудит целей, т.е. эволюции объекта защиты в условиях агрессивной среды. Следует различать средства и методы обеспечения информационного взаимодействия внутри объекта защиты, т.е. внутренние, и средства и методы обеспечения информационной безопасности за периметром объекта защиты, т.е. внешние. Цели объекта защиты формируются на будущее время, поэтому необходимо иметь прогноз, как о состоянии внешней среды, так и об изменении ресурсов и потенциала объекта защиты. Причём время прогнозирования должно быть достаточным для постановки и реализации тактических и стратегических целей, адекватных внешним условиям и собственному потенциалу объекта защиты. Современная физика придерживается той модели, при которой физический объект оказывает некоторое возмущение на окружающую его внешнюю среду, а внешняя среда на объект. При информационном взаимодействии с внешней средой объект защиты так же оказывает влияние на внешнюю среду, а внешняя среда воздействует на объект защиты. Поэтому цели объекта защиты должны быть такими, что бы производимые им возмущения внешней среды создавали наиболее благоприятные условия для реализации поставленных целей. С другой стороны необходимо минимизировать возмущающее воздействие внешней среды на объект защиты, препятствующее достижению целей объекта защиты.

Собственное (т.е. внутреннее) состояние объекта защиты так же должно способствовать достижению поставленной цели. Фактически любой информационный субъект, являющийся частью объекта защиты, оказывает возмущающее воздействие на другие частные информационные субъекты объекта, следовательно, на состояние объекта в целом. Поэтому необходим мониторинг, аудит и обеспечение информационной безопасности внутреннего информационного взаимодействия объекта защиты.

Условия внешней среды и состояние объекта защиты, являются не стационарными, но динамическими, т.е. изменяющимися с течением времени, поэтому для реализации эффективного и адекватного управления необходим мониторинг внешней среды и состояния объекта защиты, и своевременная корректировка информационного управляющего воздействия, т.е. корректировка положения объекта защиты в пространстве возможных состояний [25, 26].

Отметим, что не всё информационное взаимодействие сколько-нибудь сложного объекта защиты может наблюдаться средствами и методами информационной безопасности. Ещё меньшая часть информационного взаимодействия может управляться с применением методов информационной безопасности. Такие обстоятельства могут значительно затруднить реализацию информационной безопасности объекта защиты. А само информационное взаимодействие можно отобразить в виде диаграммы Эйлера-Вена (рис. 2).

Ранее было сказано, что основная функция, реализуемая при информационном взаимодействии – управление. Следовательно, более развёрнутая формулировка задачи информационной безопасности – определение целей объекта защиты и информационной безопасности объекта, обеспечение эффективного управления объектом защиты за счёт безопасности внутреннего и внешнего информационного взаимодействия на всех уровнях интероперабельности для достижения поставленной объектом цели.

Оценка управления осуществляется на основании оценки достижения поставленных целей. Следовательно, оценка информационной безопасности,

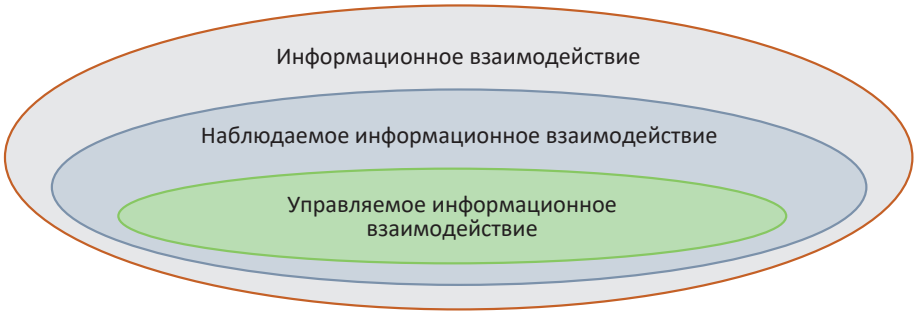


Рис. 2. Наблюдаемое и управляемое в общем информационном взаимодействии

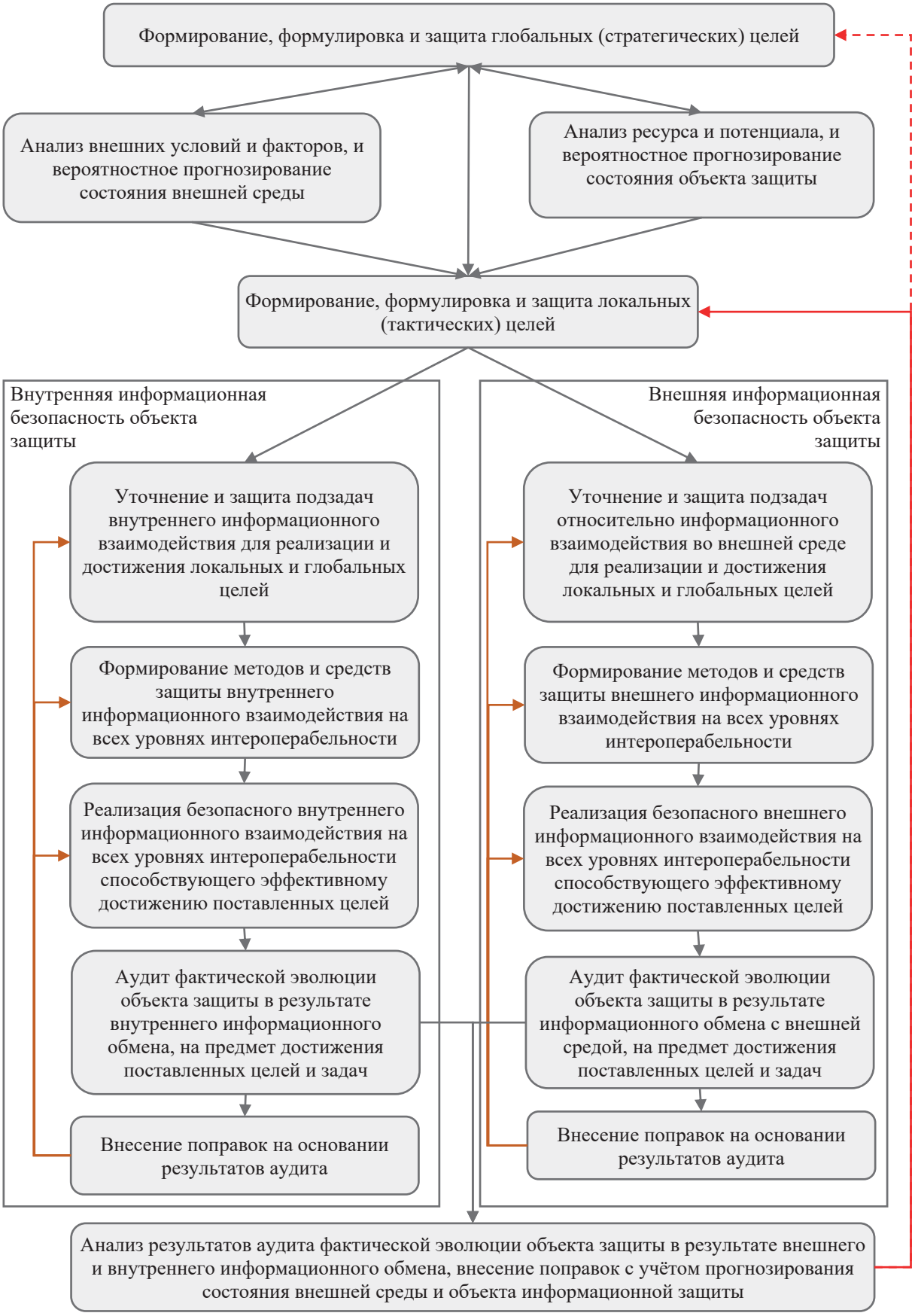


Рис. 3. Графическая модель обеспечения информационной безопасности объекта защиты

необходимая при аудите, формируется из оценок информационного взаимодействия на всех уровнях интероперабельности и способствования информационного взаимодействия достижению поставленной цели объекта защиты.

В соответствии с изложенными принципами сформирована графическая модель обеспечения информационной безопасности объекта защиты (рис. 3). Следует отметить, что данная модель является информационной и потому стрелками обозначены информационные потоки, а блоками ключевые информационные этапы обеспечения безопасности. Изменение стратегических целей объекта защиты может быть связано со значительными изменениями внутреннего состояния объекта защиты и/или внешней среды. Потому стратегические цели вынесены в отдельный блок и в случае устойчивой эволюции объекта защиты достаточно стабильны во времени, потому стрелка, обозначающая внесение поправок, обозначена пунктиром. Локальные (тактические) цели, напротив, изменяются динамично, учитывая состояние объекта защиты и внешней среды. Отдельно реализуются внешняя и внутренняя информационная безопасность объекта защиты. Корректировка подзадач, методов и средств реализации, проводится на основании аудита фактической эволюции объекта защиты в результате информационного взаимодействия внешнего и внутреннего.

Эталонная модель интероперабельности как основа классификации методов и средств информационной безопасности

В таблице 1 дана возможная модель информационной безопасности на основе эталонной модели интероперабельности. По строкам определены уровни интероперабельности (информационного взаимодействия), по столбцам определены уровни обеспечения информационной безопасности. В соответствующих клетках таблицы расположены виды обеспечения информационной безопасности, отвечающие уровню информационного взаимодействия.

Отдельного внимания требует модель аудита информационной безопасности (табл. 2). По строкам, как и в таблице 1, размещены уровни информационного взаимодействия. По столбцам – уровни аудита и оценки рисков информационной безопасности. В клетках таблицы даны соответствующие области аудита и оценки рисков информационной безопасности. Семантический уровень информационной безопасности обеспечивает смысловую защиту информации. Вопрос защиты смыслов требует дополнительного разбора, который в рамках одной публикации сделать затруднительно. В большинстве случаев задача формирования смыслового содержания информации выходит за пределы информационной безопасности, но необходимость сохранения смысла или выявление подмены и/или сокрытия

Таблица 1.

Модель обеспечения информационной безопасности на основе эталонной модели интероперабельности

		Уровни обеспечения информационной безопасности		
		Технический	**Семантический**	**Организационный**
Уровни интероперабельности	**Организационный**	Обеспечение защищённого документооборота	Терминологическое, методическое и методологическое обеспечение информационного взаимодействия объекта защиты	Правовое обеспечение информационной безопасности объекта защиты
	Семантический	Обеспечение конфиденциальности, целостности и доступности информации без учёта её смысловой нагрузки	Обеспечение информационной безопасности объекта защиты за счёт анализа смыслового наполнения при информационном взаимодействии	Формализация целей и задач информационной безопасности, согласно с целями и задачами объекта защиты
	Технический	Обеспечение реализации криптографической, физической и технической защиты	Обеспечение конфиденциальности, целостности и доступности смысловой нагрузки информации	Методическое и методологическое обеспечение технического уровня информационной безопасности

Таблица 2.

Модель угроз, аудита и оценки рисков информационной безопасности на основе эталонной модели интероперабельности

		Уровни обеспечения информационной безопасности		
		Технический	**Семантический**	**Организационный**
Уровни интероперабельности	**Организационный**	Аудит защищённого документооборота	Аудит, ревизия и совершенствование терминологического, методического и методологического обеспечение информационной безопасности	Правовая поддержка, аудита и оценки рисков информационной безопасности
	Семантический	Данные об эффективности методов и средств конфиденциальности, целостности и доступности для построения модели угроз, аудита и оценки рисков информационной безопасности	Комплексный анализ рисков внешнего и внутреннего информационного взаимодействия объекта защиты и прогнозирование на основе моделирования состояния его информационной безопасности. Смысловой аудит управляющей информации с точки зрения обеспечения безопасности объекта информационного взаимодействия.	Аудит и комплексная оценка эффективности результатов применения методов и средств информационной безопасности. Ревизия целей и задач информационной безопасности
	Технический	Аппаратные и аппаратно-программные средства реализации аудита, оценки рисков информационной безопасности и модели угроз	Аудит смысловой нагрузки информации на предмет её конфиденциальности, целостности и доступности	Аудит методического и методологического обеспечения технического уровня информационной безопасности

смысла, а также анализ соответствия смыслового наполнения информации объективному состоянию дел, вполне является задачей информационной безопасности и информационного противоборства.

Как, в целях обеспечения целостности информации производится, например, вычисление *хеш*-функции (от англ. hash function) для некоторого сообщения, и не обязательно эту работу проделывает специалист по информационной безопасности, но специалист владеющей пониманием критериев целостности информации. Так специалист владеющий пониманием целостности смысла информации, может заинтересоваться адекватностью, т.е. соответствию реальному положению дел, относительно, например, принятого сообщения, даже если хеш-функция данного сообщения показывает структурную целостность. Следуя логике защиты интересов стороны информационного взаимодействия, специалист по информационной безопасности должен понимать как те или иные смыслы, наполняющие информацию, повлияют на безопасность объекта защиты в настоящем времени и в будущем, в том числе в отдалённой перспективе, и насколько это влияние способствует эволюции состояния объекта защиты в заданном целевом направлении, т.е. достижению цели. Для такого понимания необходимо располагать целями и задачам объекта защиты интересов информационного взаимодействия на ближайшую и отдалённую перспективу. Например, для защиты государственной информационной безопасности, необходимо располагать целеуказующей идеологией государства, информационная безопасность которого

обеспечивается, с уточняющими комментариями и частными целями.

Как было показано ранее с увеличением значимости и при масштабировании отдельных смыслов формируемых относительно некоторого значительного субъекта хозяйственной деятельности, смыслы трансформируются в идеологию и определяют с одной стороны эволюцию рассматриваемого объекта, а с другой стороны позволяют производить аудит этой эволюции и ключевую управляющую роль в этом процессе имеет смысловой уровень внешнего и внутреннего информационного взаимодействия.

Итак, по мнению авторов, семантический уровень информационной безопасности заключается в защите смыслового наполнения информации. Вопрос защиты и формирования смыслового наполнения информации, является не однозначным и зависит от личностных качеств и идеологических принципов лица или группы лиц осуществляющих смысловое наполнение. Поэтому высокий уровень личной и коллективной ответственности приходится на тех, кто принимает решения о смысловом наполнении и распространении информации, особенно если информация распространяется массово и влияет на мировоззрение значительного числа людей и/или имеет стратегическое значение для развития социальных систем различного масштаба. В этом смысле компетенции специалиста по информационной безопасности имеют характер ценза, способного оценить риски деструктивного информационного влияния. Предмет таких компетенций является не простой задачей, имеет значительное число аспектов, обсуждение которых выходит за рамки данной работы. На сегодняшний день существуют органы, осуществляющие в той или иной степени смысловое регулирование, но делается это зачастую не системно и без опоры на явно обозначенные цели. Но как было показано ранее, это не означает, что целей нет, отсутствие явно сформулированных целей означает, что либо эти цели неизвестны тем, кто реализует частные задачи, либо что их явная формулировка нежелательна по тем или иным причинам.

Обсуждение

Выполненный в работе анализ и синтез моделей информационной безопасности построен на основе исследования с одной стороны современного состояния проблемы и способности к информационному взаимодействию, получившего устоявшееся англоязычное название интероперабельность; с другой стороны современного состояния в области исследований информационной безопасности. Анализ показывает, что интероперабельность, будучи молодым научным направлением за счёт усилий многих учёных по всему миру, выработала устойчивую и обоснованную модель информационного взаимодействия. Сопоставление моделей интероперабельности и информационной безопасности, показывает, что информационная безопасность, не рассматривает в явном виде угрозы и риски при информационном взаимодействии которые связаны с семантическим уровнем информационного взаимодействия, что по мнению авторов является значительной угрозой, особенно для крупных предприятий, организаций и суверенитета Родины. Необходимо отметить, что учёт смысловой составляющей информации в решении задач информационной безопасности не должен обернуться обычной цензурой, запретом, ограничением доступа к информации, подобных явлений и так предостаточно. Напротив, только при максимальной открытости и доступности знаний и объективной информации возможен баланс и безопасность информационной среды. Авторы предлагают за счёт включения в сферу компетенций специалистов по информационной безопасности семантического уровня информационного взаимодействия расширить аналитический инструментарий информационной безопасности, что в свою очередь позволит повысить эффективность информационной безопасности, как в масштабах отдельного предприятия, так и в масштабах страны в целом. Например, анализ и оценка поставленных для объекта защиты целей и оценки результатов её достижения, выявление причин и следствий срыва поставленных целей, выявление возможностей и наличия необходимых ресурсов достижения поставленной цели, критический анализ инструментария и параметров оценивания. Анализ содержательной части имеющихся и вновь принимаемых законов, постановлений на предмет их влияния на национальную безопасность в различных секторах государственной жизни. Разработка инструментария и оценка эффективности информационно-управляющей деятельности организаций и отдельных управленцев. По мнению авторов наиболее опасные информационные угрозы содержаться в смысловом наполнении информации, когда информация не соответствуют действительности, а поставленные цели и задачи не соответствуют фактическому положению дел и объективным возможностям и методам их достижения, и при этом сами сообщения, обеспечивающие необходимый информационный обмен: конфиденциальны, целостны и доступны, но небезопасны в смысле информационно-управляющего эффекта, который эти сообщения осуществляют.

Выводы

На основе анализа научных направлений интероперабельность и информационной безопасность, предлагается дополнить область интересов информационной безопасности семантическим уровнем,

в соответствии с эталонной моделью интероперабельности.

- Выполнен анализ угроз информационной безопасности объекту защиты реализуемых на семантическом уровне информационного взаимодействия.
- В ходе исследований доказана необходимость информационной безопасности на семантическом уровне для обеспечения полноты защиты информационного взаимодействия и удовлетворения интересов объекта информационной защиты.
- Предложена информационная модель разработки и реализации методов информационной безопасности, способствующая достижению целей объекта защиты.
- Предложена модель информационной безопасности на основе эталонной модели интероперабельности.
- Предложена модель аудита и оценки рисков информационной безопасности на основе эталонной модели интероперабельности.

Литература

1. Заседание дискуссионного клуба «Валдай» (дата обращения: 11.11.2024) URL: http://www.kremlin.ru/events/president/news/75521
2. Жаринов И. О. Стек сквозных цифровых технологий как фактор инновационной модернизации оборонно-промышленного комплекса России // Военный академический журнал. 2024. № 3 (43). С. 133–139.
3. Алешковский И. А. Демографический кризис как угроза национальной безопасности России // Век глобализации, 2(10). 2012. 96–114 с.
4. Третьяк О. А., Румянцева М. Н. Сетевые формы межфирменной кооперации: подходы к объяснению феномена // Российский журнал менеджмента. 2003. Т. 1. № 2. С. 25–50.
5. Грановеттер М. Сила слабых связей // Экономическая социология. 2009. Т. 10. № 4. С. 31–50.
6. Введение в теорию управления организационными системами / В. Н. Бурков, Н. А. Коргин, Д. А. Новиков / – М.: Либроком, 2009. 264 с.
7. Поле битвы – киберпространство. Теория, приемы, средства, методы и системы ведения информационной войны / С. Н. Гриняев / – М.: Харвест, 2004. 426 с.
8. О диалекте сдерживания и предотвращения военных конфликтов в информационную эру / И. Н. Дылевский, В. О. Запивахин, С. А. Комов, С. В. Коротков, А. А. Кривченко // Военная мысль. 2016. № 7. С. 3–11.
9. Информационное противоборство и радиоэлектронная борьба в сете-центрических войнах начала XXI века / С. И. Макаренко / – С.-Пб.: Наукоемкие технологии, 2017. 546 с.
10. Макаренко С. И., Олейников А. Я., Черницкая Т. Е. Модели интероперабельности информационных систем // Системы управления, связи и безопасности. 2019. № 4. С. 215–245. DOI: 10.24411/2410-99162019-10408.
11. Гришенцев А. Ю., Коробейников А. Г., Дукельский К. В. Метод численной оценки технической интероперабельности. Кибернетика и программирование. 2017. № 3. С. 23–38.
12. Гришенцев А. Ю., Коробейников А. Г. Средства интероперабельности в распределенных геоинформационных системах. Журнал радиоэлектроники. 2015. № 3. С. 1–18.
13. Интероперабельность человеко-машинных интерфейсов. / С. И. Макаренко / С.-Пб.: Наукоемкие технологии, 2023. 185 с.
14. Вопросы создания единого информационного пространства в системе здравоохранения РАН / Н. Г. Гончаров, Я. И. Гулиев, Ю. В. Гуляев [и др.] // Информационные технологии и вычислительные системы. 2006. № 4. С. 83–95.
15. Информационная безопасность. / С. И. Макаренко / Ставрополь: СФ МГГУ им. М. А. Шолохова, 2009. 372 с.
16. Батоврин В. К., Гуляев Ю. В., Олейников А. Я. Обеспечение интероперабельности – основная тенденция в развитии открытых систем // Информационные технологии и вычислительные системы. 2009. № 5. С. 7.
17. Выдержки из инаугурационной речи президента Дональда Трампа, касающиеся внешней политики. Посольство и консульства США в Российской Федерации. (дата обращения: 03.05.2025) URL: https://ru.usembassy.gov/ru/president-donald-trumps-inaugural-address-ru/.
18. Научная мысль как планетное явление. Избранные труды / В. И. Вернадский / Сост. Г.П. Аксенов. - М.: РОССПЭН, 2010. С.: 580–742.
19. Эпоха потрясений / А. Гринспен / – М.: Альпина Бизнес Букс, 2007. 90 с.
20. Кризис. Как это делается / Н. Стариков / – С.-Пб.: Питер, 2010. 304 с.
21. Цель жизни. Записки авиаконструктора. 5-е изд., переработ. и доп. / А. С. Яковлев / – М.: Политиздат, 1987. 511 с.
22. Моя жизнь, мои достижения / Г. Форд / Пер. под ред. В. А. Зоргенфрея; предисл. Н. С. Лаврова. – Л.: Время, 1924. 326 с.
23. Выдающийся русский ученый-электрик академик Владимир Федорович Миткевич / М. А. Шателен, Л. Р. Нейман, И. А. Зайцев [и др.] // Электричество. 2005. № 1. С. 89–91.
24. Коровкин Н. В. Академик Владимир Федорович Миткевич (к 150-летию со дня рождения). Электричество. 2022. № 8. С. 65–69. DOI 10.24160/0013-5380-2022-8-65-69.
25. Заколдаев Д. А., Гришенцев А. Ю. Методология моделирования и обеспечения информационной безопасности при управлении ресурсами // Вестник компьютерных и информационных технологий. 2021. Т. 18. № 4 (202). С. 45–52. DOI 10.14489/vkit.2021.04.pp. 045–052
26. Заколдаев Д. А., Гришенцев А. Ю. Формальная модель обеспечения информационной безопасности при управлении ресурсами на производствах // Системы управления, связи и безопасности. 2021. № 1. С. 33–61. DOI 10.24411/2410-9916-2021-10102.

INTEROPERABILITY AS A BASIS FOR SYSTEMATIZATION OF INFORMATION SECURITY METHODS AND MEANS

Grishentsev A. Yu.[9], Korovkin N. V.[10], Korobeynikov A. G.[11]

Keywords: *information protection, information interaction, open systems, models, standards.*

Purpose of the study: *development of the theoretical foundations of information security through sound systematization, methods and means of information security based on the concept of interoperability.*

Methods of research: *analysis of information interaction and threats in information interaction based on a standardized reference model of interoperability and synthesis of a systematic structured model of information security methods and tools in the context of the concept of interoperability.*

Result's: *based on the analysis of scientific areas of interoperability and information security, it is proposed to supplement the field of interests of information security with a semantic level, in accordance with the reference model of interoperability. The analysis of information security threats to the object of protection implemented at the semantic level of information interaction has been performed. In the course of research, the need for information security at the semantic level has been proved to ensure the completeness of information interaction protection and to satisfy the interests of the information protection object. An information model for the development and implementation of information security methods is proposed to help achieve the objectives of the object of protection. A model of information security based on a reference model of interoperability is proposed. A model of information security audit and risk assessment based on a reference model of interoperability is proposed.*

Scientific novelty: *It consists in a new approach to systematization of methods, means and increasing the sphere of interests of information security based on modern scientific ideas about the levels of information interaction in accordance with the concept of interoperability.*

References

1. Zasedanie diskussionnogo kluba «Valdaj» (2024, November 07). URL: http://www.kremlin.ru/events/president/news/75521.
2. Zharinov, I. O. (2024). Stek Skvoznyx Cifrovyx Texnologij Kak Faktor Innovacionnoj Modernizacii Oboronno-Promyshlennogo Kompleksa Rossii. Voennyj Akademicheskij Zhurnal, 2(10), 133–139.
3. Aleshkovskij, I. A. (2012). Demograficheskij Krizis Kak Ugroza Nacional'noj Bezopasnosti Rossii. Vek Globalizacii, 2(10), 96–114. https://www.socionauki.ru/journal/articles/147957/.
4. Tret'yak, O. A., & Rumyanceva, M. N. (2003). Setevye Formy Mezhfirmennoj Kooperacii: Podxody k Ob''yasneniyu Fenomena. Rossijskij Zhurnal Menedzhmenta, 2, 25–50. https://rjm.spbu.ru/article/view/812/707.
5. Granovetter, M. (2003). Sila Slabyx Svyazej (Z. V. Kotel'nikova, Trans.). Ekonomicheskaya Sociologiya, 10(4), 31–50. https://ecsoc.hse.ru/2009-10-4/26591138.html.
6. Burkov V. N., Korgin N. A., Novikov D. A. (2009). Vvedenie v teoriyu upravleniya organizacionnymi sistemami. Librokom. 264 p.
7. Grinyaev S. N. (2004). Pole bitvy – kiberprostranstvo. Teoriya, priemy, sredstva, metody i sistemy vedeniya informacionnoj vojny. Harvest. 426 p.
8. Dylevskij I. N., Zapivaxin V. O., Komov S. A., Korotkov S. V. & Krivchenko A. A. (2016) O dialekte sderzhivaniya i predotvrashheniya voennyx konfliktov v informacionnuyu eru. Voennaya mysl'. 7, 3–11.
9. Makarenko S. I. (2017). Informacionnoe protivoborstvo i radioelektronnaya bor'ba v sete-centricheskix vojnax nachala XXI veka. Naukoemkie texnologii. 546 p.
10. Makarenko S. I., Olejnikov A. Ya., Chernickaya T. E. (2019). Modeli interoperabel'nosti informacionnyx system. Sistemy upravleniya, svyazi i bezopasnosti. 4, 215–245. DOI: 10.24411/2410-99162019-10408.
11. Grishencev A. Yu., Korobejnikov A. G., Dukel'skij K. V. (2017). Metod chislennoj ocenki texnicheskoj interoperabel'nosti. Kibernetika i programmirovanie. 3, 23–38.
12. Grishencev A. Yu., Korobejnikov A. G. (2015). Sredstva interoperabel'nosti v raspredelennyx geoinformacionnyx sistemax. Zhurnal radioelektroniki. 3, 1–18.
13. Makarenko S. I. (2023). Interoperabel'nost' cheloveko-mashinnyx interfejsov. Naukoemkie texnologii. 185 p.
14. Goncharov N. G., Guliev Y. I., Gulyaev Y. V., Kavinskaya A. A., Olejnikov A. Y., & Xatkevich M. I. (2006). Voprosy Sozdaniya Edinogo Informacionnogo Prostranstva v Sisteme Zdravooxraneniya RAN. Informacionnye Texnologii i Vychislitel'nye Sistemy, 4, 83–95. https://jitcs.frccsc.ru/arhiv/2006/release_4/voprosy_sozdaniya_edinogo_informatsionnogo_prostranstva_v_sisteme_zdravoohraneniya_ran.html.

9 Alexey Yu. Grishentsev, Dr.Sc. of Technical Sciences, Associate Professor, Corresponding Member of the Academy of Electrotechnical Sciences of the Russian Federation, Associate Professor of the Federal State Autonomous Educational Institution of Higher Education ITMO National Research University. St. Petersburg, Russia. E-mail:AGrishentsev@yandex.ru

10 Nikolay V. Korovkin, Dr.Sc. of Technical Sciences, Professor, Full Member of the Academy of Electrotechnical Sciences of the Russian Federation, Professor of the Federal State Autonomous Educational Institution of Higher Education Peter the Great St. Petersburg Polytechnic University. St. Petersburg, Russia. E-mail: Nikolay.Korovkin@gmail.com

11 Anatoly G. Korobeynikov, Dr.Sc. of Technical Sciences, Professor, Deputy Director for Science of the St. Petersburg Branch of the Pushkov Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation of the Russian Academy of Sciences, Professor of the Federal State Autonomous Educational Institution of Higher Education ITMO National Research University. St. Petersburg, Russia. E-mail: Korobeynikov_A_G@mail.ru

15. Makarenko S. I. Informacionnaya bezopasnost'. (2009). SF MGGU im. M. A. Sholoxova. 372 p.
16. Batovrin V. K., Gulyaev Y. V., & Olejnikov A. Y. (2006). Obespechenie Interoperabel'nosti – Osnovnaya Tendenciya v Razvitii Otkrytyx Sistem. Informacionnye Texnologii I Vychislitel'nye Sistemy, 2009. № 5. Pp. 7. http://www.jitcs.ru/index.php?option=com_content&view=article&id=310.
17. Trump, D. (2025, January 22). President Donald Trump's Inaugural Address. Ru.Usembassy.Gov. https://ru.usembassy.gov/president-donald-trumps-inaugural-address/.
18. Vernadskij V. I. (2010). Nauchnaya Mysl' Kak Planetnoe Yavlenie. Izbrannye Trudy (Aksenov G. P.). ROSSPEN. 742 p.
19. Grinspen, A. (2007). Epoxa Potryasenij. Al'pina Biznes Buks. 90 p.
20. Starikov N. (2010). Krizis. Kak eto delaetsya. Piter. 304 p.
21. Yakovlev A. S. (1987). Cel' Zhizni. Zapiski Aviakonstruktora (5th ed.). Politizdat. 511 p.
22. Ford, G. (1922). My Life and Work. Stone Hedge. 304 p.
23. Shatelen, M. A., Nejman, L. R., & Zajcev, I. A. (2005). Vydayushhijsya Russkij Uchenyj-Elektrik Akademik Vladimir Fedorovich Mitkevich. Elektrichestvo, 1, 89–91.
24. Korovkin, N. V. (2022). Akademik Vladimir Fedorovich Mitkevich (k 150-letiyu so dnya rozhdeniya). Elektrichestvo, 2, 65–69. DOI 10.24160/0013-5380-2022-8-65-69.
25. Zakoldaev, D. A., Grishentsev, A. Yu. (2021). Methodology for modeling and ensuring information security in resource management. Herald of computer and information technologies, 4(202), 45–52. DOI: 10.14489/vkit.2021.04.pp.045-052.

МЕТОДИКА ОЦЕНКИ ОПАСНОСТИ ДЕСТРУКТИВНЫХ ПРОГРАММНЫХ ВОЗДЕЙСТВИЙ НА АВТОМАТИЗИРОВАННЫЕ СИСТЕМЫ СПЕЦИАЛЬНОГО НАЗНАЧЕНИЯ ОРГАНОВ ВНУТРЕННИХ ДЕЛ

Мельников А. В.[1], Кобяков Н. С.[2]

DOI: 10.21681/2311-3456-2025-5-28-40

Цель исследования: *моделирование показателя опасности деструктивных программных воздействий, с учетом актуальности поведенческих паттернов вредоносных программ для автоматизированных систем специального назначения органов внутренних дел.*

Методы исследования: *для формирования моделей оценки опасности деструктивных программных воздействий и определения численных значений признаков АССН ОВД используется метод анализа иерархий.*

Результат исследования: *определены базовые и частные признаки АССН ОВД, характеризующие актуальность поведенческих паттернов вредоносных программ в зависимости от функциональных особенностей АССН ОВД. Разработаны базовые и частные модели оценки опасности деструктивных программных воздействий на АССН ОВД с учетом актуальности поведенческих паттернов вредоносных программ. Разработан алгоритм планирования и реализации процессов жизненного цикла АССН ОВД в условиях деструктивных программных воздействий. Выполнена верификация разработанной методики на примере формирования моделей оценки опасности деструктивного программного воздействия вредоносных программ класса «Вредоносные утилиты» на тестовую автоматизированную систему специального назначения. Верификация разработанных моделей выполнена на тестовом наборе данных, сформированном путем опроса экспертов.*

Практическая значимость: *разработанная методика может быть использована администраторами безопасности автоматизированных систем специального назначения при оценке опасности деструктивных программных воздействий и определении целей и перечня реализуемых мер обеспечения защиты информации при появлении неизвестных вредоносных программ.*

Ключевые слова: *вредоносные программы, признаки автоматизированных систем, защита информации, метод анализа иерархий.*

Введение

В настоящее время цифровизация процессов обработки информации в силовых ведомствах при всех ее преимуществах приводит к повышению активности злоумышленников по нанесению ущерба информации, которая хранится и обрабатывается в автоматизированных системах специального назначения органов внутренних дел (АССН ОВД). АССН ОВД – это система, состоящая из комплекса средств автоматизации оперативно-служебной и (или) повседневной деятельности, реализующая информационную технологию выполнения установленных функций, а также сотрудников органов внутренних дел, обеспечивающих её функционирование, с учетом требований по защите информации.

Одним из важных критериев, который необходимо учесть при формировании моделей оценки опасности деструктивных программных воздействий (ООДПВ) на АССН ОВД (J), это их функциональные особенности (признаки) (H,F), например, как это учтено в методическом документе ФСТЭК России[3]. В зависимости от признаков АССН ОВД можно определить актуальные поведенческие паттерны вредоносных программ (p), для реализации деструктивного программного воздействия. Под поведенческими паттернами вредоносных программ понимаются деструктивные функции, реализуемые вредоносной программой. В рамках работы будут рассмотрены базовые признаки АССН ОВД, характеризующие

1 Мельников Александр Владимирович, доктор технических наук, доцент, профессор кафедры автоматизированных информационных систем органов внутренних дел Воронежского института Министерства внутренних дел Российской Федерации, г. Воронеж, Россия. ORCID: https://orcid.org/0000-0001-5080-1162. E-mail: meln78@mail.ru

2 Кобяков Николай Сергеевич, адъюнкт кафедры автоматизированных информационных систем органов внутренних дел Воронежского института Министерства внутренних дел Российской Федерации, г. Воронеж, Россия. ORCID: https://orcid.org/0000-0002-4950-7879. E-mail: kkobyakov1234@gmail.com

3 Методический документ «Методика оценки уровня критичности уязвимостей программных, программно-аппаратных средств». Утвержден ФСТЭК России 28 октября 2022 г.

классы защищенности автоматизированных систем (подключение к сервису электронной почты и др.) и частные признаки, характеризующие конкретную автоматизированную систему (тип используемой операционной системы и др.).

Авторы в работах [1, 2] отмечают необходимость постоянного совершенствования системы защиты информации автоматизированных систем, в том числе от новых угроз. В работе [3] описаны следующие подходы к формированию требований в области информационной безопасности:

1. Экспертный анализ угроз безопасности информации, их идентификация, с последующей обработкой рисков и их снижения до приемлемого уровня.
2. Распространение на систему действия некоторого набора нормативных документов, в которых требования по информационной безопасности заранее определены.

Также, на практике может применяться комбинированный подход. В работах [4–8] представлены результаты исследований по моделированию угроз безопасности информации, но, в них не рассматриваются вопросы оценки опасности деструктивных программных воздействий с учетом признаков автоматизированных систем.

Для реализации достаточных дополнительных мер защиты информации в АССН ОВД необходимо оценить опасность деструктивных программных воздействий вредоносных программ. В рамках данной работы для формирования моделей ООДПВ будут использованы экспертные и многокритериальные методы принятий решений.

Авторы в работе [9] рассматривают современные подходы к моделированию с использованием метода анализа иерархий и делают вывод о том, что применение метода анализа иерархий, в ситуациях, когда исследуемая область характеризуется связанными признаками, может привести к ошибкам при верификации моделей. В работе [10] разработан численный метод модификации моделей, разработанных на основе метода анализа иерархий, с использованием искусственной нейронной сети. Данный численный метод может быть использован для модификации моделей, в которых определены пары связанных признаков, совместная реализация которых, приводит к повышению значения показателя качества. Применение данного метода позволяет учесть связь признаков при формировании моделей ООДПВ с использованием метода анализа иерархий и повысить точность сформированных моделей.

Цель исследования

Моделирование оценки опасности деструктивных программных воздействий на АССН ОВД с учетом их базовых и частных признаков.

Для достижения цели работы необходимо решить следующие задачи:

1. Определить базовые и частные признаки АССН ОВД, характеризующие актуальность поведенческих паттернов вредоносных программ в зависимости от функциональных особенностей АССН ОВД.
2. Выполнить моделирование оценки опасности деструктивных программных воздействий с учетом базовых, частных признаков АССН ОВД.
3. Разработать алгоритм планирования и реализации процессов жизненного цикла АССН ОВД в условиях деструктивных программных воздействий.
4. Выполнить вычислительный эксперимент, по оценке опасности деструктивных программных воздействий на АССН ОВД.

Порядок разработки моделей ООДПВ на АССН ОВД представлен на рисунке 1.

Для формирования моделей оценки опасности деструктивных программных воздействий необходимо использовать следующие исходные данные на каждом этапе, в соответствии с рисунком 1:

1. **Данные о вредоносных программах.** Вредоносные программы реализуют характерные для них поведенческие паттерны. Множество поведенческих паттернов $P = \{p_1, p_2, ..., p_n\}$, n – количество поведенческих паттернов вредоносных программ.

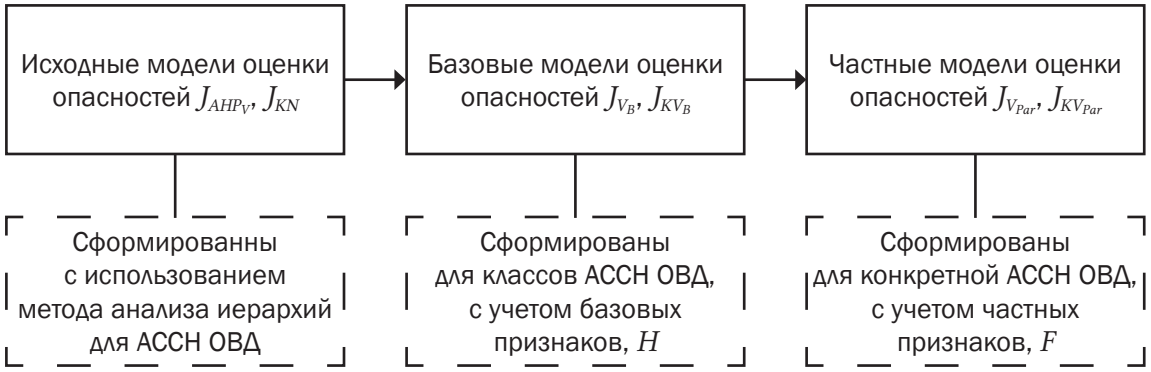


Рис. 1. Порядок разработки моделей оценки опасности деструктивных программных воздействий

Исходными моделями для формирования базовых и частных являются модели оценки опасности деструктивных программных воздействий: без учета связи поведенческих паттернов (J_{AHP_V}), сформированная в результате исследования [11], и с учетом связи поведенческих паттернов (J_{KN}), исследование [10].

2. **Данные об экспертной группе**. Экспертная группа формируется в соответствии с требованиями методического документа ФСТЭК России. Множество экспертов $T = \{t_1, t_2, ..., t_\tau\}$, τ – количество экспертов в экспертной группе.
3. **Данные об АССН ОВД для формирования базовых моделей**. Для формирования базовых моделей (J_{V_B}, J_{KV_B}) экспертной группой будет рассмотрен набор базовых признаков, $H = \{h_1, h_2, ..., h_\beta\}$, где β – количество базовых признаков, от которых зависит актуальность деструктивных программных воздействий вредоносных программ и мультимножество значений данных признаков для тестовой АССН ОВД $H^* = \{h_1^*, h_2^*, ..., h_\beta^*\}$. В результате работы экспертной группы формируется мультимножество значений базовых признаков для класса защищенности АС $\Psi = \{\psi_1, \psi_2, ..., \psi_n\}$.
4. **Данные об АССН ОВД для формирования частных моделей**. Для формирования частных моделей ($J_{V_{Par}}$, $J_{KV_{Par}}$) экспертной группой будет рассмотрен набор частных признаков $F = \{f_1, f_2, ..., f_\varphi\}$, φ – количество частных признаков, от которых зависит актуальность деструктивных программных воздействий вредоносных программ и мультимножество значений данных признаков для тестовой АССН ОВД $F^* = \{f_1^*, f_2^*, ..., f_\varphi^*\}$. В результате работы экспертной группы формируется мультимножество значений частных признаков для конкретной АССН ОВД $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$.

Допущения и ограничения, принятые в методике:

- на период эксплуатации состав компонентов АС остается неизменным;
- оцениваемые вредоносные программы реализуют поведенческие паттерны, описанные в работе [11]. В случае появления новых поведенческих паттернов необходимо уточнение исходных, базовых и частных моделей;
- в рамках данной работы будут рассмотрены вредоносные программы, используемые при проведении нецелевых компьютерных атак.
- в связи с большим количеством исследуемых поведенческих паттернов вредоносных программ (более 25) в работе будет рассмотрен процесс оценки опасности деструктивных программных воздействий только класса «Вредоносные утилиты».

Методика может быть применена для ситуаций, когда возможна реализация деструктивных функций вредоносных программ, сигнатуры которых не определены средствами антивирусной защиты (САВЗ) АС.

1. Формирование группы экспертов

Для каждой АС необходимо сформировать группу экспертов для оценки процессов, связанных с защитой информации. В соответствии с рекомендациями методического документа ФСТЭК России, формирование моделей оценки опасности деструктивного программного воздействия вредоносных программ на автоматизированные системы должно выполняться экспертной группой. В экспертную группу по результатам исследования [12] для оценки признаков АССН ОВД рекомендуется включить:

1. Должностные лица, ответственные за обеспечение безопасности информации, обрабатываемой в АССН.
2. Должностные лица, ответственные за функционирование ИТ-инфраструктуры АССН.
3. Должностные лица, выполняющее свои должностные обязанности в ходе эксплуатации АССН (хранение, обработка информации).

В рамках работы предлагается включить в экспертную группу будут 8 экспертов ($\tau = 8$). Данный набор экспертов обеспечивает выполнение требований по количеству и порядку подчиненности членов экспертной группы.

2. Формирование базовых моделей оценки опасности деструктивных программных воздействий

В общем виде модель оценки показателя опасности, сформированная с использованием метода анализа иерархий имеет вид:

$$J_{AHP} = \sum_{i=1}^{n} w_{AHP_i} \cdot p_i, \quad (1)$$

где: w_{AHP_i} – весовые коэффициенты поведенческих паттернов вредоносных программ; p_i – поведенческие паттерны вредоносных программ. Данная переменная принимает значение 1, если поведенческий паттерн реализован в вредоносной программе и значение 0, если не реализован; n – количество поведенческих паттернов вредоносных программ.

Далее, необходимо выполнить нормировку весовых коэффициентов поведенческих паттернов и нормализацию модели для приведения значения показателя опасности к лингвистической шкале. Для нормализации модели (1) необходимо разделить на значение самого высокого показателя опасности ($J_{\max}$) и умножить на максимальное значение лингвистической шкалы (A).

Исходная модель, без учета связи поведенческих паттернов, будет иметь вид:

$$J_{AHP_V} = \mathrm{A} \cdot \frac{J^*_{AHP_V}}{J_{\max}} = \frac{\sum_{i=1}^{n} w_i \cdot p_i}{J_{\max}}, \quad (2)$$

где w_i – нормализованные весовые коэффициенты поведенческих паттернов; $J^*_{AHP_V}$ – значение показателя опасности, рассчитанное с использованием классической модели, сформированной на основе метода анализа иерархий; $J_{\max}$ – максимальное значение опасности вредоносной программы (рассчитывается значение $J^*_{AHP_V}$ для существующих вредоносных программ и выбирается максимальное).

Для оценки опасности деструктивных программных воздействий разработана лингвистическая шкала по аналогии со стандартом CVSS v3.1 в диапазоне [0–10]. Следовательно, для формулы (2) значение $A = 10$.

Также в случае, если для исследуемого класса вредоносных программ характерны связанные поведенческие паттерны, то необходимо учесть данный факт в исходной модели. Исходная модель с учетом связи поведенческих паттернов будет иметь вид:

$$J_{KN} = 10 \cdot \frac{J^*_{KN}}{J_{KN_{\max}}} = \frac{\sum_{i=1}^{n} w_i \cdot p_i + \sum_{j=1}^{u} K_j}{J_{KN_{\max}}}, \quad (3)$$

где J^*_{KN} – значение показателя опасности, рассчитанное с использованием не нормализованной модели, с учетом связи поведенческих паттернов; $J_{KN_{\max}}$ – значение опасности самой опасной вредоносной программы; K – коэффициент, характеризующий связь поведенческих паттернов вредоносных программ; u – количество пар связанных поведенческих паттернов вредоносных программ.

Базовые модели формируются на основе исходных и предназначены для ООДПВ вредоносных программ с учетом класса защищенности АС. В работе [13] определена классификация АССН ОВД исходя из требований руководящих документов по классификации информационных и автоматизированных систем. Результаты работы представлены в таблице 1.

2.1. Определение весовых коэффициентов для классов АССН ОВД

Для каждого класса АССН необходимо определить весовой коэффициент, который будет учтен при построении частных моделей. Для этого предлагается выполнить эксперимент с использованием метода анализа иерархий, предложенный Т. Саати. После обобщения результатов опроса получена таблица парных сравнений, представленная в таблице 2.

С использованием программного обеспечения Mathcad получен первый собственный вектор матрицы парных сравнений $v_1 = (3{,}30193; 1{,}71712; 1)$. Выполнение нормировки первого собственного вектора может быть выполнено путем деления значения

Таблица 1.

Классы защиты (уровни доверия) средств защиты информации и соответствующие классы защищенности (КЗ) информационных систем (ИС)

КЗ ФСТЭК	Предназначение	КЗ ИСПДН	КЗ ГИС	КЗ ИС ОП	КЗ АССН ОВД
1, 2, 3	Предназначен для установки в средствах вычислительной техники и автоматизированных системах, обрабатывающих сведения, составляющие государственную тайну	–	–	–	1, 2, 3
4	Предназначен для установки в средствах вычислительной техники и автоматизированных системах, входящих в состав государственных информационных систем, информационных систем общего пользования и информационных систем, обрабатывающих персональные данные.	1	1	2	4
5		2	2	–	5
6		3	3	–	6

Таблица 2.

Парные сравнения классов АССН

	4 класс	5 класс	6 класс
4 класс	1	2	3
5 класс	1/2	1	2
6 класс	1/3	1/2	1

каждого элемента на их сумму [14]. Получим следующие весовые коэффициенты D для классов АССН ОВД:

4 класс АССН – 0,54;
5 класс АССН – 0,3;
6 класс АССН – 0,16.

Для определения итоговых весовых коэффициентов классов АССН ОВД необходимо учесть, что наиболее важная информация хранится и обрабатывается в АССН 4 класса. Следовательно, весовой коэффициент D для него примем за 1, а коэффициенты для остальных классов рассчитаем, используя пропорцию:

4 класс АССН – 1;
5 класс АССН – 0,56;
6 класс АССН – 0,3.

2.2. Формирование базовых моделей оценки опасности деструктивных программных воздействий в общем виде

Базовые модели оценки опасности деструктивных программных воздействий разрабатываются для классов АС на основе исходных моделей и имеет вид:

Без учета связи поведенческих паттернов:

$$J_{V_B} = 10 \cdot \frac{J^*_{V_B}}{J_{B_{\max}}} = \frac{\sum\limits_{i=1}^{n} \psi_i \cdot w_i \cdot p_i}{J_{B_{\max}}}, \quad (4)$$

где J_{V_B} – скорректированное значение показателя опасности для базовой модели без учета связи поведенческих паттернов; ψ – значение весовых коэффициентов базовых признаков актуальности для соответствующих поведенческих паттернов; $J_{B_{\max}}$ – значение опасности самой опасной вредоносной программы для базовой модели без учета связи поведенческих паттернов.

С учетом связи поведенческих паттернов:

$$J_{KV_B} = 10 \cdot \frac{J^*_{KV_B}}{J_{KB_{\max}}} = \frac{\sum\limits_{i=1}^{n} \psi_i \cdot w_i \cdot p_i + \sum\limits_{j=1}^{u} \psi_j \cdot K_j}{J_{KB_{\max}}}, \quad (5)$$

где: $J^*_{KV_B}$ – скорректированное значение показателя опасности для базовой модели с учетом связи поведенческих паттернов; $J_{KB_{\max}}$ – значение опасности самой опасной вредоносной программы для базовой модели с учетом связи поведенческих паттернов.

Для формирования базовых моделей ООДПВ на АС необходимо выделить признаки, которые влияют на актуальность поведенческих паттернов вредоносных программ. Исходя из результатов исследований [15, 16] определены базовые признаки автоматизированных систем. В таблице 3 представлены базовые признаки АС и значения, которые они могут принимать и весовые коэффициенты для каждого значения.

Полученные значения коэффициентов меньше 0,1 будем считать незначительными, и приравнивать к 0.

Каждый из этих признаков влияет на актуальность поведенческих паттернов вредоносных программ. Для каждого класса АС необходимо сформировать мультимножество значений базовых признаков актуальности для поведенческих паттернов.

$$\Psi = \{\psi_1, \psi_2, \ldots, \psi_n, \ldots, \psi_u\}, \quad (6)$$

где $\psi = \frac{\Sigma h^*}{b}$, h^* – весовые коэффициенты значений базовых признаков АС, влияющих на поведенческий

Таблица 3.

Базовые признаки автоматизированных систем

Наименование признака АССН ОВД, h	Принимаемые значения признака	Значение коэффициента, h^*
Технология, используемая для построения системы хранения данных (СХД), h_1	NAS	1
	SAN	0,49
Доступ к сервису электронной почты (СЭП), h_2	Возможна отправка и получение писем внутри организации и в сети общего доступа	1
	Возможность отправки и получения писем только внутри организации	0,44
	Отсутствует доступ к сервису электронной почты	0,08
Подключение к сетям общего доступа (СОД), h_3	Возможность доступа ко всем ресурсам сети общего доступа	1
	Возможность доступа только к разрешенным ресурсам сети общего доступа	0,58
	Отсутствует подключение к сети общего доступа	0,09

Таблица 4.

Характеристики поведенческих паттернов вредоносных утилит

Поведенческий паттерн	Свойство информации	Базовый признак АССН ОВД	Частный признак АССН ОВД
Проникновение на компьютер-жертву, p_1	К	h_1,h_3	f_4,f_5,f_6
Скрытие следов присутствия преступников в системе, p_2	К	h_3	f_4,f_5
Внесение в список разрешенных посетителей системы новых пользователей, p_3	К	h_1,h_3	f_4,f_5,f_6
Прекращение работы системы, p_4	Д	h_1,h_3	f_1,f_3
Проведение атак типа «Отказ в обслуживании», p_5	Д	h_3	f_2,f_3,f_7
Сбор и анализ сетевых пакетов, p_6	К	h_1,h_3	f_1,f_2
Подмена адреса отправителя письма по электронной почте, p_7	Ц	h_2	f_6
Создание вредоносных программ, p_8	Ц	h_3	f_4
Навязывание ложной информации (уведомление об опасности, нарушениях), p_9	Ц	h_1	f_4
Модификация вредоносных программ, p_{10}	Ц	h_3	f_4
Распространение флуда (бесполезных сообщений по каналам электронной почты), p_{11}	Ц	h_2	f_6

паттерн, b – количество влияющих на поведенческий паттерн признаков АС (таблица 4).

2.3. Формирование базовых моделей оценки опасности деструктивных программных воздействий для АССН «Тестовая АССН ОВД»

Рассмотрим пример формирования базовых моделей оценки опасности деструктивных программных воздействий для АССН «Тестовая АССН ОВД» для вредоносных программ класса «Вредоносные утилиты».

Поведенческие паттерны вредоносных утилит влияют на конфиденциальность (К), целостность (Ц) и доступность (Д) информации, обрабатываемой в АССН ОВД. В таблице 4 представлены свойства информации, на которые воздействует паттерн и признак АССН ОВД, от которого зависит актуальность поведенческого паттерна.

В работе [17] разработана исходная модель для оценки опасности деструктивного программного воздействия вредоносных утилит:

$$J_{AHP_V} = 10 \cdot (0{,}258 \cdot p_1 + 0{,}181 \cdot p_2 + 0{,}121 \cdot p_3 + 0{,}121 \cdot p_4 + 0{,}077 \cdot p_5 + 0{,}077 \cdot p_6 + 0{,}077 \cdot p_7 + 0{,}027 \cdot p_8 + 0{,}027 \cdot p_9 + 0{,}019 \cdot p_{10} + 0{,}015 \cdot p_{11}) / 0{,}516. \quad (7)$$

В работе [10] определено множество связанных поведенческих паттернов вредоносных утилит

$$L = \{\{p_4,p_5\};\{p_7,p_{11}\};\{p_8,p_{10}\}\}.$$

С учетом связи поведенческих паттернов модель (7) примет следующий вид:

1. Если во вредоносной утилите совместно реализуется пара паттернов p_4,p_5:

$$J_{KN_{4,5}} = 10 \cdot (0{,}224 \cdot p_1 + 0{,}157 \cdot p_2 + 0{,}105 \cdot p_3 + 0{,}105 \cdot p_4 + 0{,}067 \cdot p_5 + 0{,}067 \cdot p_6 + 0{,}067 \cdot p_7 + 0{,}024 \cdot p_8 + 0{,}024 \cdot p_9 + 0{,}017 \cdot p_{10} + 0{,}013 \cdot p_{11} + 0{,}13 \cdot p_{4,5}) / 0{,}487. \quad (8)$$

2 .Если во вредоносной утилите совместно реализуется пара паттернов p_7,p_{11}:

$$J_{KN_{7,11}} = 10 \cdot (0{,}232 \cdot p_1 + 0{,}163 \cdot p_2 + 0{,}109 \cdot p_3 + 0{,}109 \cdot p_4 + 0{,}07 \cdot p_5 + 0{,}07 \cdot p_6 + 0{,}07 \cdot p_7 + 0{,}024 \cdot p_8 + 0{,}024 \cdot p_9 + 0{,}017 \cdot p_{10} + 0{,}014 \cdot p_{11} + 0{,}1 \cdot p_{7,11}) / 0{,}504. \quad (9)$$

3. Если во вредоносной утилите совместно реализуется пара паттернов p_8,p_{10}:

$$J_{KN_{8,10}} = 10 \cdot (0{,}244 \cdot p_1 + 0{,}171 \cdot p_2 + 0{,}114 \cdot p_3 + 0{,}114 \cdot p_4 + 0{,}073 \cdot p_5 + 0{,}073 \cdot p_6 + 0{,}073 \cdot p_7 + 0{,}026 \cdot p_8 + 0{,}026 \cdot p_9 + 0{,}018 \cdot p_{10} + 0{,}014 \cdot p_{11} + 0{,}056 \cdot p_{8,10}) / 0{,}529. \quad (10)$$

АССН «Тестовая АССН» относится к 4 классу и имеет следующие базовые признаки:

- Класс АССН – 4.
- СХД (h_1) – SAN ($h_1^* = 0{,}49$).
- Доступ к СЭП (h_2) – возможность отправки и получения писем только внутри организации (ведомства) ($h_2^* = 0{,}44$).
- Подключение АС к СОД (h_3) – отсутствует подключение к сети общего доступа ($h_3^* = 0{,}09$).

Составим мультимножество значений базовых признаков актуальности поведенческих паттернов для данного класса АССН ОВД:

$$\Psi_4 = \{0{,}25;0;0{,}25;0{,}25;0;0{,}25;0{,}44;0;0{,}49;0;0{,}44;0{,}44\}. \quad (11)$$

Значения базовых признаков актуальности для коэффициентов K определяются как средние значения базовых признаков для соответствующих связанных поведенческих паттернов.

Выполнив вычисления и нормировку элементов [18] получим следующие базовые модели ООДПВ:

1. При отсутствии связанных признаков:

$$J^*_{V_4} = 0{,}326 \cdot p_1 + 0{,}153 \cdot p_3 + 0{,}153 \cdot p_4 + \\ + 0{,}098 \cdot p_6 + 0{,}172 \cdot p_7 + 0{,}067 \cdot p_9 + 0{,}03 \cdot p_{11}. \quad (12)$$

2. При совместной реализации поведенческих паттернов p_7,p_{11}:

$$J^*_{KV_{4_{7,11}}} = 0{,}26 \cdot p_1 + 0{,}12 \cdot p_3 + 0{,}12 \cdot p_4 + 0{,}079 \cdot p_6 + \\ + 0{,}138 \cdot p_7 + 0{,}053 \cdot p_9 + 0{,}03 \cdot p_{11} + 0{,}2 \cdot K_{7,11}. \quad (13)$$

В данном случае не рассматриваются остальные пары связанных поведенческих паттернов, поскольку они не актуальны для данной АССН ОВД.

Затем необходимо определить значение $J_{4_{\max}}$ исходя из весовых коэффициентов поведенческих паттернов. Самой опасной вредоносной утилитой будет Linux.Siggen.172223, которая реализует поведенческие паттерны p_1, p_4, p_6, p_9. Для систем 4 класса ее опасность равна 0,644. При появлении новых вредоносных утилит, опасность которых будет выше, чем у Linux.Siggen.172223, необходимо будет выполнить уточнение сформированных моделей. Для формулы (10) самой опасной также будет вредоносная утилита Linux.Siggen.172223 и ее опасность равна 0,512.

Таким образом, базовые модели оценки опасности деструктивных программных воздействий вредоносных утилит для 4 класса АССН ОВД будут иметь вид:

$$J_{V_4} = 10 \cdot (0{,}326 \cdot p_1 + 0{,}153 \cdot p_3 + 0{,}153 \cdot p_4 + \\ + 0{,}098 \cdot p_6 + 0{,}172 \cdot p_7 + 0{,}067 \cdot p_9 + \\ + 0{,}03 \cdot p_{11}) / 0{,}644. \quad (14)$$

$$J_{KV_{4_{7,11}}} = 10 \cdot (0{,}26 \cdot p_1 + 0{,}12 \cdot p_3 + 0{,}12 \cdot p_4 + \\ + 0{,}079 \cdot p_6 + 0{,}138 \cdot p_7 + 0{,}053 \cdot p_9 + 0{,}03 \cdot p_{11} + \\ + 0{,}2 \cdot K_{7,11}) / 0{,}512. \quad (15)$$

3. Формирование частных моделей оценки опасности деструктивных программных воздействий

Частные модели предназначены для оценки опасности деструктивных программных воздействий вредоносных программ на конкретную АС.

3.1. Формирование частных моделей оценки опасности деструктивных программных воздействий в общем виде

Частные модели оценки опасности деструктивных программных воздействий разрабатываются для конкретной АС на основе базовых моделей и имеют вид:

Без учета связи поведенческих паттернов:

$$J_{V_B} = 10 \cdot \frac{J^*_{V_{Par}}}{J_{Par_{\max}}} = \frac{\sum_{i=1}^{n} \omega_i \cdot w_{ci} \cdot p_i}{J_{Par_{\max}}}, \quad (16)$$

где $J^*_{V_{Par}}$ – скорректированное значение показателя опасности для частной модели без учета связи признаков; ω – значение весовых коэффициентов частных признаков актуальности для соответствующих поведенческих паттернов; w_{ci} – скорректированное значение весовых коэффициентов поведенческих паттернов, рассчитанное в базовой модели по формуле: $w_{ci} = \psi_i \cdot w_i$; $J_{Par_{\max}}$ – значение опасности самой опасной вредоносной программы для частной модели без учета связи поведенческих паттернов.

С учетом связи поведенческих паттернов:

$$J_{KV_{Par}} = 10 \cdot \frac{J^*_{KV_{Par}}}{J_{KPar_{\max}}} = \frac{\sum_{i=1}^{n} \omega_i \cdot w_{ci} \cdot p_i + \sum_{j=1}^{u} \omega_j \cdot K_{cj}}{J_{KPar_{\max}}}, \quad (17)$$

где K_{cj} – скорректированное значение коэффициента для учета связи признаков, рассчитанное в базовой модели по формуле $K_{cj} = \psi_j \cdot K_j$; $J^*_{KV_{Par}}$ – скорректированное значение показателя опасности для частной модели без учета связи признаков; $J_{KPar_{\max}}$ – значение опасности самой опасной вредоносной программы для частной модели без учета связи поведенческих паттернов.

Для формирования частных моделей экспертам необходимо оценить влияние частных признаков АС на актуальность поведенческих паттернов. В работах [19, 20] определены признаки автоматизированных систем, которые рекомендуется рассматривать при моделировании процессов, связанных с информационной безопасностью: Состав и принимаемые значения для данных признаков могут изменяться исходя из особенностей построения и функционирования автоматизированных систем.

В результате работы экспертной группы для конкретной АС будет сформировано мультимножество значений частных признаков актуальности для поведенческих паттернов:

$$\Omega = \{\omega_1,\omega_2,...,\omega_n,...,\omega_u\}, \quad (18)$$

где $\omega = \frac{\Sigma f^*}{r}$, f^* – весовые коэффициенты значений базовых признаков АС, влияющих на поведенческий паттерн; r – количество влияющих на поведенческий паттерн признаков АС.

3.2. Формирование частных моделей оценки опасности деструктивных программных воздействий для АССН «Тестовая АССН ОВД»

Частные признаки АССН ОВД и принимаемые ими значения, с соответствующими весовыми коэффициентами (рассчитанными с использованием метода анализа иерархий) представлены в таблице 5. Весовые коэффициентов частных признаков могут быть уточнены в процессе формирования моделей, при изменении признаков, или принимаемых ими значений.

В нашем случае экспертная группа будет оценивать признаки АССН «Тестовая АССН ОВД», которая относится к 4 классу и имеет следующие частные признаки:

Назначение автоматизированной системы (f_1) – обеспечение оперативно-служебной деятельности $(f_1^* = 1)$.

Используемая топология построения автоматизированной системы (f_2) – древовидная структура $(f_2^* = 0{,}49)$.

Задачи, решаемые интеграцией с внешними автоматизированными системами (f_3) – отсутствует интеграция с внешними системами $(f_3^* = 0)$.

Тип используемой операционной системы (f_4) – ОС семейства Linux в защищенном исполнении $(f_4^* = 0{,}19)$.

Таблица 5.

Частные признаки АССН ОВД

Наименование признака АССН ОВД, , f	**Принимаемые значения признака**	**Значение коэффициента, f^***
Назначение автоматизированной системы, f_1	Обеспечение оперативно-служебной деятельности	1
	Обеспечение повседневной деятельности	0,57
	Обеспечение других видов деятельности	0,21
Используемая топология построения автоматизированной системы, f_2	Звезда-шина	1
	Древовидная структура	0,49
Задачи, решаемые интеграцией с внешними автоматизированными системами, f_3	Обмен файлами	1
	Общая база данных	0,44
	Отсутствует интеграция с внешними системами	0,08
Тип используемой операционной системы, f_4	ОС семейства Windows	1
	ОС семейства Linux	0,53
	ОС семейства Linux в защищенном исполнении	0,19
Тип используемой системы управления базами данных, f_5	Oracle	1
	MySQL	1
	PostgreSQL	0,33
Количество доменов безопасности в автоматизированной системе, f_6	Применение одного домена безопасности в нескольких АССН	1
	Отдельный домен безопасности для каждой АССН	0,49
Процент задействования ресурсов (оперативная, постоянная память), автоматизированной системы, в моменты пиковой нагрузки, f_7	[75% – 100%]	1
	[50% – 75%)	0,41
	[0% – 50%)	0,16

Тип используемой системы управления базами данных (f_5) – PostgreSQL ($f_5^* = 0{,}33$).

Количество доменов безопасности в автоматизированной системе (f_6) – отдельный домен безопасности для каждой АССН ($f_6^* = 0{,}49$).

Процент задействования ресурсов (оперативная, постоянная память), автоматизированной системы, задействованной в моменты пиковой нагрузки (f_7) – [50 % – 75 %] ($f_7^* = 0{,}41$).

Таким образом, мультимножество значений частных признаков актуальности поведенческих паттернов для АССН «Тестовая АССН ОВД» будет иметь вид:

$$\Omega_{Par} = \{0{,}34;0{,}26;0{,}34;0{,}5;0{,}3;0{,}75; 0{,}49;0{,}19;0{,}19;0{,}19;0{,}49;0{,}49\}. \quad (19)$$

Выполнив вычисления и нормировку элементов множества путем деления на их сумму, получим следующие частные модели оценки опасности деструктивных программных воздействий:

1. При отсутствии связанных поведенческих паттернов:

$$J^*_{V_{Par}} = 0{,}261 \cdot p_1 + 0{,}123 \cdot p_3 + 0{,}18 \cdot p_4 + 0{,}173 \cdot p_6 + 0{,}199 \cdot p_7 + 0{,}03 \cdot p_9 + 0{,}034 \cdot p_{11}. \quad (20)$$

2. При совместной реализации поведенческих паттернов p_7, p_{11}:

$$J^*_{KV_{Par_{7,11}}} = 0{,}21 \cdot p_1 + 0{,}09 \cdot p_3 + 0{,}14 \cdot p_4 + 0{,}14 \cdot p_6 + 0{,}15 \cdot p_7 + 0{,}02 \cdot p_9 + 0{,}03 \cdot p_{11} + 0{,}22 \cdot K_{7,11}. \quad (21)$$

Самой опасной вредоносной утилитой для автоматизированной системы «Тестовая АССН ОВД» также будет являться Linux.Siggen.172223, и ее опасность $J_{Par_{\max}}$ равна 0,647, а $J^*_{KPar_{7,11_{\max}}}$ равна 0,51. Также

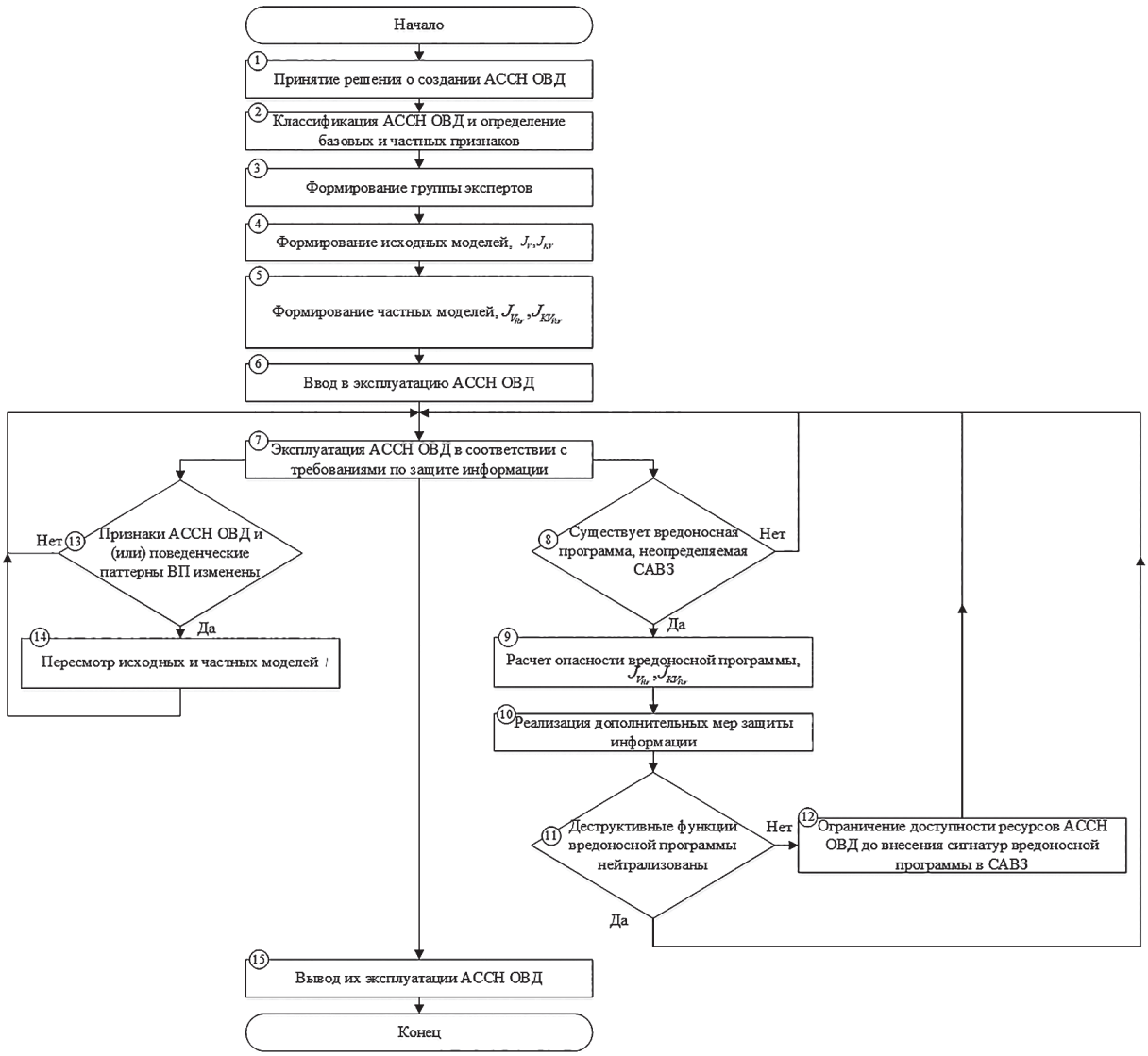


Рис. 2. Алгоритм планирования и реализации процессов жизненного цикла АССН ОВД в условиях деструктивных программных воздействий

необходимо учесть коэффициенты для корректировки частной модели в зависимости от класса (уровня) защищенности.

Автоматизированная система «Тестовая АССН ОВД» относится к 4 классу АССН, следовательно, частные модели будет иметь вид:

1. Без учета связи поведенческих паттернов:

$$J_{KV_{Par}} = 1 \cdot (10 \cdot (0{,}261 \cdot p_1 + 0{,}123 \cdot p_3 + 0{,}18 \cdot p_4 + 0{,}173 \cdot p_6 + 0{,}199 \cdot p_7 + 0{,}03 \cdot p_9 + 0{,}034 \cdot p_{11}) / 0{,}647). \quad (22)$$

2. С учетом связи поведенческих паттернов p_7,p_{11}:

$$J_{KV_{Par_{7,11}}} = 1 \cdot (10 \cdot (0{,}21 \cdot p_1 + 0{,}09 \cdot p_3 + 0{,}14 \cdot p_4 + 0{,}14 \cdot p_6 + 0{,}15 \cdot p_7 + 0{,}02 \cdot p_9 + 0{,}03 \cdot p_{11} + 0{,}22 \cdot K_{7,11}) / 0{,}51). \quad (23)$$

Таким образом, алгоритм планирования и реализации процессов жизненного цикла АССН ОВД в условиях деструктивных программных воздействий представлен на рисунке 2.

4. Верификация сформированных моделей оценки опасности деструктивных программных воздействий

Верификация сформированных моделей выполнена на тестовом наборе данных вредоносных утилит и представлена в таблице 6.

В ходе вычислительного эксперимента определено, что учет связи поведенческих паттернов вредоносных программ и признаков автоматизированных систем влияет на уровень опасности вредоносных утилит. Например, для вредоносной утилиты № 1 Constructor.DarkHorse в исходной модели определен уровень опасности «Критический», а для базовой и частной «Средний». Во вредоносной утилите № 3 DDoS.Siggen.41 в исходной модели определен уровень опасности «Средний», а в базовой и частной «Низкий». Снижение уровня опасности вызвано тем, что реализуемые в данных примерах связанные поведенческие паттерны не актуальны для АССН «Тестовая АССН ОВД», а признаки АССН снижают актуальность для остальных реализуемых паттернов. Для вредоносной утилиты № 5 Tool.TermService уровень опасности, рассчитанный с использованием исходной модели «Средний», а с использованием базовой и частной «Критический». Отличие в значениях показателя опасности вызвано тем, что признаки АССН повышают актуальность реализуемых во вредоносной программе поведенческих паттернов. Формирование частных моделей для каждой АССН ОВД позволит принимать адекватные и достаточные меры по обеспечению безопасности информации при появлении неизвестных вредоносных программ.

Заключение

Разработанная методика ООДПВ, отличается от существующих учетом признаков АССН ОВД, способствует формированию единого подхода к реализации мер по защиты информации. При разработке методики были решены следующие задачи:

1. Определены базовые и частные признаки АССН ОВД, характеризующие актуальность поведенческих паттернов вредоносных программ в зависимости от функциональных особенностей АССН ОВД.

Таблица 6.

Верификация сформированных моделей

№ п/п	Название вредоносной утилиты	Реализуемые поведенческие паттерны	Исходные модели J_{AHP_V},J_{KN}	Базовые модели J_{V_4},J_{KV_4}	Частные модели $J_{V_{Par}}$,$J_{KV_{Par}}$
1.	Constructor.DarkHorse	p_1, p_2, p_8, p_{10}	9,74	5,06	4,03
2.	Spy-Net 0.9	p_1, p_2	8,51	5,06	4,03
3.	DDoS.Siggen.41	p_4, p_5, p_{10}	6,2	2,38	2,78
4.	Linux.Siggen.5542	p_1, p_6	6,49	6,58	6,7
5.	Tool.TermService	p_3, p_7, p_{11}	5,8	9,53	9,61
6.	Linux.Siggen.322	p_1	5	5,06	4,03
7.	Tool.UDPFlood	p_3, p_{11}	2,64	2,84	2,43
8.	Tool.InstallToolbar.5	p_6, p_9	2,02	2,56	3,13
9.	Tool.Wpakill.4	p_7, p_9	2,02	3,71	3,6
10.	Tool.Spamer.18	p_9, p_{11}	0,81	1,09	0,99

2. Выполнено моделирование оценки опасности деструктивных программных воздействий с учетом базовых, частных признаков АССН ОВД.
3. Разработан алгоритм планирования и реализации процессов жизненного цикла АССН ОВД в условиях деструктивных программных воздействий.
4. Выполнен вычислительный эксперимент, по оценке опасности деструктивных программных воздействий для АССН ОВД. В ходе вычислительного эксперимента получены непротиворечивые результаты, которые подтверждают зависимость уровня опасности вредоносных программ от признаков АССН ОВД.

Результаты исследования могут быть применены администраторами безопасности АССН ОВД, для обеспечения бесперебойного функционирования на всех этапах эксплуатации систем.

Перспективы дальнейших исследований:

- определение набора достаточных дополнительных мер защиты информации, при появлении неизвестных вредоносных программ;
- формирование моделей для оценки опасности деструктивных программных воздействий с учетом актуальности поведенческих паттернов для других классов вредоносных программ.

Литература

1. F. Alkhudhayr, S. Alfarraj, B. Aljameeli and S. Elkhdiri, «Information Security: A Review of Information Security Issues and Techniques», 2019. 2nd International Conference on Computer Applications & Information Security (ICCAIS), Riyadh, Saudi Arabia, 2019, pp. 1–6. DOI:10.1109/CAIS.2019.8769504.
2. Methodical approach to reducing the dimensionality of the task of requirements substantiation for protection of information systems against unauthorized access in organizational-technical systems / T. V. Meshcheryakova, A. V. Batskikh, O. A. Gulyaev, A. A. Abdullin // Journal of Physics: Conference Series: Applied Mathematics, Computational Science and Mechanics: Current Problems, Voronezh, 11–13 ноября 2019 года. – Bristol: Institute of Physics Publishing, 2020. – P. 012013. – DOI 10.1088/1742-6596/1479/1/012013.
3. Оценка соответствия модели угроз и требований доверия систем Интернета вещей массового применения / А. А. Бахтин, Д. С. Брагин, А. А. Конев, А. В. Шарамок // Наноиндустрия. – 2020. – Т. 13, № S4(99). – С. 137–138. – DOI 10.22184/1993-8578.2020.13.4s.137.138.
4. Язов Ю. К. Составные сети Петри-Маркова со специальными условиями построения для моделирования угроз безопасности информации / Ю. К. Язов, А. П. Панфилов // Вопросы кибербезопасности. – 2024. – № 2(60). – С. 53–65. – DOI 10.21681/2311-3456-2024-2-53-65.
5. Язов Ю. К. Составные сети Петри – Маркова на основе полумарковских процессов и их применение при моделировании динамики реализации угроз безопасности информации в информационных системах / Ю. К. Язов, А. О. Авсентьев, А. П. Панфилов, В. Н. Пржегорлинский // Вестник Воронежского института МВД России. – 2024. – № 2. – С. 63–78. – EDN UWINDW.
6. Мещеряков Р. В. Перспективные направления применения технологий искусственного интеллекта при защите информации // Мещеряков Р. В., Мельников С. Ю., Пересыпкин В. А., Хорев А. А. // Вопросы кибербезопасности. – 2024. – № 4(62). – С. 2–12. – DOI 10.21681/2311-3456-2024-4-02-12.
7. Models and methods of information reliability and data protection / G. I. Korshunov, V. A. Lipatnikov, V. A. Tichonov [et al.] // IOP Conference Series: Materials Science and Engineering: International Workshop «Advanced Technologies in Material Science, Mechanical and Automation Engineering – MIP: Engineering – 2019», Krasnoyarsk, 04–06 апреля 2019 года. – London: Institute of Physics and IOP P8ublishing Limited, 2019. – P. 52001. – DOI 10.1088/1757-899X/537/5/052001.
8. Атакищев О. И. Метаграмматический подход анализа иерархий для синтеза систем безопасности атомных электростанций / О. И. Атакищев, В. Г. Грибунин, И. Л. Борисенков, М. Н. Лысачев // Вопросы кибербезопасности. – 2023. – № 1(53). – С. 82–92. – DOI 10.21681/2311-3456-2023-1-82-92.
9. Munier N. Uses and Limitations of the AHP Method/ N. Munier, E. Hontoria // Management for Professionals. Springer Cham – 2021. – 130 pp. DOI 10.1007/978-3-030-60392-2.
10. Мельников А. В. Численный метод модификации моделей, разработанных на основе метода анализа иерархий, с использованием искусственной нейронной сети / А. В. Мельников, Н. С. Кобяков // Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. – 2024. – № 4. – С. 5–22. – DOI 10.17308/sait/1995-5499/2024/4/5-21.
11. Мельников А. В. Подход к оценке опасности деструктивных воздействий вредоносных программ на автоматизированные системы специального назначения / А. В. Мельников, Н. С. Кобяков // Безопасность информационных технологий. – 2023. – Т. 30, № 3. – С. 51–60. – DOI 10.26583/bit.2023.3.03.
12. Мельников А. В. Модели и алгоритмы реализации организационных мер защиты информации в АССН от деструктивных воздействий ранее неизвестных вредоносных программ / А. В. Мельников, Н. С. Кобяков, Р. А. Жилин // Вестник Воронежского института МВД России. – 2023. – № 3. – С. 80–87. – EDN ZILKNA.
13. Кобяков, Н. С. Алгоритм классификации автоматизированных систем специального назначения / Н. С. Кобяков, В. Н. Париев // Альманах Пермского военного института войск национальной гвардии. – 2024. – № 2(14). – С. 15–21. – EDN PKWCCP.
14. Жилин Р. А. Численный метод предварительной экспертизы альтернатив нарушителей охраны объектов общекриминальной направленности / Р. А. Жилин, А. В. Мельников, И. В. Щербакова // Вестник Воронежского института МВД России. – 2019. – № 3. – С. 46–54. – EDN NEYIJN.
15. Авраменко В. С., Маликов А. В. Методика диагностирования компьютерных инцидентов безопасности в автоматизированных системах специального назначения. Наукоемкие технологии в космических исследованиях Земли. – 2020. – Т. 12, № 1. – С. 44–52. – DOI 10.36724/2409-5419-2020-12-1-44-52.
16. Долгачев, М. В., Костюнин В. А. Комплексный анализ поведения системы Windows для обнаружения киберугроз. Вопросы кибербезопасности. – 2025. – № 2(66). – С. 71–77. – DOI 10.21681/2311-3456-2025-2-71-77.
17. Мельников А. В. Модель оценки опасности вредоносных утилит / А. В. Мельников, В. И. Сумин, Н. С. Кобяков // Промышленные АСУ и контроллеры. – 2023. – № 7. – С. 33–40. – DOI 10.25791/asu.7.2023.1448.

18. Melnikov, A. V. Method of forming expert coalitions in the context of solving the expertise problem of alternatives with weakly formalized criteria / A. V. Melnikov, I. V. Shcherbakova, R. A. Zhilin // Journal of Physics: Conference Series: Applied Mathematics, Computational Science and Mechanics: Current Problems, Voronezh, 11–13 ноября 2019 года. – Bristol: Institute of Physics Publishing, 2020. – P. 012071. – DOI 10.1088/1742-6596/1479/1/012071.
19. Язов Ю. К., Соловьев С. В. Методология оценки эффективности защиты информации в информационных системах от несанкционированного доступа. Санкт-Петербург: Издательство «Наукоемкие технологии», 2023. – 258 с. – ISBN 978-5-907618-36-7. – EDN WVCHKW.
20. Язов Ю. К., Авсентьев О. С., Авсентьев А. О., Рубцова И. О. Метод оценивания эффективности защиты электронного документооборота с применением аппарата сетей Петри – Маркова. Труды СПИИРАН. – 2019. – Т. 18, № 6. – С. 1269–1300. – DOI 10.15622/sp.2019.18.6.1269-1300.

METHOD OF ASSESSING THE DANGER OF DESTRUCTIVE SOFTWARE IMPACTS ON AUTOMATED SPECIAL-PURPOSE SYSTEMS OF INTERNAL AFFAIRS BODIES

Melnikov A. V.[4], Kobyakov N. S.[5]

Keywords: *malware, automated systems features, information security, analytic hierarchy process.*

The objective of the study: *modeling the hazard indicator of destructive software impacts, taking into account the relevance of the behavioral patterns of malware for automated special-purpose systems of the internal affairs agencies.*

Research methods: *the hierarchy analysis method is used to form models for assessing the hazard of destructive software impacts and to determine the numerical values of the attributes of the automated special-purpose systems of the internal affairs agencies.*

Research result: *the basic and specific attributes of the automated special-purpose systems of the internal affairs agencies are determined, characterizing the relevance of the behavioral patterns of malware depending on the functional features of the automated special-purpose systems of the internal affairs agencies. Basic and specific models for assessing the hazard of destructive software impacts on the automated special-purpose systems of the internal affairs agencies have been developed, taking into account the relevance of the behavioral patterns of malware. An algorithm for planning and implementing the life cycle processes of the automated special-purpose systems of the internal affairs agencies in the context of destructive software impacts has been developed. The developed methodology has been verified using the example of forming models for assessing the hazard of destructive software impacts of malware of the «Malicious Utilities» class on a test automated special-purpose system. The developed models have been verified on a test data set generated by interviewing experts.*

Practical significance: *the developed methodology can be used by security administrators of automated special-purpose systems when assessing the danger of destructive software impacts and determining the goals and list of measures to be implemented to ensure information protection when unknown malicious programs appear.*

References

1. F. Alkhudhayr, S. Alfarraj, B. Aljameeli and S. Elkhdiri, «Information Security:A Review of Information Security Issues and Techniques», 2019. 2nd International Conference on Computer Applications & Information Security (ICCAIS), Riyadh, Saudi Arabia, 2019, pp. 1–6, doi: 10.1109/CAIS.2019.8769504.
2. Methodical approach to reducing the dimensionality of the task of requirements substantiation for protection of information systems against unauthorized access in organizational-technical systems / T. V. Meshcheryakova, A. V. Batskikh, O. A. Gulyaev, A. A. Abdullin // Journal of Physics: Conference Series: Applied Mathematics, Computational Science and Mechanics: Current Problems, Voronezh, 11–13 ноября 2019 года. – Bristol: Institute of Physics Publishing, 2020. – P. 012013. – DOI 10.1088/1742-6596/1479/1/012013.
3. Ocenka sootvetstviya modeli ugroz i trebovanij doveriya sistem Interneta veshchej massovogo primeneniya / A. A. Bakhtin, D. S. Bragin, A. A. Konev, A. V. Sharamok // Nanoindustry. – 2020. – T. 13, № S4(99). – pp. 137-138. – DOI 10.22184/1993-8578.2020.13.4s.137.138.
4. Yazov, Yu. K. Sostavnye seti Petri-Markova so special'nymi usloviyami postroeniya dlya modelirovaniya ugroz bezopasnosti informacii / Yu. K. Yazov, A. P. Panfilov // Cybersecurity issues. – 2024. – № 2(60). – pp. 53-65. – DOI 10.21681/2311-3456-2024-2-53-65.
5. Sostavnye seti Petri – Markova na osnove polumarkovskih processov i ih primenenie pri modelirovanii dinamiki realizacii ugroz bezopasnosti informacii v informacionnyh sistemah / Yu. K. Yazov, A. O. Avsentiev, A. P. Panfilov, V. N. Przhegorlinsky // Vestnik Voronezhskogo instituta MVD Rossii. – 2024. – № 2. – pp. 63-78. – EDN UWINDW.
6. Perspektivnye napravleniya primeneniya tekhnologij iskusstvennogo intellekta pri zashchite informacii / R. V. Meshcheryakov, S. Yu. Melnikov, V. A. Peresypkin, A. A. Khorev // Cybersecurity issues. – 2024. – № 4(62). – pp. 2–12. – DOI 10.21681/2311-3456-2024-4-02-12.

4 Alexander V. Melnikov, Doctor of Technical Sciences, Associate Professor, Professor of the Department of Automated Information Systems of Internal Affairs Bodies, Voronezh Institute of the Ministry of Internal Affairs of the Russian Federation, Voronezh, Russia. ORCID: https://orcid.org/0000-0001-5080-1162. E-mail: meln78@mail.ru
5 Nikolai S. Kobyakov, postgraduate student of the Department of Automated Information Systems of Internal Affairs Bodies, Voronezh Institute of the Ministry of Internal Affairs of the Russian Federation, Voronezh, Russia. ORCID: https://orcid.org/0000-0002-4950-7879. E-mail: kkobyakov1234@gmail.com

7. Models and methods of information reliability and data protection / G. I. Korshunov, V. A. Lipatnikov, V. A. Tichonov [et al.] // IOP Conference Series: Materials Science and Engineering: International Workshop «Advanced Technologies in Material Science, Mechanical and Automation Engineering – MIP: Engineering – 2019», Krasnoyarsk – London: Institute of Physics and IOP P8ublishing Limited, 2019. – P. 52001. – DOI 10.1088/1757-899X/537/5/052001.
8. Metagrammaticheskij podhod analiza ierarhij dlya sinteza sistem bezopasnosti atomnyh elektrostancij / O. I. Atakishchev, V. G. Gribunin, I. L. Borisenkov, M. N. Lysachev // Cybersecurity issues. – 2023. – № 1(53). – pp. 82–92. – DOI 10.21681/2311-3456-2023-1-82-92.
9. Munier N. Uses and Limitations of the AHP Method/ N. Munier, E. Hontoria // Management for Professionals. Springer Cham – 2021. – 130 pp. DOI 10.1007/978-3-030-60392-2.
10. Mel'nikov A. V. Kobjakov N. S. Chislennyj metod modifikacii modelej, razrabotannyh na osnove metoda analiza ierarhij, s ispol'zovaniem iskusstvennoj nejronnoj seti.VSU Bulletin. Series: System analysis and information technologies – 2024. – № 4. – S. 5–22. – DOI 10.17308/sait/1995-5499/2024/4/5-21.
11. Melnikov A. V. Podhod k ocenke opasnosti destruktivnyh vozdejstvij vredonosnyh programm na avtomatizirovannye sistemy special'nogo naznacheniya / A. V. Melnikov, N. S. Kobyakov // Bezopasnost' informacionnyh tekhnologij. – 2023. – T. 30, № 3. – pp. 51–60. – DOI 10.26583/bit.2023.3.03. – EDN RJWWZH.
12. Melnikov, A. V. Modeli i algoritmy realizacii organizacionnyh mer zashchity informacii v ASSN ot destruktivnyh vozdejstvij ranee neizvestnyh vredonosnyh programm / A. V. Melnikov, N. S. Kobyakov, R. A. Zhilin // Vestnik Voronezhskogo instituta MVD Rossii. – 2023. – № 3. – pp. 80–87. – EDN ZILKNA.
13. Zhilin R. A. CHislennyj metod predvaritel'noj ekspertizy al'ternativ narushitelej ohrany ob"ektov obshchekriminal'noj napravlennosti / R. A. Zhilin, A. V. Melnikov, I. V. Shcherbakova // Vestnik Voronezhskogo instituta MVD Rossii. – 2019. – № 3. – pp. 46–54. – EDN NEYIJN
14. Kobyakov, N. S. Algoritm klassifikacii avtomatizirovannyh sistem special'nogo naznacheniya / N. S. Kobyakov, V. N. Pariev // Al'manah Permskogo voennogo instituta vojsk nacional'noj gvardii. – 2024. – № 2(14). – pp. 15–21. – EDN PKWCCP.
15. Avramenko V. S., Malikov A. V. Metodika diagnostirovanija komp'juternyh incidentov bezopasnosti v avtomatizirovannyh sistemah special'nogo naznachenija. Naukoemkie tehnologii v kosmicheskih issledovanijah Zemli. – 2020. – T. 12, № 1. – S. 44–52. – DOI 10.36724/2409-5419-2020-12-1-44-52.
16. Dolgachev, M. V., Kostjunin V. A. Kompleksnyj analiz povedenija sistemy Windows dlja obnaruzhenija kiberugroz. Voprosy kiberbezopasnosti. – 2025. – № 2(66). – S. 71–77. – DOI 10.21681/2311-3456-2025-2-71-77.
17. Melnikov, A. V. Model' ocenki opasnosti vredonosnyh utilit / A. V. Melnikov, V. I. Sumin, N. S. Kobyakov // Promyshlennye ASU i kontrollery. – 2023. – № 7. – pp. 33–40. – DOI 10.25791/asu.7.2023.1448.
18. Melnikov, A. V. Method of forming expert coalitions in the context of solving the expertise problem of alternatives with weakly formalized criteria / A. V. Melnikov, I. V. Shcherbakova, R. A. Zhilin // Journal of Physics: Conference Series: Applied Mathematics, Computational Science and Mechanics: Current Problems, Voronezh, 11–13 ноября 2019 года. – Bristol: Institute of Physics Publishing, 2020. – P. 012071. – DOI 10.1088/1742-6596/1479/1/012071.
19. Yazov, Yu. K. Soloviev S.V. Metodologiya ocenki effektivnosti zashchity informacii v informacionnyh sistemah ot nesankcionirovannogo dostupa. Sankt-Peterburg: Izdatel'stvo «Naukoemkie tekhnologii» 2023. – 258 P. – ISBN 978-5-907618-36-7. – EDN WVCHKW.
20. Yazov Yu. K., Avsentiev O. S., Avsentiev A. O., Rubtsova I. O. Metod ocenivaniya effektivnosti zashchity elektronnogo dokumentooborota s primeneniem apparata setej Petri – Markova / Proceedings of SPIIRAS. – 2019. – T. 18, № 6. – pp. 1269–1300. – DOI 10.15622/sp.2019.18.6.1269-1300.

О ПРОГНОЗИРОВАНИИ УСТОЙЧИВОСТИ ФУНКЦИОНИРОВАНИЯ СИСТЕМЫ БЕЗОПАСНОСТИ ОБЪЕКТОВ КРИТИЧЕСКОЙ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЫ, ПОДВЕРЖЕННОЙ ВОЗДЕЙСТВИЮ УГРОЗ

Воеводин В. А.[1]

DOI: 10.21681/2311-3456-2025-5-41-49

Цель исследования: *обосновать актуальность, сформулировать и формализовать научную задачу прогнозирования устойчивости функционирования системы безопасности объектов критической информационной инфраструктуры (КИИ), подверженных воздействию угроз.*

Методы исследования: *эвристические, экстраполяционные, экспертные, сравнение и сопоставление, дифференциальное исчисление, информационная диагностика.*

Полученные результаты: *сформулированы вербальная и формальная постановки научной задачи прогнозирования устойчивости функционирования системы безопасности объектов КИИ, подверженных воздействию угроз, предложен алгоритм ее решения задач.*

Научная новизна: *предлагается инструмент для прогнозирования, основанный на построении прогностической модели, представляющей собой линейную комбинацию одноименных параметров объекта прогнозирования и объектов-аналогов. На основе данных о значении параметров объектов-аналогов, когда одноименные параметры имеют линейную корреляцию, делается прогноз. Например, можно априори предсказать устойчивость функционирования системы безопасности объекта критической инфраструктуры после планируемого реинжиниринга его системы безопасности.*

Практическая значимость: *постановка научной задачи может служить основой для формулирования технического задания на реинжиниринг систем безопасности объектов КИИ с заданными требованиями к устойчивости их функционирования.*

Ключевые слова: *прогнозирование, устойчивость функционирования, объект-аналог, реинжиниринг, объект критической информационной инфраструктуры, угрозы.*

Введение

Для поддержания требуемого уровня кибербезопасности и устойчивости функционирования (устойчивость) объектов критической информационной инфраструктуры (КИИ) в соответствии со складывающейся обстановкой осуществляется реинжиниринг их систем безопасности. Системы безопасности (СБ) создаются субъектами КИИ для обеспечения информационной безопасности (защиты информации)[2].

Органам управления информационной безопасностью (ИБ) при принятии решения по реинжинирингу СБ объектов КИИ на ранних стадиях разработки проекта, требуется инструмент для прогнозирования, в том числе, устойчивости функционирования СБ. Прогнозирование устойчивости на ранних этапах разработки сложных и уникальных систем осуществляется в условиях высокой неопределенности, когда нет точных сведений о будущем составе и структуре объекта. На практике могут быть известны: 1) требуемые значения отдельных параметров СБ (требования заказчика), подвергаемой реинжинирингу; 2) устойчивости функционирования одного или нескольких реальных объектов-аналогов, которые могут быть признаны экспертами как прототипы рассматриваемой СБ. В соответствии с этими предположениями задача прогнозирования имеет следующую вербальную постановку.

Постановка задачи

Пусть требуется априори оценить устойчивости функционирования объекта КИИ после ее реинжиниринга его СБ. Для чего было принято решение организовать и провести внутренний аудит с целью прогнозирования устойчивости функционирования объекта КИИ после реинжиниринга СБ. Для организации аудита имеются исходные данные, характеризующие: 1) устойчивости функционирования реальных аналогичных объектов КИИ, находящихся в сравнимых

1 Воеводин Владислав Александрович, кандидат технических наук, доцент МИЭТ, Москва, Россия. AuthorID: 1012813, ORCID 0009-0003-9431-1685. E-mail: vva541@mail.ru

2 Требования к созданию систем безопасности значимых объектов критической информационной инфраструктуры Российской Федерации и обеспечению их функционирования. Утверждены приказом ФСТЭК России от 21 декабря 2017 г. № 235.

условиях в отношении воздействия угроз; 2) требования к устойчивости функционирования оцениваемого объекта, подвергаемого реинжинирингу.

Требуется разработать методику, позволяющую получить прогноз устойчивости функционирования объекта КИИ с учетом возможного воздействия угроз. Определения понятия «кибербезопасность» приводится в [1], структурной и функциональной устойчивости – в [2]. Далее по тексту объект КИИ будет позиционироваться как объект информатизации (ОИ).

Пусть заданы:

1. Исходные условия (данные):

а) требования к устойчивости функционирования, заданные с помощью нестационарного коэффициента живучести $K_{ж}(t)$ ОИ, который характеризует вероятность того, что в момент времени $t \in (0, T]$ он будет находиться в состоянии «функционален», при условии что в начальный момент времени t_0 он находился в состоянии «функционален».

б) требования к обеспечению устойчивости функционирования оцениваемого ОИ $K_{ж}(t) \geq K_{ж}^*$, $t \in 0, T$, где $K_{ж}^*$ – требуемое пороговое значение нестационарного коэффициента живучести;

в) значения параметров, оцениваемого ОИ, существенных для обеспечения требуемой устойчивости его функционирования $X^* = \{x_i^*\}$, $i = \overline{(1, n)}$, где n – число учитываемых параметров;

г) устойчивость функционирования *реальных* ОИ – объектов-аналогов $U = \{U_l\}$, $l = \overline{(1, L)}$, где L – число объектов-аналогов;

д) значения аналогичных параметров для объектов-аналогов $X_l = \{x_{li}\}$, $i = \overline{(1, n)}$, $l = \overline{(1, L)}$;

е) из множества потенциальных объектов-аналогов L экспертным методом выбирается один, который обладает наиболее близким сходством по отношению к оцениваемому, поэтому при дальнейших рассуждениях, если это не изменяет их смысл, индекс l будет опускаться.

2. Исходные гипотезы:

1. В общем виде прогноз устойчивости оцениваемого ОИ U^* (позиционируется как зависимый или искомый параметр) может быть получен с помощью некоторой обобщенной функции

$$U^* = U(g_1, g_2, ..., g_i, ..., g_n), \quad (1)$$

где $g_i = \frac{x_i^*}{x_i}$, $i = \overline{(1, n)}$, g_i – нормирующий коэффициент соответствующих одноименных параметров, который вводится для приведения сравниваемых параметров к одной относительной шкале прогнозирования.

2. При прогнозировании устойчивости необходимо учитывать, что задача решается в условиях:

а) высокой информационной неопределенности, при которой нет точных сведений о сценарии воздействия угроз, вероятностных характеристиках $K_{ж}(t)$;

б) формализованный вид зависимости (1) для сложных объектов вывести неоправданно сложно или практически невозможно;

в) при этом имеется некоторая информация по реальным объектам-аналогам $X_l = \{x_{li}\}$ и $U = \{U_l\}$, см. исходные данные.

В этих условиях конструктивным является подход поиска приближенной функциональной зависимости для выражения (1), построенной на основе процедуры сравнения с аналогами.

3. Исходные допущения:

а) все нормирующие коэффициенты (параметры) $g_{li} \geq 0$ и независимы;

б) $g_{i,j}$ предпочтительно максимизируют;

в) исходная функция (1): неотрицательна и монотонно возрастающая по всем аргументам;

г) $U(r, g_i) = \gamma(r)U(g_i), i = \overline{(1, n)}$, при одновременном увеличении значений всех параметров в r раз устойчивость возрастает в $\gamma(r)$ раз, где $\gamma(r)$ – некоторая дифференцируемая возрастающая функция $r \geq 1$, принимающая при $r = 1$ значение $\gamma(r) = 1$.

Обосновать принятие перечисленных исходных допущений на практике можно с учетом следующих обстоятельств:

1. Условия а) и б) могут быть выполнены соответствующим выбором сравниваемых параметров объекта-аналога. При этом в случае минимизируемых параметров можно брать их обратные значения.
2. Условия в) и г) могут быть приняты на основании логики вещей, так как увеличение (улучшение) значений параметров на практике всегда связано с дополнительными затратами, зависящими от степени их улучшения.

4. Требуется сформировать прогноз устойчивости U^* оцениваемого ОИ, подверженного воздействию угроз. Известные способы оценивания устойчивости ориентированы на условия штатного применения, для которых возможно получить репрезентативную статистику, характеризующую надежность элементов ОИ. Для условий воздействия угроз, когда репрезентативная статистика в принципе отсутствует, а сами процессы не являются стационарными такие методы могут приводить к ошибочным оценкам.

Ретроспективный анализ методов решения подобных задач

В настоящее время опубликовано значительное число работ, посвященных оцениванию результативности эксплуатации сложных технических систем[3],

3 Саркисян С. А., Ахундов В. М., Минаев Э. С. Анализ и прогноз развития больших технических систем. – М.: Наука, 1982. – 282 с.

в том числе в условиях штатного применения [3–9]. В условиях штатного применения возникают *отказы*, порядок устранения которых строго регламентирован в эксплуатационной документации и обеспечен всеми имеющимися силами и средствами и априори известен. Применяемые методы оценивания устойчивости КИИ основаны на моделях *надежности* функционирования восстанавливаемых технических устройств, которые построены на результатах наблюдения за *стационарными* условиями их эксплуатации. При этом следует учитывать, что теория надежности – это наука экспериментальная, рабочие гипотезы и выводы которой базируются на конкретных экспериментальных данных. Экспериментальные данные и выводы всегда отражают некоторую предысторию, а принимаемые решения базируются на отрицательном опыте прошлого. Предметом теории надёжности являются методы обеспечения *стабильности* работы объектов в процессе проектирования, производства, приёмки, транспортировки, эксплуатации и хранения. Вопросам оценивания устойчивости функционирования объектов экономики в чрезвычайных ситуациях методом прогнозирования и моделирования производственных функций посвящены работы [10, 11]. В [10] автор делает вывод, что применение основных положений оценивания устойчивости функционирования объектов экономики методом прогнозирования может оптимизировать использование сил и средств при ликвидации последствий чрезвычайных ситуаций с помощью коэффициентов надёжности. В [11] автор предлагает подход к моделированию производственных функций в условиях стохастической неопределенности с целью оценивания риска. Во всех упомянутых выше публикациях в основу оценивания был положен циклический процессный подход с обучением.

Для прогнозирования применительно к воздействию угроз нет и не может быть репрезентативной статистики, прежде всего по причине того, что обстановка динамична, сами явления весьма редки и зависят от сценария воздействия угроз (поведенческой неопределенности). В этом случае справедливо предварительное утверждение (гипотеза), что применение классических методов теории надежности, основанных на процессном подходе с обучением, для прогнозирования могут давать ошибочные результаты. Отсюда следует, что применительно к условиям воздействия угроз требуются обобщение методов теории надежности на случаи, когда процесс применения не является стационарным, ограничен по времени и ему присуща поведенческая неопределенность и отсутствуют условия для создания обучающей выборки.

Решение задачи

Учитывая приведенные выше обстоятельства, предлагается авторский подход к решению поставленной задачи. Для решения задачи предлагается использовать модель прогнозируемого объекта в виде полного дифференциала функции (1)

$$dU^* = \sum_{i=1}^{n} \frac{\partial U^*}{\partial g_i} \Delta g_i,$$

где $\Delta g_i = dg_i$ с учетом того, что g_i являются аргументами в (1).

В соответствии с исходным допущением в) при приращении параметров

$$\Delta g_i = (\mathrm{r} - 1)\, g_i,\ i = \overline{1, n},$$

полный дифференциал

$$dU^* = [\gamma(r) - 1]U^* - \varepsilon,$$

где ε – величина, которую можно считать величиной высшего порядка малости, которую при расчетах можно не учитывать.

В результате

$$[\gamma(r) - 1]U^* - \varepsilon = \sum_{i=1}^{n} \frac{\partial U^*}{\partial g_i} (r - 1) g_i$$

или

$$\frac{\gamma(r) - 1}{r - 1} U^* - \frac{\varepsilon}{r - 1} = \sum_{i=1}^{n} \frac{\partial U^*}{\partial g_i} g_i.$$

При $\mathrm{r} \longrightarrow 1$ получаем

$$\rho U^* = \sum_{i=1}^{n} \frac{\partial U^*}{\partial g_i} g_i, \qquad (2)$$

где $\rho = \frac{\partial \gamma(r)}{\partial r}$ – показатель скорости роста устойчивости, характеризующий возрастание ее уровня при увеличении значений параметров, характеризующих СБ.

Таким образом, определение значений функции (1) сводится к решению дифференциального уравнения (2) в частных производных.

Для получения решения определяются начальные условия.

1. Пусть $g_1, g_2, \ldots, g_n = 1$, тогда $U^* = U_l$, так как $g_i = \frac{x_i^*}{x_i}$ и $g_i = 1$, то $x_i^* = x_i$ и $U^* = U(1, 1, \ldots, 1) = U_l$.

При данных начальных условиях можно получить следующие частные решения уравнения (2):

$$U^* = U_l \prod_{i=1}^{n} g_i^{\lambda},\ \lambda > 0; \qquad (3)$$

$$U^* = U_l [\sum_{i=1}^{n} \beta_i g_i]^{\omega},\ \omega > 0; \qquad (4)$$

$$U^* = U_l \prod_{i=1}^{n} g_i^{\alpha i},\ \sum_{i=1}^{n} \alpha_i = 1; \qquad (5)$$

$$U^* = U_l \sum_{i=1}^{n} \beta_i g_i,\ \sum_{i=1}^{n} \beta_i = 1; \qquad (6)$$

где α, β – показатели степени, характеризующие влияние произведения и взвешенной суммы значений

параметров на устойчивость U^*; α_i, β_i – коэффициенты важности, характеризующие влияние каждого отдельного параметра g_i на устойчивость ОИ.

Полученные выражения (3–6) удовлетворяют уравнению (2) при различных значениях коэффициента роста ρ. Выражению (3) соответствует $\rho = \lambda n$, т. е. при одновременном возрастании параметров в r раз устойчивость возрастает в $\gamma(r) = \lambda n r$ раз. Выражению (4) соответствует $\rho = \omega$ и $\gamma(r) = \omega r$. Выражения (5) и (6) удовлетворяют уравнению (2) при $\beta = 1$ и $\gamma(r) = r$.

Данные выражения могут быть использованы для определения прогнозируемого значения устойчивости U_π^* посредством подстановки соответствующих значений параметров $g_{\pi i} = \frac{x_{\pi i}^*}{x_{li}}$. Однако для этого необходимо определить значения λ, ω, α_i, β_i и выбрать из числа полученных соотношений $U_\pi^* = U(g_{\pi 1}, g_{\pi 2}, ..., g_{\pi n})$, обеспечивающих минимальную погрешность прогнозирования.

Последовательность прогнозирования. В соответствии с исходными данными и полученными соотношениями процесс прогнозирования можно осуществить в следующей последовательности:

1) определение показателя λ в выражении (3);
2) определение коэффициентов важности α_i и β_i в выражениях (4), (5), (6);
3) определение показателя ω в выражении (4);
4) выбор соотношения $U_\pi^* = U(g_{\pi 1}, g_{\pi 2}, ..., g_{\pi n})$, обеспечивающего минимальную погрешность прогнозирования;
5) определение прогнозируемого уровня устойчивости U_π^*.

Определение показателя λ производится на основании исходных данных об устойчивостях U_l и соответствующих параметров $x_{l,i}$ для объектов-аналогов, $l = \overline{1, L}$, $i = \overline{1, n}$.

Выбирается один из объектов-аналогов, имеющий максимальный уровень устойчивости U_k, в качестве базового. Для каждого l-го объекта оценивания (ОО) вычисляется значение

$$\pi_l = \prod_{i=1}^{n} g_{l,i}$$

и соответствующие отношения

$$\delta_l = \frac{U_l}{U_k}, \; l = \overline{1, L},$$

где $g_{l,i} = \frac{x_{li}}{x_{\kappa,i}}$, $x_{l,i}$ – значение i-го параметра l-го ОО; $x_{\kappa,i}$ – значение i-го параметра базового к-го объекта-аналога.

В результате получаем L точек, характеризующих зависимости $\delta = \varphi(\pi)$, в которой при $l = \kappa$, $\pi_l = 1$ и $\delta_l = 1$. Предполагается, что данная зависимость сохраняется для объектов-аналогов, которые являются прототипами:

$$\delta_\Pi = \frac{U_\Pi}{U_l} = \varphi(\pi_\Pi),$$

где $\pi_\Pi = \prod_{i=1}^{n} g_{\Pi i}$; $g_{\Pi i} = \frac{x_{\Pi i}}{x_{l,i}}$.

Если при этом $U_\Pi = U_l$, то $\delta_\Pi = 1$, $\pi_\Pi = 1$, что соответствует начальной точке при $l = \kappa$ для объектов-аналогов.

Далее сглаживается зависимость $\delta = \varphi(\pi)$ с помощью функции вида $\delta = \pi^\lambda$. Для сглаживания используется метод наименьших квадратов, в соответствии с которым необходимо определить такое значение λ, которое обеспечивает минимум квадратичной формы[4]:

$$S(\lambda) = \sum_{l=1}^{L} (\pi_l^\lambda - \delta_l)^2.$$

Значение λ определяется в результате решения уравнения

$$\frac{dS}{d\lambda} = 2\sum_{l=1}^{L} (\pi_l^\lambda - \delta_l)\, \pi_l^\lambda \ln \pi_l = 0$$

или

$$\sum_{l=1}^{L} (\pi_l^\lambda - \delta_l)\, \pi_l^\lambda \ln \pi_l = 0 \qquad (7)$$

Определение значений коэффициентов важности α_i и β_i производится в следующей последовательности. Вначале с использованием более простого выражения (6) определяются значения β_i, а затем на основании равенства выражений (5) и (6) при $g_i = g_{\Pi i}$ подбираются значения α_i. Для снижения сложности расчетов целесообразно взять логарифмы от этих выражений:

$$\sum_{i=1}^{n} \alpha_i \lg g_{\Pi i} = \lg \sum_{i=1}^{n} \beta_i g_{\Pi i}. \qquad (8)$$

Аналитические зависимости, характеризующие коэффициент важности β_i, можно определить путем подстановки функции (6) в уравнение (2) при $\rho = 1$. В результате получим

$$U_l \sum_{i=1}^{n} \beta_i g_i = \sum_{i=1}^{n} \frac{\partial U}{\partial g_i}\bigg|_{g_i = 1} g_i$$

или

$$\sum_{i=1}^{n} \beta_i g_i = \sum_{i=1}^{n} \frac{v_i}{U_l} g_i, \qquad (9)$$

где $v_i = \frac{\partial U}{\partial g_i}\Big|_{g_i = 1}$ – интенсивность изменения устойчивости при изменении параметра g_i относительно начального значения $g_i = 1$.

Равенство (9) сохраняется в случае

$$\beta_i = \frac{v_i}{U_l}. \qquad (10)$$

Таким образом, значение коэффициента важности β_i прямо пропорционально приращению устойчивости функционирования объекта на единицу приращения параметра g_i^*. Для определения конкретных

4 Линник Ю. В. Метод наименьших квадратов и основы теории обработки наблюдений. – М.: Наука, 1962. 349 с.

значений β_i можно использовать следующие методы: 1) метод, основанный на решении системы линейных уравнений; 2) метод наименьших квадратов; 3) методы экспертных оценок. Для повышения достоверности и при наличии ресурса рекомендуется использовать все методы в комплексе с последующим сравнением полученных результатов.

Система линейных уравнений составляется на основе исходных данных и имеет вид:

$$\left.\begin{array}{l}\beta_1 g_{11} + \beta_2 g_{12} + \ldots + \beta_n g_{1n} = \delta_1;\\ \beta_1 g_{21} + \beta_2\, g_{22} + \ldots + \beta_n g_{2n} = \delta_2;\\ \ldots\\ \beta_1 g_{n-1,1} + \beta_2 g_{n-1,2} + \ldots + \beta_n g_{n-1,n} = \delta_{n-1};\\ \beta_1 + \beta_2 + \ldots + \beta_n = 1)\end{array}\right\} \quad (11)$$

где $g_{l,i} = \frac{x_{li}}{x_{к,i}}$, $x_{l,i}$ – значение i-го параметра l-го объекта-аналога; $x_{к,i}$ – значения i-го параметра объекта-аналога, выбранного в качестве базового; $\delta_l = \frac{U_l}{U_к}$ – приведенная устойчивость l-й объекта-аналога к устойчивости базового объекта-аналога ; $U_к$ – устойчивость функционирования для базового к объекта-аналога; $l = 1, n-1$; $i = 1, n$; $U_к < U_l$.

Определение значений β_i, $i = 1, n$, посредством решения системы (11) возможно, если $L \geq n$. т. е. число объектов-аналогов L не меньше числа параметров n. Неравенство $L \geq n$ получается из условия, что один из объектов аналогов выбирается в качестве базового, а одно из уравнений представляет собой условие

$$\sum_{i=1}^{n} \beta_i = 1.$$

При использовании метода наименьших квадратов необходимо определить также значения β_i, которые обеспечивают минимум квадратичной формы:

$$S(\beta_1, \beta_2, \ldots, \beta_n) = \sum_{l=1}^{L} (\varphi_l - \delta_l)^2,$$

где

$$\varphi_l = \sum_{i=1}^{n} \beta_i\, g_{l,i}.$$

В результате получаем систему уравнений:

$$\left.\begin{array}{l}\sum\limits_{l=1}^{L} (\varphi_l - \delta_l)\, g_{l,1} = 0;\\ \sum\limits_{l=1}^{L} (\varphi_l - \delta_l)\, g_{l,2} = 0;\\ \ldots\\ \sum\limits_{l=1}^{L} (\varphi_l - \delta_l)\, g_{l,n-1} = 0;\\ \beta_1 + \beta_2 + \ldots + \beta_n = 1\end{array}\right\} \quad (12)$$

Подставляем в данную систему выражение

$$\sum_{i=1}^{n} \beta_i\, g_{l,i} = \varphi_l$$

и путем решения полученной системы определяем неизвестный параметры β_i.

Метод наименьших квадратов имеет существенные преимущества, так как позволяет определять значения β_i при $L < n$ и дает меньшую погрешность приближения.

Метод экспертных оценок целесообразно использовать в тех случаях, при которых число L объектов-аналогов незначительно или когда значения коэффициентов важности β_i параметров объектов аналогов не соответствуют значениям β_i оцениваемого объекта.

Анализ методов экспертного оценивания позволяет утверждать, что в наибольшей степени подходит метод анкетного опроса, в ходе которого каждый эксперт попарно оценивает степень влияния каждого параметра на устойчивость функционирования ОО.

Таблица 1.

Параметры существенные для прогнозирования затрат

Наименование параметра	**Физический смысл**	**Тип шкалы**	**Диапазон изменения**
g_1 – отношение затрат на ИБ к затратам на ИТ	Доля ресурса, выделенного на обеспечение устойчивости функционирования на период $(0, T]$	$g_1 = \frac{C_{ИБ} - C_{ИТ}}{C_{ИБ}}$	(–1, +1)
g_2 – минимум функции устойчивости (функции живучести)	Максимум вероятности того, что в результате воздействия угроз на периоде $(0, T]$ объект потеряет свою функциональность	$u_m = \min\limits_{t \in (0, T]} u(t)$	(0,1)
g_3 – доля резерва невозобновляемого ресурса	Доля резерва от общей ресурсоемкости восстановительных работ оцениваемого объекта, выделенного на период $(0, T]$, для обеспечения восстановления функциональности	$g_3 = \frac{C_Р - C_{ИБ}}{C_{ИБ}}$	(–1, +1)
g_4 – вариабельность времени восстановления	Отношение разницы нижней и верхней границ прогностических оценок времени восстановления функциональности после успешного воздействиям угроз к периоду $(0, T]$	$g_4 = \frac{\tau - \hat{\tau}}{T}$	(0,1)

Период $(0, T]$ – прогнозируемая продолжительность периода воздействия угроз, T – прогнозируемый момент времени завершения восстановления функциональности объекта оценивания от начала отчета – 0.

СИБ – стоимость мероприятий по обеспечению ИБ на периоде $(0, T]$;

СИТ – стоимость мероприятий по обеспечению ИТ на периоде $(0, T]$;

СР – стоимость резерва (невозобновляемого ресурса, находящегося в резерве);

$\hat{\tau}$ – верхняя прогностическая оценка времени восстановления функциональности;

$\underline{\tau}$ – нижняя прогностическая оценка времени восстановления функциональности;

Физический смысл и содержание показателей приведены в табл. 1

Далее эксперту предлагается заполнить квадратную таблицу-матрицу, например, табл. 2. Каждой строке и столбцу матрицы соответствует определенный параметр. Заполнение матрицы осуществляется по следующему правилу.

Выделяется первая строка, соответствующая параметру g_1. Выделенный параметр последовательно сравнивается с другими параметрами, указанными в соответствующих столбцах. Если, по мнению эксперта, приращение устойчивости на единицу первого параметра g_1 больше, чем на единицу сравниваемого параметра g_r, то на пересечении первой строки и r-го столбца проставляется 1, в противном случае – 0 и т. д. Затем сравнение повторяется для остальных строк матрицы. После проведения экспертного опроса группы из m экспертов прогнозист получает m заполненных матриц

$$D_j = \|d_{i,r,j}\|,\ i,\ r = \overline{1,\ n};\ j = \overline{1,\ m},$$

где $d_{i,r,j}$ – обобщающий элемент j-й матрицы, характеризующий предпочтение j-го эксперта, по параметру g_i в отношении параметра g_r по затратам.

Статистическая обработка полученных результатов осуществляется следующим образом:

1. Суммируются элементы каждой i-й строки матрицы D_j и определяется балл

$$d_{i,j} = \sum_{r=1}^{n} d_{i,r,j},\ i = \overline{1,\ n},$$

характеризующий предпочтение параметра g_i перед остальными параметрами по затратам.

По полученным величинам $d_{i,j}$ вычисляется коэффициент важности $\beta_{i,j}$, определяющий, по мнению эксперта j, относительное предпочтение параметра g_i перед остальными параметрами по затратам:

$$\beta_{i,j} = \frac{d_{i,j}}{\sum_{i=1}^{n} d_{i,i}}. \qquad (13)$$

2. Затем вычисляется среднее значение

$$\bar{\beta}_i = \frac{1}{m} \sum_{j=1}^{m} \beta_{i,j},\ i = \overline{1,\ n}, \qquad (14)$$

и дисперсия

$$D_{\beta_i} = \frac{1}{m-1} \sum_{j=1}^{m} (\beta_{i,j} - \bar{\beta}_i)^2,\ i = \overline{1,\ n}, \qquad (15)$$

каждого коэффициента важности β_i.

3. Определение показателя ω в выражении (4) производится на основании исходных данных о значениях устойчивости функционирования объектов-аналогов, а также полученных значений коэффициентов важности β_i. Определение осуществляется по аналогии, как и при вычисления показателя λ, только для каждого l-го объекта-аналога вместо π_l вычисляется значение

$$\sigma_l = \sum_{i=1}^{n} \beta_i\, g_{l,i}.$$

Таблица 2.

Результат анкетного опроса (вариант)

Наименование показателя	**g_1 – отношение затрат на ИБ к затратам на ИТ**	**g_2 – минимум функции устойчивости**	**g_3 – доля резерва невозобновляемого ресурса**	**g_4 – вариабельность времени восстановления**
g_1 – отношение затрат на ИБ к затратам на ИТ	–	0	0	1
g_2 – минимум функции устойчивости		–	1	1
g_3 – доля резерва невозобновляемого ресурса			–	1
g_4 – вариабельность времени восстановления				–

В результате получаем L точек (σ_l, δ_l), характеризующих значение графика $\delta = f(\sigma)$. Сглаживаем зависимость $\delta = f(\sigma)$ функцией $\delta = \sigma^{\omega}$ с использованием метода наименьших квадратов. Значение ω определяется путем решения уравнения

$$\sum_{l=1}^{L}(\sigma_l^{\omega} - \delta_l)\,\sigma_l^{\omega}\ln\sigma_l = 0.$$

Выбор соотношения $U_{\text{П}} = U(g_{n,1}, g_{n,2}, \ldots, g_{n,n})$, обеспечивающего наименьшую погрешность прогнозирования, производится из семейства выражений (3), (4), (5), (6). Выбор осуществляется на основании критерия минимума суммы квадратов отклонений вычисленных уровней устойчивости соответствующих объектов-аналогов, т. е. необходимо выбрать такое соотношение j, которое бы обеспечивало

$$\sum_{l=1}^{L}(\varphi_{l,j} - \delta_l)^2 = \min_j,\ j = 1, 3, 3, 4, \ldots, \qquad (16)$$

где

$$\varphi_{l,1} = \prod_{i=1}^{n} g_{li}^{\lambda};\ \varphi_{l,2} = \left[\sum_{i=1}^{n}\beta_i\, g_{l,i}\right]^{\omega};\ \varphi_{l,3} = \prod_{i=1}^{n} g_{li}^{\alpha_i};$$

$$\varphi_{l,4} = \sum_{i=1}^{n}\beta_i\, g_{l,i};\ g_{l,i} = \frac{x_{l,i}}{x_{\text{к},i}};\ \delta = \frac{U_l}{U_{\text{к}}},$$

где к – индекс базового объекта-аналога.

Выбранное с помощью данного критерия соотношение может быть рекомендовано в качестве прогноза устойчивости ОО после его реинжиниринга (обеспечение устойчивости его функционирования в условиях воздействия угроз).

Определение прогнозируемых затрат $U_{\text{П}}$ производится посредством подстановки соответствующих значений параметров $g_{\text{П}i} = \frac{x_{\text{П},i}}{x_{l,i}}$ в выбранное соотношение и вычисление УП. Например, пусть в соответствии с критерием (16) наименьшую погрешность обеспечивает соотношение (4). Тогда

$$U_{\text{П}} = U_l\left[\sum_{i=1}^{n}\beta_i\, g_{\text{П},i}\right]. \qquad (17)$$

Если значение $g_{\text{П},i}$ задаются приблизительно с помощью законов распределения или числовых характеристик, то с использованием соотношения (17) также определяются соответствующие законы распределения или числовые характеристики (математическое ожидания и дисперсия) величины $U_{\text{П}}$. При этом необходимо учитывать среднюю погрешность, допускаемую при использовании приближенного выражения (17).

Для априорной оценки погрешности прогноза статистическими методами можно использовать размер доверительного интервала. Модель прогноза считается более точной, если при одной и той же доверительной вероятности она даёт более узкий доверительный интервал по сравнению с другой моделью.

Выводы

В результате настоящего исследования предложен авторский подход к прогнозированию уровня устойчивости объектов КИИ после реинжиниринга его СБ, который позволяет решить задачу обоснования проектного решения на ранних этапах. Предложенный подход отличается от известных применением методов математического анализа совместно с экстраполяционными и экспертными методами, что в совокупности позволяет построить: а) результативную прогностическую модель ОО; б) прогноз за пределами обучения на основе параллельного балансирования влияния выбранных параметров.

При применении метода следует учитывать следующие обстоятельства:

- если между одноименными параметрами существует нелинейная зависимость, линейная модель может давать результаты с большой погрешностью;
- модель чувствительна к выбросам значений одноименных параметров.

При выборе метода следует учитывать, что предлагаемая модель уступает статистическим моделям в точности оценивания зависимой переменной, поэтому ее следует применять, если отсутствует репрезентативная статистика.

В качестве объекта прогнозирования могут выступать и другие факторы, например:

- *управляемые факторы*, которые можно корректировать посредством организации целенаправленных воздействий, например стоимость владения, вероятность поражения, устойчивость функционирования и т. п;
- *неуправляемые* факторы, будущее проявление которых можно только прогнозировать, но изменять практически невозможно, например сценарий воздействия противника, физико-географические особенности и т. п.

Работа выполнена при поддержке Фонда Потанина.

Литература
1. Язов Ю. К. Об определении понятия «кибербезопасность» и связанных с ним терминов // Вопросы кибербезопасности. 2025. № 1(65). С. 2–6. DOI:10.21681/2311-3456-2025-1-2-6.
2. Воеводин В. А. Генезис понятия структурной устойчивости информационной инфраструктуры автоматизированной системы управления производственными процессами к воздействию целенаправленных угроз информационной безопасности. Вестник Воронежского института ФСИН России. 2023. № 2, апрель–июнь. С. 30–41.
3. Стародубцев Ю. И. Структурно-функциональная модель киберпространства/ Ю. И. Стародубцев, П. В Закалкин, С. А Иванов // Вопросы кибербезопасности. 2021. № 4(44), с. 16–24. DOI:10.21681/2311-3456-2021-4-16-24.
4. Фатин А. Д., Павленко Е. Ю. Анализ моделей представления киберфизических систем в задачах обеспечения информационной безопасности // Проблемы информационной безопасности. Компьютерные системы. 2020. № 2. С. 109–121.
5. Зегжда Д. П. Кибербезопасность цифровой индустрии. Теория и практика функциональной устойчивости к кибератакам / под ред. Д. П. Зегжды. – М.: Горячая линия – Телеком. 2022. – 560 с.
6. Коноваленко С. А. Методика оценивания функциональной устойчивости гетерогенной системы обнаружения, предупреждения и ликвидации последствий компьютерных атак // Системы управления, связи и безопасности. 2023. № 4. С. 157–195. doi: 10.24412/2410-9916-2023-4-157-195.
7. Ерохин С. Д., Петухов А. Н., Пилюгин П. Л. Управление безопасностью критических информационных инфраструктур. М.: Горячая линия – Телеком. 2023. – 240 с.
8. Коцыняк М. А., Осадчий А. И., Коцыняк М. М., Лаута О. С., Дементьев В. Е., Васюков Д. Ю. Обеспечение устойчивости информационно-телекоммуникационных сетей в условиях информационного противоборства. СПб.: ЛО ЦНИИС, 2014. – 126 с.
9. Одоевский С. М., Лебедев П. В. Методика оценки устойчивости функционирования системы технологического управления инфокоммуникационной сетью специального назначения с заданной топологической и функциональной структурой // Системы управления, связи и безопасности. 2021. № 1. – С. 152–189.
10. Евстропов В. М. Основные положения, используемые при оценке устойчивости функционирования объектов экономики в чрезвычайных ситуациях методом прогнозирования / В. М. Евстропов // Заметки ученого. – 2021. № 10. – С. 321–325.
11. Долгов А. В. Анализ современных подходов к моделированию производственных функций в условиях неопределенности / А. В. Долгов // Вестник Волжского университета им. В. Н. Татищева. – 2024, Т. 2, № 1(53). – С. 37–45. – DOI 10.51965/2076-7919_2024_2_1_37.

ON FORECASTING COSTS FOR THE RE-ENGINEERING OF THE SECURITY SYSTEM OF CRITICAL INFORMATION INFRASTRUCTURE OBJECTS EXPOSED TO THREATS

Voevodin V. A.[5]

Keywords: *cost forecasting, operation sustainability, object-analog, re-engineering, critical information infrastructure object, threat.*

The objective of the study: *to justify the relevance, formulate and formalize the scientific task of forecasting costs for re-engineering the security system of critical information infrastructure (CII) objects exposed to threats.*

Methods of research: *heuristic, extrapolation, expert, comparison and comparison, differential calculus, informational diagnosis.*

The results obtained: *a verbal and formal formulation of the scientific task of forecasting costs for re-engineering the security system of KIA objects exposed to threats was formulated and an algorithm for its solution was proposed.*

Scientific novelty: *a tool for forecasting is proposed, based on the structure of a prognostic model, which represents a linear combination of the same-name parameters of the object of prediction and objects-analogues. Based on data about the value of parameters of objects-analogues, when namesake parameters have linear correlation, a forecast is made. For example, it is possible to predict the cost of reengineering, the sustainability of operation, the reserve stock, etc.*

Practical significance: *the scientific problem can serve as a basis for formulating a technical task to reengineer the safety systems of the KIA with the requirements for their sustainability.*

References
1. Jazov Ju. K. Ob opredelenii ponjatija «kiberbezopasnost'» i svjazannyh s nim terminov // Voprosy kiberbezopasnsoti. 2025. № 1(65). S. 2–6. DOI:10.21681/2311-3456-2025-1-2-6.
2. Voevodin V. A. Genezis ponjatija strukturnoj ustojchivosti informacionnoj infrastruktury avtomatizirovannoj sistemy upravlenija proizvodstvennymi processami k vozdejstviju celenapravlennyh ugroz informacionnoj bezopasnosti. Vestnik Voronezhskogo instituta FSIN Rossii. 2023. № 2, aprel'–ijun'. S. 30–41.
3. Starodubcev Ju. I. Strukturno-funkcional'naja model' kiberprostranstva/ Ju. I. Starodubcev, P. V. Zakalkin, S. A. Ivanov // Voprosy kiberbezopasnosti. 2021. № 4(44), s. 16–24. DOI:10.21681/2311-3456-2021-4-16-24.

5 Vladislav A. Voevodin, Ph.D. in Technical Sciences, MIET, Moscow, Russia. AuthorID: 1012813, ORCID 0009-0003-9431-1685. E-mail: vva541@mail.ru

4. Fatin A. D., Pavlenko E. Ju. Analiz modelej predstavlenija kiberfizicheskih sistem v zadachah obespechenija informacionnoj bezopasnosti // Problemy informacionnoj bezopasnosti. Komp'juternye sistemy. 2020. № 2. S. 109–121.
5. Zegzhda D. P. Kiberbezopasnost' cifrovoj industrii. Teorija i praktika funkcional'noj ustojchivosti k kiberatakam / pod red. D. P. Zegzhdy. – M.: Gorjachaja linija – Telekom. 2022. – 560 s.
6. Konovalenko S. A. Metodika ocenivanija funkcional'noj ustojchivosti geterogennoj sistemy obnaruzhenija, preduprezhdenija i likvidacii posledstvij komp'juternyh atak // Sistemy upravlenija, svjazi i bezopasnosti. 2023. № 4. S. 157–195. DOI: 10.24412/2410-9916-2023-4-157-195.
7. Erohin S. D., Petuhov A. N., Piljugin P. L. Upravlenie bezopasnost'ju kriticheskih informacionnyh infrastruktur. M.: Gorjachaja linija – Telekom. 2023. – 240 s.
8. Kocynjak M. A., Osadchij A. I., Kocynjak M. M., Lauta O. S., Dement'ev V. E., Vasjukov D. Ju. Obespechenie ustojchivosti informacionno-telekommunikacionnyh setej v uslovijah informacionnogo protivoborstva. SPb.: LO CNIIS, 2014. – 126 s.
9. Odoevskij S. M., Lebedev P. V. Metodika ocenki ustojchivosti funkcionirovanija sistemy tehnologicheskogo upravlenija infokommunikacionnoj set'ju special'nogo naznachenija s zadannoj topologicheskoj i funkcional'noj strukturoj // Sistemy upravlenija, svjazi i bezopasnosti. 2021. № 1. – S. 152–189.
10. Evstropov V. M. Osnovnye polozhenija, ispol'zuemye pri ocenke ustojchivosti funkcionirovanija ob#ektov jekonomiki v chrezvychajnyh situacijah metodom prognozirovanija / V. M. Evstropov // Zametki uchenogo. – 2021. № 10. – S. 321–325.
11. Dolgov A. V. Analiz sovremennyh podhodov k modelirovaniju proizvodstvennyh funkcij v uslovijah neopredelennosti / A. V. Dolgov // Vestnik Volzhskogo universiteta im. V. N. Tatishheva. – 2024, T. 2, № 1(53). – S. 37–45. – DOI: 10.51965/2076-7919_2024_2_1_37.

КОЛЛАБОРАТИВНОЕ ПОСТРОЕНИЕ МОДЕЛИ ГРЕБНЕВОЙ ЛИНЕЙНОЙ РЕГРЕССИИ В РАСПРЕДЕЛЕННОЙ СИСТЕМЕ С ВИЗАНТИЙСКИМИ ОТКАЗАМИ

Волкова Е. С.[1], Гисин В. Б.[2]

DOI: 10.21681/2311-3456-2025-5-50-57

Цель исследования: *разработка алгоритма построения модели гребневой линейной регрессии в распределенной системе с византийскими отказами узлов.*

Методы исследования: *применение техники работы со статистическими данными высокой размерности и применение протоколов организации вычислений в распределенных сетях.*

Полученный результат: *описан механизм достижения усредненного согласия узлами асинхронной сети и его применение для построения регрессионной модели. Приведены оценки параметров сети, при которых алгоритм достижения усредненного согласия применим: распределение данных между узами может быть неоднородным; византийские узлы могут отклоняться от исполнения сетевого протокола произвольным образом; ни один честный узел не знает, какие из остальных узлов являются честными; византийские узлы знают друг друга и могут вступать в сговор. Ошибки линейной регрессии предполагаются субгауссовскими и независимыми.*

Научная новизна: *разработан метод достижения усредненного согласия относительно параметров регрессии в асинхронной системе.*

Ключевые слова: *федеративное машинное обучение, регуляризация по Тихонову, консенсус.*

Введение

Успехи современного машинного обучения были достигнуты в условиях, когда модель обучается на большом объеме данных. Растущий объем доступных данных, а также растущая сложность моделей машинного обучения привели к созданию схем обучения, требующих больших вычислительных ресурсов. Как следствие, многие реализации машинного обучения промышленного уровня в настоящее время являются распределенными (см. [1–3]).

Механизм извлечения и обобщения знаний из различных источников и обновления алгоритмов принятия решений должен обеспечить обработку данных и не быть уязвимым для вредоносной входной информации. Данные на отдельных платформах могут быть использованы для обучения, но не могут быть переданы в открытое пользование. Включение таких данных в обучение возможно в рамках распределенных моделей. Данные, доступные узлам распределенной сети, могут быть неоднородными. Существует множество факторов, которые влияют на характеристики данных, таких как предпочтения пользователей, методы сбора данных и характеристики клиентов. Неоднородность данных, вообще говоря, приводит к увеличению затрат на связь, и неравномерному обновлению локальных моделей [4–6].

Одним из вариантов обучения распределенных моделей является федеративное обучение. При федеративном обучении клиенты сотрудничают в процессе обучения, и, не раскрывая своих данных, проводят обучение модели на всей совокупности данных. Федеративное обучение не допускает прямой передачи необработанных данных и, более того, может потребовать применения тех или иных методов защиты данных, поскольку в ряде случаев должно обеспечивать безопасность и конфиденциальность данных. Абстрагируясь от вопросов информационной безопасности, как это сделано в [7], можно выделить две архитектуры систем федеративного машинного обучения:

- данные распределены между узлами, все вычисления выполняются самими участниками в ходе выполнения протокола, отсутствует третья сторона, которой передаются данные для вычислений;
- имеется большое число агентов, которые передают свои данные в центры обработки информации, где и осуществляется обучение моделей.

1 Волкова Елена Сергеевна, кандидат физико-математических наук, доцент, Финансовый университет при Правительстве Российской Федерации, Москва, Россия. E-mail: evolkova@fa.ru

2 Гисин Владимир Борисович, кандидат физико-математических наук, профессор, Финансовый университет при Правительстве Российской Федерации, Москва, Россия. E-mail: vgisin@fa.ru

Распределение вычислений по сети узлов повышает риск нарушений протокола. К ним относятся сбои и ошибки в вычислениях, остановка процессов, искажения в способе распределения выборок данных между процессами, а также, в худшем случае, попытки злоумышленников скомпрометировать процессы. В банке данных угроз безопасности информации (БДУ)[3] указаны две угрозы, связанные с машинным обучением: угроза модификации модели машинного обучения путем искажения («отравления») обучающих данных (УБИ.221) и угроза подмены модели машинного обучения (УБИ.222).

В [7] дана классификация моделей федеративного машинного обучения в зависимости от потенциально опасных внешних факторов и категории нарушителей и указаны инструменты обеспечения конфиденциальности данных для систем разных типов. Основным инструментом является криптография. Конкретный вариант ее применения зависит от определения безопасности решаемой задачи. В частности, например, при горизонтальной схеме требуется неразглашение данных, хранящихся в отдельных узлах, в то время как нарушителями может выступать часть участников (узлов) сети.

Согласно знаменитой FLP-теореме о невозможности (Фишер, Линч, Паттерсон) детерминированный консенсус в асинхронных средах невозможен даже в том случае, когда среди узлов имеется хотя бы один, который может прекратить работу. Альтернативой детерминированному консенсусу может служить приблизительный консенсус (см. [8]). В этом случае значения, предлагаемые честными узлами, сходятся к значениям, близким друг к другу, оставаясь в выпуклой оболочке значений, предложенных честными узлами. Достижение консенсуса требует при этом порядка n^d локальных вычислений, где d – размерность пространства параметров, а n – число процессов (узлов). Допустимая доля «нечестных» узлов должна быть при этом обратно пропорциональна размерности пространства параметров [9].

На сегодня предложено значительное количество алгоритмов агрегирования параметров локальных моделей. Наиболее популярным из них является алгоритм федеративного усреднения (см. [10]). Детальный анализ состояния исследований по методам федеративного обучения можно найти в [11].

Модель распределенной сети

В настоящей работе рассматривается стандартная одноранговая децентрализованная модель распределенных вычислений (в [13] для обозначения таких моделей предложен термин «роевые»). Сеть состоит из n узлов, из которых h являются честными, а $f = n - h$ – византийскими, т.е. такими, которые могут отклоняться от исполнения сетевого протокола произвольным образом. Множество всех узлов обозначим N, множество всех честных узлов – H, множество всех византийских узлов – F. Ни один честный узел не знает, какие из остальных узлов являются честными. Византийские узлы знают друг друга и могут вступать в сговор. Можно считать, что все византийские узлы контролирует один противник (злоумышленник). В дальнейшем будем считать, что $f > 0$.

Византийские узлы могут отправлять произвольные сообщения, они могут отправлять разные сообщения на разные узлы. Злоумышленник имеет доступ ко всей информации об обучении, включая цель обучения, используемый алгоритм, а также набор данных. Злоумышленник может задерживать отправку сообщений на честные узлы. Тем не менее, мы предполагаем, что злоумышленник не в состоянии задерживать все сообщения на неопределенный срок. Кроме того, злоумышленник не в состоянии изменить сообщения от честных узлов, защищенные подписью отправителя.

Система предполагается асинхронной. Работа каждого честного узла разбита на раунды (подобно тому, как это сделано в [12]). В каждом раунде каждый честный узел передает сообщение с указанием номера раунда и ожидает, пока удастся собрать сообщения от $q \le h$ других узлов (с правильно указанным раундом), прежде чем выполнить некоторые локальные вычисления и перейти к следующему раунду. Несмотря на то, что сеть асинхронна, каждый раунд в конечном итоге завершится для всех честных узлов, поскольку все сообщения от h честных узлов будут в конечном итоге доставлены. Некоторые из них могут быть доставлены после того, как узел получит q сообщений (включая сообщения византийских узлов). Такие сообщения не принимаются во внимание.

Каждому узлу $j \in N$ доступны данные, представленные ограниченным распределением в $\mathbb{R}^d$ с нулевым средним. Распределения для разных узлов, вообще говоря, могут различаться. Такие распределения принято называть non-IID. Описание методов оценки неоднородности распределения данных можно найти в [13].

Для каждого j существует вектор параметров $\beta_j \in \mathbb{R}^d$ такой, что ответ Y_j на случайный запрос X_j имеет вид

$$Y_j(X_j) = \beta_j^T X_j + \xi_j,$$

где ξ_j – σ-субгауссовская случайная величина с нулевым средним. Кроме того, предполагается, что матрица $\sum_j = \mathbb{E}[X_j X_j^T]$, где $X_j \in \mathbb{R}^d$ – случайный запрос, положительно определена.

3 ФСТЭК России. Банк данных угроз безопасности информации. ФАУ «ГНИИИ ПТЗИ ФСТЭК России», https://bdu.fstec.ru/threat?page=22

Участники сети решают задачу построения общей модели, т. е. поиска параметров $\theta_j^* \in \mathbb{R}^d$, $j \in H$, так, чтобы векторы θ_j^* были достаточно близки у честных участников, а функции потерь не слишком возрастали при переходе от локальных моделей к модели, выработанной коллаборативно.

Для честного узла $j \in H$ локальные потери при $\theta \in \mathbb{R}^d$, $x \in \mathbb{R}^d$ составляют

$$l_j(\theta,x) = \frac{1}{2}\left((\theta - \beta)^T x - \xi_j\right)^2.$$

Пусть $\{x_i\}_{i=1,\dots,n_j}$ – выборка из распределения данных для честного узла j. Общие потери узла j при $\theta \in \mathbb{R}^d$ определим формулой

$$L_j(\theta) = \nu\|\theta\|_2^2 + \frac{1}{n_j}\sum_{i=1}^{n_j} l_j(\theta,x_i). \quad (1)$$

Функция потерь (1) соответствует так называемой гребневой регрессии (регуляризация Тихонова), применяемой в системах машинного обучения и искусственного интеллекта (см. [14]).

Координация между узлами и обновление локальных моделей осуществляются путем передачи сообщений по сети и исполнения протокола честными узлами. При этом потоки информации должны удовлетворять определенным требованиям конфиденциальности (см. [7]).

Алгоритм усреднения

В этом разделе рассматривается общая схема согласования параметров честными узлами сети за счет усреднения в условиях противодействия византийских узлов.

Начнем с обозначений. Пусть для каждого узла $j \in N$ задан вектор $\rho^{(j)} \in \mathbb{R}^d$, т.е., задано отображение $\rho: N \longrightarrow \mathbb{R}^d$, и $S \subseteq N$ – некоторое подмножество множества узлов. Положим

$$diam(\vec{\rho},S) = \max_{j,k\in S} \|\rho^{(j)} - \rho^{(k)}\|.$$

Пусть $\overline{\rho}$ обозначает среднее значение семейства векторов ρ, принадлежащих честным узлам:

$$\overline{\rho} = \frac{1}{h}\sum_{j\in H}\rho^{(j)}.$$

Будем предполагать, что

$$6f + 1 \le n. \quad (2)$$

Пусть q – заранее выбранное целое число, такое что

$$\frac{n}{2} + 2f < q \le n - f. \quad (3)$$

Положим

$$\varepsilon = \frac{2q - n - 4f}{q - f}. \quad (4)$$

Несложно проверить, что

$$0 < \varepsilon < 1.$$

Далее, положим

$$C = \frac{(n + f - q)q + (q - 2f)f}{(n - f)(q - f)}. \quad (5)$$

Опишем алгоритм согласованного усреднения, который позволяет каждому честному узлу j получить вектор φ_j, удовлетворяющий следующим соотношениям:

$$diam(\vec{\varphi},H) \le (1 - \varepsilon)\cdot diam(\vec{\rho},H), \quad (6)$$

$$\|\overline{\varphi} - \overline{\rho}\| \le C \cdot diam(\vec{\rho},H). \quad (7)$$

Опишем алгоритм работы честных узлов.

Каждый честный узел j направляет остальным узлам свой вектор $\rho^{(j)}$. Честный узел i, получив q векторов, формирует множество $Q \subset N$, соответствующее отправителям полученных им векторов. В этом множестве узел i выделяет подмножество $S \subset Q$, содержащее $q - f$ элементов и дающее наиболее «компактное» семейство векторов в том смысле, что

$$S = \arg\min\{diam(\vec{\rho},S') \mid S' \subset Q, |S'| = q - f\}. \quad (8)$$

Свой вектор $\varphi^{(j)}$ узел i получает усреднением:

$$\varphi^{(j)} = \frac{1}{q - f}\sum_{j\in H}\rho^{(j)}. \quad (9)$$

Так как $Q = (Q \cap H) \cup (Q \cap F)$ и $|Q \cap F| \le f$, среди подмножеств $S' \subset Q$ мощности $q - f$ имеется хотя бы одно, содержащее только честные узлы. Следовательно,

$$diam(\vec{\rho},S) \le diam(\vec{\rho},H). \quad (10)$$

Далее, очевидно, число честных узлов в множестве S не менее, чем $q - 2f$, т.е.

$$|S \cap H| \ge q - 2f.$$

Пусть теперь i_1 и i_2 – честные узлы, а $\varphi^{(i_1)}$ и $\varphi^{(i_2)}$ – векторы, полученные усреднением. Оценим расстояние между векторами $\varphi^{(i_1)}$ и $\varphi^{(i_2)}$.

Сначала введем обозначения и сделаем необходимые предварительные оценки.

Обозначим через Q_1, Q_2, S_1, S_2 соответствующие множества узлов. Пусть $H_1 \subset S_1$ и $H_2 \subset S_2$ – множества мощности $q - 2f$, состоящие только из честных узлов. Заметим, что множества H_1 и H_2 имеют непустое пересечение. В самом деле, в силу (2) получаем:

$$|H_1 \cap H_2| = |H_1| + |H_2| - |H_1 \cup H_2| \ge 2(q - 2f) - h =$$
$$= 2q - (n + 3f) > 0.$$

Положим $F_1 = S_1 \setminus H_1$ и $F_2 = S_2 \setminus H_2$. Множества F_1 и F_2 имеют одинаковую мощность f. Пусть $v: F_1 \longrightarrow F_2$ – биективное отображение. Аналогично, обозначим через u биективное отображение $H_1 \setminus H_2$ в $H_2 \setminus H_1$. Имеем:

$$(q - f)\|\varphi^{(i_1)} - \varphi^{(i_2)}\| = \Big\|\sum_{j\in S_1}\rho^{(j)} - \sum_{j\in S_2}\rho^{(j)}\Big\| =$$
$$= \Big\|\sum_{j\in F_1}\rho^{(j)} - \sum_{j\in F_2}\rho^{(j)} + \sum_{j\in H_1}\rho^{(j)} - \sum_{j\in H_2}\rho^{(j)}\Big\| \le$$
$$\le \Big\|\sum_{j\in F_1}\rho^{(j)} - \sum_{j\in F_2}\rho^{(j)}\Big\| + \Big\|\sum_{j\in H_1}\rho^{(j)} - \sum_{j\in H_2}\rho^{(j)}\Big\| =$$
$$\le \Big\|\sum_{j\in F_1}\rho^{(j)} - \rho^{(v(j))}\Big\| + \Big\|\sum_{j\in H_1\setminus H_2}\rho^{(j)} - \rho^{(u(j))}\Big\|.$$

Далее, если $j \in F_1$, то для $j' \in H_1 \cap H_2$ в соответствии с (8) получаем

$$\left\|\rho^{(j)} - \rho^{(v(j))}\right\| \le \left\|\rho^{(j)} - \rho^{(j')}\right\| + \left\|\rho^{(j')} - \rho^{(v(j))}\right\| \le 2diam(\vec{\rho},H).$$

Отсюда

$$\left\|\sum_{j\in F_1}\rho^{(j)} - \rho^{(v(j))}\right\| \le 2f\cdot diam(\vec{\rho},H).$$

Наконец,

$$\left\|\sum_{j\in H_1\setminus H_2}\rho^{(j)} - \rho^{(u(j))}\right\| \le (n+f-q)\cdot diam(\vec{\rho},H),$$

поскольку

$$|H_1 \setminus H_2| = |H_1| \setminus |H_1 \cap H_2| \le (q-2f) - (2q-n-3f) = n+f-q.$$

Таким образом,

$$\left\|\varphi^{(i_1)} - \varphi^{(i_2)}\right\| \le \frac{n+3f-q}{q-f}\cdot diam(\vec{\rho},H).$$

Следовательно,

$$diam(\vec{\varphi},H) \le \frac{n+3f-q}{q-f}\cdot diam(\vec{\rho},H),$$

что с учетом (4) совпадает с (6).

Покажем теперь, что

$$\left\|\overline{\varphi} - \overline{\rho}\right\| \le C\cdot diam(\vec{\rho},H). \qquad (11)$$

Рассмотрим честный узел i. Пусть S_i – множество узлов мощности $q-f$, выбранных узлом i в соответствии с (4), и $S_i = H_i \cup F_i$, а множество H_i имеет мощность $q-2f$ и состоит из честных узлов.

Положим

$$\overline{\rho}' = \frac{1}{q-2f}\sum_{j\in H_i}\rho^{(j)},\ \overline{\rho}'' = \frac{1}{h-q+2f}\sum_{j\in H\setminus H_i}\rho^{(j)}$$

$$\text{и } \overline{\rho}''_- = \frac{1}{f}\sum_{j\in F_i}\rho^{(j)}.$$

Тогда

$$\left\|\varphi^{(i)} - \overline{\rho}\right\| = \left\|\frac{(q-2f)\overline{\rho}' + f\overline{\rho}''_-}{q-f} - \frac{(q-2f)\overline{\rho}' + (h-q+2f)\overline{\rho}''}{h}\right\| =$$

$$= \left\|\frac{(q-2f)\overline{\rho}' + f\overline{\rho}''_-}{(q-2f)+f} - \frac{(q-2f)\overline{\rho}' + (h-q+2f)\overline{\rho}''}{(q-2f)+(h-q+2f)}\right\| =$$

$$= \left\|\frac{(q-2f)((h-q+2f)-f)\,\overline{\rho}'}{h(q-f)+f} + \right.$$

$$+ \frac{f(q-2f)+(h-q+2f)\overline{\rho}''_-}{h(q-f)} -$$

$$\left. - \frac{(h-q+2f)((q-2f)+f)\overline{\rho}''}{h(q-f)}\right\| =$$

$$= \left\|\frac{(q-2f)(h-q+2f)(\overline{\rho}' - \overline{\rho}'')}{h(q-f)} - \right.$$

$$\left. - \frac{f(q-2f)(\overline{\rho}' - \overline{\rho}''_-)}{h(q-f)} + \frac{(h-q+2f)f(\overline{\rho}''_- - \overline{\rho}'')}{h(q-f)}\right\|.$$

Оценим по отдельности длину каждого из трех слагаемых векторов.

Имеем:

$$(q-2f)(h-q+2f)\left\|\overline{\rho}' - \overline{\rho}''\right\| =$$

$$= \left\|(h-q+2f)\sum_{j\in H_i}\rho^{(j)} - (q-2f)\sum_{j\in H\setminus H_i}\rho^{(j)}\right\| =$$

$$= \left\|\sum_{j\in H\setminus H_i}\sum_{j\in H_i}\rho^{(j)} - \sum_{j\in H_i}\sum_{j\in H\setminus H_i}\rho^{(j)}\right\| =$$

$$= \left\|\sum_{k\in H\setminus H_i}\sum_{j\in H_i}(\rho^{(j)} - \rho^{(k)})\right\| \le \sum_{k\in H\setminus H_i}\sum_{j\in H_i}\left\|(\rho^{(j)} - \rho^{(k)})\right\| \le$$

$$\le (q-2f)(h-q+2f)\cdot diam(\vec{\rho},H).$$

Таким образом,

$$\left\|\frac{(q-2f)(h-q+2f)(\overline{\rho}' - \overline{\rho}'')}{h(q-f)}\right\| \le$$

$$\le \frac{(q-2f)(h-q+2f)(\overline{\rho}' - \overline{\rho}'')}{h(q-f)}\cdot diam(\vec{\rho},H).$$

Точно так же,

$$\left\|\overline{\rho}' - \overline{\rho}''_-\right\| \le diam(\vec{\rho},H),$$

и, значит,

$$\left\|\frac{f(q-2f)(\overline{\rho}' - \overline{\rho}''_-)}{h(q-f)}\right\| \le \frac{f(q-2f)}{h(q-f)}\cdot diam(\vec{\rho},H).$$

Наконец,

$$\left\|\overline{\rho}''_- - \overline{\rho}''\right\| \le \left\|\overline{\rho}''_- - \overline{\rho}'\right\| + \left\|\overline{\rho}'' - \overline{\rho}'\right\| \le 2diam(\vec{\rho},H),$$

и, значит,

$$\left\|\frac{(h-q+2f)f(\overline{\rho}''_- - \overline{\rho}'')}{h(q-f)}\right\| \le$$

$$\le \frac{2(h-q+2f)f}{h(q-f)}\cdot diam(\vec{\rho},H).$$

В итоге получаем

$$\left\|\varphi^{(i)} - \overline{\rho}\right\| \le$$

$$\le \frac{(q-2f)(h-q+2f) + f(q-2f) + 2(h-q+2f)f}{h(q-f)} \qquad (12)$$

Несложно убедиться в том, что дробь в правой части (12) совпадает с C, заданным формулой (5). Отсюда следует (11).

Величина $C = C(q,n,f)$ убывает по q. Следовательно,

$$C(n-f,n,f) \le C < C(\frac{n}{2}+2f,n,f),$$

т.е.,

$$\frac{(3n-f)f}{(n-f)n-2f)} \le C < \frac{n^2+4fn-8f^2}{2(n-f)(n+2f)}.$$

Верхняя граница этой оценки убывает с уменьшением отношения $\frac{f}{n}$. Таким образом, при выполнении условия (2) имеем $C < \frac{13}{20}$, так что

$$\left\|\varphi^{(i)} - \overline{\rho}\right\| \le 0{,}65\cdot diam(\vec{\rho},H). \qquad (13)$$

Набор векторов $\vec{\varphi}$, полученный усреднением (9) из набора векторов $\vec{\rho}$ обозначим $\boldsymbol{A}(\vec{\rho})$. Компоненты

набора векторов $\boldsymbol{A}(\vec{\rho})$ будем обозначать $\boldsymbol{A}^{(i)}(\vec{\rho})$. В этих обозначениях для вектора, определенного в (9) имеем $\varphi^{(i)} = \boldsymbol{A}^{(i)}(\vec{\rho})$.

Результат повторного применения усреднения к набору векторов $\boldsymbol{A}(\vec{\rho})$ обозначим $\boldsymbol{A}^2(\vec{\rho})$ и т.д. В силу (6) имеем:

$$diam(\boldsymbol{A}^p(\vec{\rho}),H) < (1-\varepsilon)^p \cdot diam(\vec{\rho},H), \qquad (14)$$

так что при достаточно большом p можно гарантировать выполнение неравенства

$$diam(\boldsymbol{A}^p(\vec{\rho}),H) < \delta,$$

каково бы ни было $\delta > 0$.

Далее,

$$\left\|\overline{\boldsymbol{A}^p(\vec{\rho})} - \overline{\boldsymbol{A}^{p-1}(\vec{\rho})}\right\| < C \cdot diam(\boldsymbol{A}^{p-1}(\vec{\rho}),H).$$

При этом

$$\left\|(\overline{\boldsymbol{A}^p(\vec{\rho})} - \overline{\rho}\right\| \le \sum_{s=1}^{p}\left\|(\overline{\boldsymbol{A}^s(\vec{\rho})} - \overline{\boldsymbol{A}^{s-1}(\vec{\rho})}\right\| <$$

$$< \sum_{s=1}^{\infty} C \cdot (1-\varepsilon)^s \cdot diam(\vec{\rho},H)$$

и, значит,

$$\left\|(\overline{\boldsymbol{A}^p(\vec{\rho})} - \overline{\rho}\right\| < \frac{C}{\varepsilon} \cdot diam(\vec{\rho},H) =$$

$$= \frac{2f(q-f)+q(n-q)}{(n-f)(2q-n-4f)} \cdot diam(\vec{\rho},H).$$

Покажем, что $\frac{C}{\varepsilon}$ убывает по переменной q. В самом деле,

$$\frac{\partial}{\partial q}\frac{C}{\varepsilon} = -\frac{2q^2 - 2(n+4f)q + (n^2+6fn+4f^2)}{(n-f)(4f+n-2q)^2}, \qquad (15)$$

а при $n \ge 6f+1$ числитель дроби в правой части (15) принимает положительные значения, поскольку

$$(n+4f)^2 - 2(n^2+6fn+4f^2) < 0.$$

При $\frac{n}{2}+2f < q < n-f$ величина $\frac{C}{\varepsilon}$ убывает от $+\infty$ до

$$\left(\frac{C}{\varepsilon}\right)\Big|_{q=n-f} = \frac{3\omega(1-3\omega)}{(1-\omega)(1-6\omega)}, \qquad (16)$$

где $\omega = \frac{f}{n}$. На промежутке $0 < \omega < 1/6$ величина (16) возрастает от 0 до $+\infty$ и принимает значения меньшие, чем 1, при $\omega < \frac{5-\sqrt{10}}{15} \approx 0{,}1225$.

Построение глобальной модели гребневой линейное регрессии

Покажем, как используя усреднение, можно провести коллаборативное построение гребневой линейной регрессии на всей совокупности данных. Задача – сформировать у каждого честного узла j набор параметров $\theta_*^{(j)}$ так, чтобы $\mathrm{diam}(\vec{\theta}_*,H)$ и усредненный градиент функции потерь $\nabla\bar{L}(\bar{\theta})$ были достаточно малы.

Далее мы опишем общую схему построения и по ходу изложения сделаем необходимые уточнения.

В дополнение к сделанным предположениям будем считать, что область в $\mathbb{R}^d$, из которой выбираются значения параметров θ, ограничена (это обычное предположение для моделей федеративного и коллаборативного обучения, см. [15]).

Начнем с нескольких предварительных замечаний, связанных с оценкой градиента функции потерь.

Для $x \in \mathbb{R}^d$ имеем

$$\nabla_\theta l_j(\theta,x) = 2\nu\theta + \left((\theta-\beta_j)^T x - \xi\right)\cdot x =$$

$$= 2\nu\theta + (x^T x)(\theta-\beta_j) - \xi x.$$

Чтобы не загромождать обозначения, мы иногда опускаем указание на узел, когда ясно, о каком узле идет речь.

Соответственно

$$\nabla_\theta l_j(\theta) = 2\nu\theta + \left(\frac{1}{n_j}\sum_{i=1}^{n_j} x_i^T x_i\right)(\theta-\beta_j) - \frac{1}{n_j}\sum_{i=1}^{n_j}\xi_i x_i.$$

Положим

$$\hat{\Sigma}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} x_i^T x_i,\ \widehat{\xi x}^{(j)} = \frac{1}{n_j}\sum_{i=1}^{n_j}\xi_i x_i.$$

Покажем, что функции потерь всех узлов являются L-гладкими для некоторого $L > 0$.

В самом деле,

$$\left\|\nabla_\theta l_j(\theta') - \nabla_\theta l_j(\theta)\right\| = \left\|\hat{\Sigma}_j(\theta'-\theta)\right\| \le \left\|\hat{\Sigma}_j\right\| \cdot \left\|\theta'-\theta\right\|,$$

где $\|\ \|$ – операторная l_2-норма. Таким образом, если выбрать постоянную L так, чтобы она превосходила все собственные значения матриц $\hat{\Sigma}_j$, $j = 1,\ldots,n$, получим

$$\left\|\nabla_\theta l_j(\theta') - \nabla_\theta l_j(\theta)\right\| \le L \cdot \left\|\theta'-\theta\right\|. \qquad (17)$$

Далее,

$$\nabla_\theta l_j(\theta,x) - \nabla_\theta l_j(\theta) = (x^T x - \hat{\Sigma}_j)(\theta-\beta_j) - (\xi x - \widehat{\xi x}^{(j)}). \qquad (18)$$

Оценим слагаемые в правой части (18).

При достаточно естественных (и не слишком ограничительных) предположениях выборочная оценка матрицы Σ_j оказывается достаточно хорошей в том смысле, что операторная l_2-норма ограничена. Например, если ξ – σ-субгауссовская случайная величина, имеются такие постоянные c_1, c_2, c_3, что

$$\mathbf{P}\left[\frac{\left\|(|\Sigma_j - \hat{\Sigma}_j|)\right\|_2}{\sigma_2} \ge c_1\left\{\sqrt{\frac{d}{n_j}} + \frac{d}{n_j}\right\} + \delta\right] \le$$

$$\le c_2 \exp(-c_3\, n_j\, min\{\delta,\delta^2\}), \qquad (19)$$

для всех $\delta \ge 0$ (см. [16, 17]).

Таким образом, можно считать, что $\mathbf{E}\left\|(|\Sigma_j - \hat{\Sigma}_j|)\right\|_2$ и $\mathbf{E}\left\|(|\Sigma_j - \hat{\Sigma}_j|)\right\|_2^2$ конечны и могут быть сделаны достаточно малыми за счет выбора n_j.

Точно так же, распределение компонент вектора ξx субэспоненциально, так что аналогичное утверждение верно для $\mathbf{E}\|\xi x - \widehat{\xi x}\|$ и $\mathbf{E}\|\xi x - \widehat{\xi x}\|^2$.

Учитывая, что

$$\left\|(|(\Sigma_j - \hat{\Sigma}_j)(\theta-\beta_j)|)\right\|_2 \le \left\|(|\Sigma_j - \hat{\Sigma}_j|)\right\|_2 \cdot \left\|\theta-\beta_j\right\|,$$

а множество возможных значений параметров ограничено, имеется постоянная S, такая, что

$$\mathbf{E}\|\nabla_\theta l_j(\theta,x) - \nabla_\theta L_j(\theta)\|^2 \le S^2. \quad (20)$$

Построение регрессионной модели происходит в несколько раундов. В начальном (первом) раунде все честные узлы используют заранее установленный вектор параметров $\theta_1 \in \mathbb{R}^d$. Вектор параметров, который честный узел, формирует в раунде t и направляет другим узлам, будем обозначать $\theta_t^{(j)}$. Если из контекста ясно, о каком узле идет речь, верхний индекс будем опускать и писать просто θ_t.

Опишем работу честного узла j в раунде t. Прежде всего узел формирует пакет запросов и находит усредненное значение градиента. Для этого он делает случайную выборку векторов $z_1,\ldots,z_m$ (в доступном ему множестве данных) и по этой выборке вычисляет среднее значение локального градиента

$$g_t^{(j)} = \frac{1}{m}\sum_{i=1}^{m} \nabla_\theta l_j(\theta_t^{(j)}, x_{t,i}^{(j)}).$$

Заметим, что

$$\|g_t^{(j)} - \nabla_\theta L_j(\theta_t^{(j)})\| = \frac{1}{m}\sum_{i=1}^{m} \|\nabla_\theta l_j(\theta_t^{(j)}, x_{t,i}^{(j)}) - \nabla_\theta L_j(\theta_t^{(j)})\|.$$

Следовательно,

$$\mathbf{E}\|g_t^{(j)} - \nabla_\theta L_j(\theta_t^{(j)})\|^2 \le \frac{S^2}{2m}.$$

К локальным значениям $g_t^{(j)}$ применяется алгоритм усреднения. Алгоритм усреднения повторяется столько раз, чтобы обеспечить выполнение условия

$$diam(\boldsymbol{A}^p(\vec{g_t}),H) < \frac{1}{t}\, diam(\vec{g_t},H).$$

При этом также

$$\|\overrightarrow{\boldsymbol{A}^p(\vec{g_t})} - \overline{g_t}\| < \frac{C}{\varepsilon}\cdot diam(\vec{g_t},H).$$

Полученное усредненное значение градиента $\gamma_t^{(j)} = \boldsymbol{A}^p(\vec{g_t})$ используется для градиентного спуска. Обновленное значение параметров $\theta_{t+1}^{(j)}$ получается усреднением векторов с компонентами $\theta_t^{(j)} - \eta\gamma_t^{(j)}$, где η – скорость обучения. Таким образом,

$$\theta_{t+1}^{(j)} = \boldsymbol{A}^r(\vec{\theta}_t - \eta\vec{\gamma}_t),$$

где r выбрано так, что $(1-\varepsilon)^r < 1/2$.

Если число итераций описанных раундов достаточно велико, разброс значений параметров может быть сделан малым. Например, можно показать, что

$$diam(\vec{\theta}_t,H) \le \delta$$

при $t > 1/\delta^3$. При этом градиент среднего значения функции потерь

$$\bar{L}(\bar{\theta}_t) = \frac{1}{h}\sum_{j\in H} L^{(j)}(\bar{\theta}_t)$$

с $\cdot\ \bar{\theta}_t = \frac{1}{h}\sum_{j\in H}\theta_t^{(j)}$ также оказывается не слишком большим.

Заметим, что с учетом (17) и (20) можно воспользоваться оценками градиента $\bar{L}(\bar{\theta}_t)$ из [18].

Заключение

В статье рассмотрена задача построения модели гребневой регрессии узлами распределенной сети на распределенных данных (возможно, неоднородных). Глобальная модель строится методом градиентного спуска. Среди узлов имеются узлы с византийским поведением, противодействующие решению задачи. Показано, что коллаборативное построение модели возможно, если византийских узлов не слишком много. Описанный в статье алгоритм решает задачу, если число византийских узлов не превышает 12 %.

В эту же схему вписывается ситуация, когда некоторые узлы, добросовестно исполняющие протокол, используют отравленные данные. Вообще говоря, ситуация с византийским поведением является более сложной для достижения согласованного решения. Однако, как показано в [15], использование отравленных данных и византийское поведение при градиентном спуске оказываются эквивалентными при выполнении некоторых условий. Проверка выполнения этих условий при построении гребневой регрессии методом градиентного спуска – тема дальнейших исследований.

Литература

1. A survey on federated learning: challenges and applications / J. Wen, Z. Zhang, Y. Lan Y.[и др.] // International Journal of Machine Learning and Cybernetics 2023. №. 2(14). P. 513–535. https://doi.org/10.1007/s13042-022-01647-y.
2. Collaborative Distributed Machine Learning / D. Jin D., N. Kannengießer, S. Rank, A. Sunyaev // ACM Computing Surveys. 2024. №. 4(57). P. 1–36. https://doi.org/10.1145/3704807.
3. Model aggregation techniques in federated learning: A comprehensive survey / P. Qi, D. Chiaro, A. Guzzo [и др.] // Future Generation Computer Systems. 2024. v. 150. P. 272–293. https://doi.org/10.1016/j.future.2023.09.008.
4. Federated learning with non-iid data: A survey / Z. Lu, H. Pan, Y. Dai [и др.] // IEEE Internet of Things Journal. 2024. №. 11(11). P. 19188–19209. DOI: 10.1109/JIOT.2024.3376548.
5. Decentralized federated learning: A survey and perspective / L. Yuan, Z. Wang, L. Sun [и др.] //IEEE Internet of Things Journal. 2024. №. 21(11). P. 34617–34638. DOI: 10.1109/JIOT.2024.3407584.
6. From distributed machine learning to federated learning: A survey / J. Liu., J. Huang, Y. Zhou [и др.] // Knowledge and Information Systems. 2022. №. 4(64). P. 885–917. https://doi.org/10.1007/s10115-022-01664-x.
7. Запечников С. В. Модели и алгоритмы конфиденциального машинного обучения // Безопасность информационных технологий. 2020. №. 1(27). С. 51–67. DOI: 10.26583/bit.2020.1.05.
8. Reaching approximate agreement in the presence of faults / D. Dolev, N. A. Lynch, S. S. Pinter [и др.] // Journal of the Association for Computing Machinery (JACM). 1986. №. 3 (33). P. 499–516. https://doi.org/10.1145/5925.5931.

9. Mendes H., Herlihy M. Multidimensional approximate agreement in byzantine asynchronous systems // Proceedings of the forty-fifth annual ACM symposium on Theory of computing. 2013. P. 391-400. https://doi.org/10.1145/2488608.2488657.
10. Распределенная система обнаружения сетевых атак на основе федеративного трансферного обучения / В. И. Васильев, А. М. Вульфин, В. М. Картак [и др.] // Вопросы кибербезопасности. 2024. № 6 (64). С. 117–129. DOI: 10.21681/2311-3456-2024-6-117-129.
11. Аналитический обзор подходов к обнаружению вторжений, основанных на федеративном обучении: преимущества использования и открытые задачи / Е. С. Новикова, Е. В. Федорченко, И. В. Котенко, И. И. Холод // Информатика и автоматизация. 2023. №. 5(22). С. 1034–1082. DOI: https://doi.org/10.15622/ia.22.5.4.
12. Bracha G. Asynchronous Byzantine agreement protocols // Information and Computation. 1987. №. 2 (75). P. 130–143. https://doi.org/10.1016/0890-5401(87)90054-X
13. Методы оценки уровня разнородности данных в федеративном обучении / Е. С. Новикова, Я. Чен., А. В. Мелешко // Международная конференция по мягким вычислениям и измерениям: Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина). 2024. Т. 1. С. 447–450.
14. Theory of ridge regression estimation with applications / A.K.Md. Ehsanes Saleh, Mohamad Arashi, B.M. Golam Kibria – John Wiley & Sons, 2019. 384 p. ISBN: 978-1-118-64461-4.
15. Farhadkhani S., Guerraoui R., Villemaud O. An equivalence between data poisoning and byzantine gradient attacks // Proceedings of the 39th International Conference on Machine Learning. Proceedings of Machine Learning Research (PMLR), 2022. P. 6284–6323.
16. High-dimensional statistics: A non-asymptotic viewpoint / M. J. Wainwright – Cambridge university press, 2019. V. 48. 552 p.
17. Rigollet P., Hütter J. C. High-dimensional statistics // arXiv preprint arXiv:2310.19244. 2023. 161 p.
18. Collaborative learning in the jungle (decentralized, byzantine, heterogeneous, asynchronous and nonconvex learning) / E. M. El-Mhamdi, S. Farhadkhani, R. Guerraoui [и др.] //Advances in neural information processing systems. 2021. v. 34. p. 25044–25057.

COLLABORATIVE RIDGE REGRESSION IN A DISTRIBUTED SYSTEM WITH BYZANTINE FAILURES

Volkova E. S.[4], Gisin V. B.[5]

Keywords: *federated machine learning, Tikhonov regularization, consensus.*

Purpose of the study: *designing an algorithm for federated building a ridge regression in a distributed system with Byzantine node failures.*

Methods of research: *combining tools of high-dimensional data processing and protocols in distributed networks.*

Result(s): *the mechanism of an average agreement in an asynchronous network and an application of the average agreement for constructing a ridge regression model are described. Estimates of network parameters are given for which the algorithm of an average agreement is applicable: the distribution of data may be heterogeneous; Byzantine nodes may deviate from the execution of the network protocol in an arbitrary way; no honest node knows which of the other nodes are honest. Byzantine nodes know each other and may collude. Linear regression errors are assumed to be sub-Gaussian and independent.*

Scientific novelty: *a method has been developed to achieve an average agreement on ridge regression parameters in an asynchronous system.*

References

1. Wen, J., Zhang, Z., Lan, Y., Cui, Z., Cai, J., & Zhang, W. (2023). A survey on federated learning: challenges and applications. International Journal of Machine Learning and Cybernetics, 14(2), 513–535. https://doi.org/10.1007/s13042-022-01647-y.
2. Jin, D., Kannengießer, N., Rank, S., & Sunyaev, A. (2024). Collaborative Distributed Machine Learning. ACM Computing Surveys, 57(4), 1–36. https://doi.org/10.1145/3704807.
3. Qi, P., Chiaro, D., Guzzo, A., Ianni, M., Fortino, G., & Piccialli, F. (2024). Model aggregation techniques in federated learning: A comprehensive survey. Future Generation Computer Systems, 150, 272–293. https://doi.org/10.1016/j.future.2023.09.008.
4. Lu, Z., Pan, H., Dai, Y., Si, X., & Zhang, Y. (2024). Federated learning with non-iid data: A survey. IEEE Internet of Things Journal. 19188–19209. 10.1109/JIOT.2024.3376548.
5. Yuan, L., Wang, Z., Sun, L., Yu, P. S., & Brinton, C. G. (2024). Decentralized federated learning: A survey and perspective. IEEE Internet of Things Journal, 11(21), 34617–34638. DOI: 10.1109/JIOT.2024.3407584.
6. Liu, J., Huang, J., Zhou, Y., Li, X., Ji, S., Xiong, H., & Dou, D. (2022). From distributed machine learning to federated learning: A survey. Knowledge and Information Systems, 64(4), 885–917. https://doi.org/10.1007/s10115-022-01664-x.
7. Zapechnikov, S. V. (2020). Modeli i algoritmy konfidencial'nogo mashinnogo obucheniya // Bezopasnost' informacionnyx texnologij, 1(27), 51–67. DOI: 10.26583/bit.2020.1.05.
8. Dolev, D., Lynch, N. A., Pinter, S. S., Stark, E. W., & Weihl, W. E. (1986). Reaching approximate agreement in the presence of faults. Journal of the ACM (JACM), 33(3), 499–516. https://doi.org/10.1145/5925.5931.
9. Mendes, H., & Herlihy, M. (2013, June). Multidimensional approximate agreement in byzantine asynchronous systems. In Proceedings of the forty-fifth annual ACM symposium on Theory of computing (pp. 391–400). https://doi.org/10.1145/2488608.2488657.

4 Elena S. Volkova, Ph.D., Associate Professor, Financial University under the Government of the Russian Federation, Moscow, Russia. E-mail:evolkova@fa.ru
5 Vladimir B. Gisin, Ph.D., Professor, Financial University under the Government of the Russian Federation, Moscow, Russia. E-mail:vgisin@fa.ru

10. Vasil'ev, V. I., Vul'fin, A. M., Kartak, V. M., Bashmakov, N. M., & Kirillova, A. D. (2024). Raspredelennaya sistema obnaruzheniya setevyx atak na osnove federativnogo transfernogo obucheniya. Voprosy kiberbezopasnosti, (6), 64. S. 117–129. DOI: 10.21681/2311-3456-2024-6-117-129.
11. Novikova, E. S., Fedorchenko, E. V., Kotenko, I. V., & Xolod, I. I. (2023). Analiticheskij obzor podxodov k obnaruzheniyu vtorzhenij, osnovannyx na federativnom obuchenii: preimushhestva ispol'zovaniya i otkrytye zadachi. Informatika i avtomatizaciya, 22(5), 1034–1082. DOI: https://doi.org/10.15622/ia.22.5.4.
12. Bracha, G. (1987). Asynchronous Byzantine agreement protocols. Information and Computation, 75(2), 130-143. https://doi.org/10.1016/0890-5401(87)90054-X.
13. Novikova, E., chen, Ya., & meleshko, A. V. (2024). Metody ocenki urovnya raznorodnosti dannyx v federativnom obuchenii. In mezhdunarodnaya konferenciya po myagkim vychisleniyam i izmereniyam Uchrediteli: Sankt-Peterburgskij gosudarstvennyj elektrotexnicheskij universitet «LETI» im. V. I. Ul'yanova (Lenina) (Vol. 1, pp. 447–450).
14. Theory of ridge regression estimation with applications / A. K. Md. Ehsanes Saleh, Mohamad Arashi, B. M. Golam Kibria – John Wiley & Sons, 2019. 384 p. ISBN: 978-1-118-64461-4.
15. Farhadkhani, S., Guerraoui, R., & Villemaud, O. (2022, June). An equivalence between data poisoning and byzantine gradient attacks. In International Conference on Machine Learning (pp. 6284–6323). PMLR.
16. Wainwright, M. J. (2019). High-dimensional statistics. Cambridge university press, 552 p.
17. Rigollet, P., & Hütter, J. C. (2023). High-dimensional statistics. arXiv preprint arXiv:2310.19244. 161 p.
18. El-Mhamdi, E. M., Farhadkhani, S., Guerraoui, R., Guirguis, A., Hoang, L. N., & Rouault, S. (2021). Collaborative learning in the jungle (decentralized, byzantine, heterogeneous, asynchronous and nonconvex learning). Advances in neural information processing systems, 34, 25044–25057.

ОБЪЯСНИМАЯ ИНТЕРПРЕТАЦИЯ ИНЦИДЕНТОВ НА ОСНОВЕ БОЛЬШОЙ ЯЗЫКОВОЙ МОДЕЛИ И МЕТОДА ГЕНЕРАЦИИ С ДОПОЛНЕННОЙ ВЫБОРКОЙ

Котенко И. В.[1], Абраменко Г. Т.[2]

DOI: 10.21681/2311-3456-2025-5-58-67

Цель исследования: *повысить достоверность и объяснимость интерпретации оповещений системы обнаружения и предотвращения вторжений Suricata за счет онтологически обогащенного графа знаний, гетерогенных графовых представлений и метода генерации с дополненной выборкой (Retrieval-Augmented Generation, RAG) на основе локальной большой языковой модели (Large Language Model, LLM).*

Методы исследования: *построение онтологически управляемого графа знаний, связывающего данные Suricata с тактиками/техниками MITRE ATT&CK; обучение гетерогенной графовой нейросети (HGNN) для получения контекстных векторных представлений узлов; извлечение релевантного контекста по ближайшим соседям в пространстве эмбеддингов; генерация объяснений через локальную LLM (с 7B параметрами) по RAG-конвейеру; экспериментальная оценка на корпусе из 25000 оповещений Suricata с использованием метрик точности интерпретации, доли галлюцинаций и релевантности.*

Полученные результаты: *разработан онтологически управляемый метод интерпретации оповещений Suricata, обеспечивающий более полное и корректное объяснение по сравнению с базовым подходом. Показано, что использование онтологий позволяет повысить содержательность объяснений на 15 % при несущественном увеличении времени генерации. В результате, интеграция онтологии и гетерогенного графа знаний существенно повышает корректность сопоставления оповещений об инцидентах безопасности с техниками MITRE ATT&CK и снижает риск неверных объяснений.*

Научная новизна: *предложена интеграция онтологии и гетерогенного графа знаний с RAG-генерацией поверх локальной LLM для привязки низкоуровневых событий системы обнаружения и предотвращения вторжений к техникам MITRE ATT&CK; показана применимость онтологий к задачам объяснимого ИИ для кибербезопасности.*

Ключевые слова: *кибербезопасность; системы обнаружения и предотвращения вторжений; поиск угроз безопасности; объяснимый искусственный интеллект; граф знаний; глубокое обучение: гетерогенная графовая нейросеть; MITRE ATT&CK; LLM; RAG; Suricata.*

1. Введение

Современные центры мониторинга кибербезопасности (Security Operations Center, SOC) ежедневно сталкиваются с лавинообразным потоком сетевых событий и оповещений систем обнаружения и предотвращения вторжений (Intrusion Detection and Prevention System, IDPS), таких как Suricata[3]. Значительная часть этих оповещений оказывается нерелевантной или малоинформативной для оперативного реагирования, а попытка вручную соотнести низкоуровневые сигнатуры с высокоуровневыми сценариями атак приводит к перегрузке аналитиков и задержкам при принятии решений [1]. Таким образом, повышается актуальность методов, которые не только выявляют подозрительную активность, но и интерпретируют смысл каждого срабатывания в терминах тактик, техник и процедур злоумышленника, понятных специалисту SOC.

Попытка напрямую использовать большие языковые модели (Large Language Model, LLM) для построения таких объяснений выглядит естественной: они умеют излагать сложные технические детали на естественном языке. Однако LLM склонны к фактологическим ошибкам и «галлюцинациям», что в критичных сценариях кибербезопасности недопустимо [2, 3].

Для решения указанных рисков используется генерация с дополненной выборкой (Retrieval-Augmented Generation, RAG), когда языковая модель получает заранее отобранные релевантные сведения из внешнего источника знаний и генерирует формулировки, используя эти сведения [4]. В гибридных архитектурах к RAG добавляются структурированные представления и графовые нейронные сети (Graph Neural Network, GNN), что повышает точность и воспроизводимость рассуждений [5].

1 Котенко Игорь Витальевич, доктор технических наук, профессор, заслуженный деятель науки РФ, главный научный сотрудник и руководитель лаборатории проблем компьютерной безопасности, ФГБУН «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (СПб ФИЦ РАН), г. Санкт-Петербург, Россия. E mail: ivkote@comsec.spb.ru

2 Абраменко Георгий Тимофеевич, аспирант факультета безопасности информационных технологий, ФГАОУ ВО «Национальный исследовательский университет ИТМО», г. Санкт-Петербург, Россия. E-mail: gtabramenko@itmo.ru

3 Suricata – Intrusion Detection and Prevention System. https://suricata.io

Наиболее перспективным источником такого знания для задач SOC являются формальные онтологии и графы знаний (Knowledge Graph, KG) по кибербезопасности (например, MITRE ATT&CK, CAPEC и др.)[4]. Использование подхода KG-Infused RAG [6] позволит связать оповещения IDPS с тактиками и техниками атак, индикаторами компрометации, уязвимостями и инструментами [7, 8] на основе стандартизированного словарь и практик проактивного поиска угроз [9]. В [10] демонстрируются промышленные решения корреляции инцидентов с учетом миллиардов событий, опирающиеся на графовые структуры.

Вместе с тем остается недостаточно изученным вопрос, связанный с тем, насколько гетерогенная структура графа, включающая различные типы узлов и связей (например, узлы типа Оповещение, Сигнатура, Техника, Индикатор Alert, Signature, Technique, Indicator и т.д.) и явное онтологическое ядро знаний (ATT&CK/CAPEC) улучшают качество именно интерпретаций, а не только обнаружения. Это особенно важно для оповещений IDPS в реальных потоках данных [11, 12], тогда как формализованное моделирование угроз на базе MITRE ATT&CK[5] влияет на пригодность объяснений для аудита и действий по реагированию на инциденты [13].

В данной работе развивается направление, связанное с обозначенными перспективными направлениями в исследованиях, и предлагается онтологически управляемый метод интерпретации оповещений Suricata на основе гетерогенного графа знаний и LLM с использованием RAG. Ключевая идея – представить каждое оповещение как узел графа, связанный с узлами сигнатур, техник и тактик из онтологии атак; далее обученная гетерогенная графовая нейронная сеть (Heterogeneous GNN, HGNN) вычисляет векторное представление события и извлекает ближайший контекст из графа, который затем используется LLM для достоверного объяснения. Такой подход обеспечивает отслеживаемость утверждений до конкретных сущностей базы знаний, снижение доли вымышленных фактов, повышение полезности текста для принятия решений аналитиком SOC.

2. Анализ релевантных работ

Проблема устранения избыточных и ложных оповещений, а также приоритизации инцидентов в SOC является актуальным направлением исследований в области кибербезопасности. Так, в [1] приводятся данные исследования SOC, в котором отмечается, что в проанализированных предприятиях обрабатывается свыше 2,4 млн оповещений ежедневно, причем до 70 % из них являются ложными. Предлагаются подходы к интеллектуальной приоритизации оповещений с помощью больших языковых моделей, интегрированных в системы фильтрации, оценки и категоризации событий.

Некоторые публикации фокусируются на объяснимости интеллектуальных моделей для классификации сетевых атак. В частности, в [2] оценивались методы объяснимого искусственного интеллекта (ИИ) (Explainable AI, XAI), и было указано, что недостаток прозрачности мешает практическому применению моделей машинного обучения (Machine Learning, ML) для обнаружения угроз.

В [14] представлена диалоговая система, функционирующая на основе агрегированных баз знаний по уязвимостям и инцидентам, в которой сначала извлекаются подтвержденные факты, а затем формируется текстовое пояснение. Такой подход демонстрирует рост точности ответов при типичных запросах аналитика. В [15] источником контекста служат одновременно два графа: внутренний граф событий и журналов организации и внешний граф знаний по кибербезопасности. Из каждого графа выбираются согласованные фрагменты, и уже на их основе формируются пояснения.

В [16] для задач реагирования на инциденты предлагается подход, где структурированные запросы к внешним источникам сведений об угрозах сочетаются с семантическим поиском по локальным материалам. В результате реализации такого подхода формируется контекст, который используется для выдачи конкретных рекомендаций по реагированию на инциденты.

Отдельно развивается направление объединения графовых моделей с генеративным ИИ. В [17] представлена технология корреляции событий безопасности, основанная на графах знаний, в которых события связываются через уязвимости, активы и тактики атак, что обеспечивает выявление сложных атак. Для анализа данных из журналов безопасности предложены методы автоматизированного построения графов [18] и виртуальные графы знаний для анализа логов из разнородных источников [19]. Такой подход снижает затраты на предварительную интеграцию данных и повышает пригодность данных для последующего анализа. Жизнеспособность графовых структур подтверждается в системах корреляции инцидентов, работающих с большими объемами событий [10].

Заметный прогресс в исследованиях по системам обнаружения вторжений связан с переходом от классических признаков к графовым представлениям. В [20, 21] показано, что явный учет связей между объектами инфраструктуры и событиями повышает

4 Syed Z., Padia A., Finin T., Mathews L., Joshi A. UCO: A Unified Cybersecurity Ontology // Proc. AAAI Workshop on Artificial Intelligence for Cyber Security. 2016. 8 p.

5 Strom, B.; et al. MITRE ATT&CK: Design and Philosophy. MITRE Technical Report, 2018.

полноту и устойчивость обнаружения сложных атак. Также развиваются методы реализации объяснимости графовых моделей. Такие методы позволяют определять, какие вершины графа и какие ребра повлияли на итоговое решение. Кроме того, вводятся количественные показатели устойчивости объяснений [22]: проверяется, сохраняется ли интерпретация при небольших изменениях входных данных и структуры графа, а также при изменении статистики данных между обучением и использованием.

При реализации подходов, основанных на RAG, складывается общее понимание, что качество модуля поиска, полнота охвата и согласованность найденных фактов определяют применимость и полезность итогового ответа. В [5, 23, 24] предлагаются решения, в которых сам граф знаний направляет поиск и задает допустимые цепочки рассуждений, что дает возможность повысить согласованность и уменьшить риск вымышленных утверждений.

Подобные системы демонстрируют, что интеграция внешних знаний способна уменьшить галлюцинации моделей и повысить точность объяснений. Для объяснения оповещений IDPS используются подходы, основанные на комбинировании гетерогенных графов и LLM. Так в [25] представлено использование XG-NID с HGNN и больших языковых моделей для объяснения сетевых вторжений, а также предлагается подход к интерпретации решений GNN в кибербезопасности.

В отечественных исследованиях последних лет фиксируется устойчивый тренд применения методов искусственного интеллекта к задачам кибербезопасности. Развиваются подходы к федеративному обучению для систем обнаружения вторжений, например, в [26] представлена архитектура федеративной IDS/IDPS с экспериментальной валидацией на репрезентативных данных. Отдельное направление составляет атрибуция целевых атак и профилирование нарушителей, например, в [27] систематизированы модели, источники данных и методики, определены ограничения и требования к системам атрибуции атак. Параллельно формируется графовый аппарат оценки защищенности предложена модель оценки на графе эксплойтов, обеспечивающая количественные показатели для принятия решений [28]. Развитие идей федеративного обучения демонстрирует предложенная в [29] распределенная система обнаружения сетевых атак, сочетающая федеративное и трансферное обучение. Для систем обнаружения вторжений в кибер-физических системах обобщены источники данных, методы и метрики, выявлены ограничения наборов данных и типовые ошибки постановки экспериментов [30]. Для сетей Интернета вещей применяются многозадачное обучение и гибридные методы формирования выборок, повышающие полноту при выраженном дисбалансе классов [31]. Для автоматизации и приоритизации реагирования на инциденты разработаны методы прогнозирования категорий уязвимостей по конфигурациям устройств, что позволяет подготовить признаки для ранжирования задач реагирования на инциденты [32].

Обобщая рассмотренные исследования, следует отметить, что эффективная интерпретация оповещений IDPS требует использования формализованных знаний (онтологии и графы знаний). Применение гетерогенных графовых моделей для представления разнотипных сущностей и связей, а также использование RAG-подходов для обеспечения текстовых объяснений проверяемых фактов, извлеченных из внешних баз знаний, потенциально сможет уменьшить вероятность недостоверных утверждений и повысить практическую ценность выводов для аналитиков SOC. Вместе с тем недостаточно проработанными остаются вопросы построения интегрированного конвейера, в рамках которого оповещения для IDPS систем системно встраиваются в онтологически обогащенный гетерогенный граф и далее используются в качестве внешнего хранилища знаний для языковой модели при формировании кратких и проверяемых пояснений по каждому оповещению.

3. Варианты архитектур системы объяснения оповещений

В статье представлено три варианта предлагаемой архитектуры системы объяснения оповещений IDPS, основанных на комбинировании поиска информации на графе знаний с генеративными возможностями LLM.

Каждая из архитектур представлена на (рис. 1).

Сопоставление трех архитектур, представленное в (табл. 1), показывает, что усложнение модели памяти от однородного графа к гетерогенному и далее к онтологически управляемому графу последовательно снижает риск шумовых (ложных) соответствий и повышает «проверяемость» объяснений. Это выражается в переходе от выбора единственного «наиболее важного» пути (GNN-RAG) к формированию согласованного подграфа доказательств (HGNN-RAG) и, наконец, к отбору контекста под жесткими онтологическими ограничениями с трассировкой ссылок на узлы ATT&CK/CAPEC (Ontology-KG-HGNN-RAG).

При этом возрастает стоимость подготовки данных: если базовому варианту достаточно минимальной связности KG, то онтологически управляемый вариант требует поддержания соответствий сигнатур и событий элементам онтологии и регулярного обновления отображений (маппингов).

С учетом табл. 1 можно сформулировать следующие выводы.

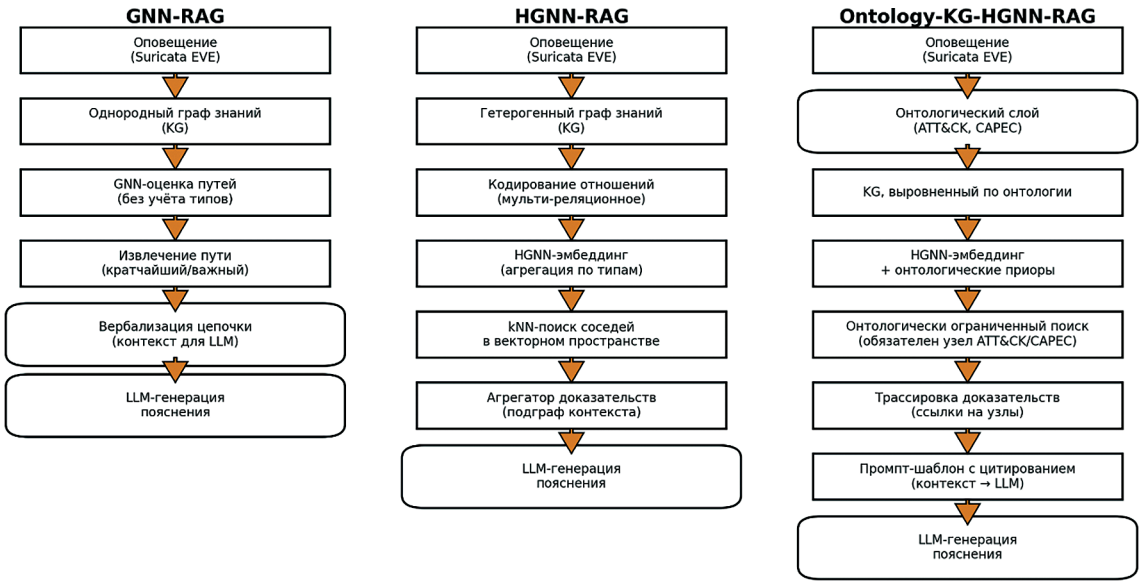


Рис. 1. Архитектуры GNN-RAG, HGNN-RAG, Ontology-KG-HGNN-RAG

Во-первых, GNN-RAG целесообразен как стартовый или резервный режим: он прост в развертывании и дает минимально необходимую интерпретируемость, но чувствителен к косвенным связям и потому хуже подходит для отчетных сценариев SOC.

Во-вторых, HGNN-RAG является рабочим компромиссом: он учитывает природу сущностей и взаимосвязей, формирует более чистый контекст и обеспечивает устойчивый прирост качества без радикального роста операционных затрат.

В-третьих, Ontology-KG-HGNN-RAG приоритетен для сред с повышенными требованиями для аудита: он обеспечивает комплекс ссылок на онтологию и минимизирует недостоверные утверждения, однако требует строгости во ведении онтологического слоя и подтверждения маппингов.

С точки зрения эксплуатационных характеристик, рост точности и релевантности при переходе к онтологически управляемой архитектуре сопровождается увеличением стоимости предлагаемого решения. Необходимо поддерживать онтологическую базу (обновления ATT&CK/CAPEC, ревизии соответствий сигнатур) в актуальном состоянии, а также периодически переобучать HGNN для учета новых типов

Таблица 1.

Сопоставление архитектур по ключевым характеристикам

Характеристика	GNN-RAG	HGNN-RAG	Ontology-KG-HGNN-RAG
Тип графа	Однородный KG	Гетерогенный KG (типы узлов/связей)	Гетерогенный KG + явное онтологическое ядро
Выбор контекста	Кратчайший/значимый путь; kNN	Мульти-реляционный агрегатор; kNN	Онтологически ограниченный поиск; kNN
Прослеживаемость к ATT&CK/CAPEC	Косвенная	Частичная	Гарантированная (узлы-опоры в каждой цепочке)
Риск шумовых связей	Повышенный	Сниженный	Минимизирован за счет онтологических ограничений
Ожидаемая доля «галлюцинаций» LLM	Выше	Ниже	Наименьшая
Практический вывод	Базовая интерпретация	Семантически богаче	Наиболее строгие и точные объяснения

связей. В обмен на эти затраты формируются объяснения, пригодные для требуемой отчетности и повторной проверки: каждое утверждение привязывается к явно идентифицируемым узлам знания, а отбор контекста является воспроизводимым.

4. Эксперименты

4.1. Описание экспериментов

Эксперименты выполнялись на воспроизводимом стенде SOC. В качестве сетевой основы использовался корпус CIC-IDS2017[6] после его обработки на Suricata (Emerging Threats, релиз 2024.1). Получено около 25000 срабатываний примерно по пяти десяткам сигнатур. Чтобы сравнение трех архитектур (и соответствующих конвейеров) было однозначным, было отобрано 30 репрезентативных правил, и для одних и тех же событий формировались объяснения на основе использования трех выбранных подходов (GNN-RAG, HGNN-RAG, Ontology-KG-HGNN-RAG) – всего 90 текстов.

Ниже приведен фрагмент EVE-события, на основе которого строится граф и формируется контекст для LLM:

```
{
   "timestamp": "2025-05-26T14:07:31.342Z",
   "src_ip": "10.24.5.12", "src_port": 53342,
   "dest_ip": "8.8.8.8",  "dest_port": 53,
   "proto": "UDP",
   "alert": {
      "signature_id": 1234567,
      "signature": "ET DNS Possible DNS
Tunneling",
      "category": "A Network Trojan was
detected",
      "severity": 2
   }
}
```

Сопоставление полей EVE с онтологией/графом выполнено так: Signature → Technique (ATT&CK T1071), Technique → Tactic (Command-and-Control), Event → Indicator (DNS/53-UDP).

4.2. Архитектура конвейера и параметры эксперимента

Все три конвейера имеют одинаковый режим обучения графовой части: двухслойная модель с размерностью представлений $d = 64$ и контрастивным обучением на парах «позитив/негатив» (позитивы – вручную размеченные связи «Событие → Техника» и «Событие → ПО/инструмент», негативы – случайные несвязанные пары). В процессе вывода для узла «событие» выбираются $k = 3$ ближайших факта из графа по косинусному сходству (формула (1)) и из них собирается текстовый контекст для LLM.

$$sim(u,v) = \frac{(h_v^T h_u)}{\|h_v\|_2 \cdot \|h_u\|_2};\ sim \in [-1;1], \qquad (1)$$

где: u – узел «событие», v – узел «факт»; h_v, $h_u \in R^d$ – их векторные представления; $\| \ \|_2$ – евклидова норма.

Чем значение sim ближе к 1, тем больше доверие, что событие и найденный факт описывают один и тот же контекст; по убыванию sim берутся k лучших соответствий.

Во второй архитектуре (HGNN-RAG) эмбеддинги считаются с учетом типов ребер: вклад переходов «Signature → Technique» и «Technique → Tactic» различается. Обновление представлений задается формулой (2), где сумма берется по всем типам $r \in R$ и по соседям соответствующего типа:

$$h_i^{(l+1)} = ReLU\left(W_{self}\, h_i^{(l)} + \sum_{r \in R} \sum_{j \in N_r(i)} \frac{1}{c_{i,r}} W_r h_j^{(l)} \right), \qquad (2)$$

где: i – узел, l – номер слоя; R – типы связей (например, «Сигнатура → Техника», «Техника → Тактика», «Событие → Индикатор»); N_r – соседи i по связи типа r; W_{self}, $W_r \in \mathbb{R}^{d \times d}$ – обучаемые матрицы; $c_{i,r} = \max(1, |N_r(i)|)$ – нормировка вклада «плотных» узлов.

Учет типов связей через W_r уменьшает число неверных соответствий по сравнению с однородным графом.

В онтологически управляемом конвейере (Ontology-KG-HGNN-RAG) вводится правило: каждая цепочка обоснования должна проходить через узел ATT&CK или CAPEC. Это обеспечивает проверяемость утверждений и устраняет выводы вне фактического контекста события.

Генерация выполнялась на локальной LLM (с примерно 7B параметрами); среднее время ответа 3–5 с (поиск контекста + генерация).

4.3. Метрики и процедура оценки

Оценка проводилась на одном и том же множестве событий, что обеспечивает сопоставимость трех модификаций конвейера. Эталонная разметка основной техники выполнялась независимыми экспертами центра мониторинга безопасности; для исключения переобучения использовалась отложенная выборка. Порог включения факта в контекст – фиксирован и одинаков для всех вариантов.

Ключевая метрика (основная техника):

$$\text{Precision@1} = \frac{1}{N} \sum_{i=1}^{N} 1\{\hat{t}_i = t_i^*\}, \qquad (3)$$

где N – число проверяемых событий; $\hat{t}_i$ – идентификатор основной техники ATT&CK, указанной в объяснении; t_i^* – эталон; $1\{\}$ – индикаторная функция.

Доверительные интервалы для средних экспертных оценок считались по стандартной формуле t-интервала; согласованность мнений проверялась коэффициентом координации Кендэлла W.

6 https://www.kaggle.com/datasets/chethuhn/network-intrusion-dataset

Таблица 2.

Качество генерируемых объяснений на 30 сигнатурах (90 текстов)

Метрика / Подход	GNN-RAG	HGNN-RAG	Ontology-KG-HGNN-RAG
Precision@1 (техника)	0,60	0,73	0,83
Recall (ключевые факты)	0,65	0,80	0,92
Длина объяснения, слов	33 ± 7	41 ± 6	46 ± 5
BLEU-2 (10 эталонов)	0,47	0,53	0,58
Корректность (эксперты)	3,4 ± 0,3	4,2 ± 0,2	4,8 ± 0,1
Полнота (эксперты)	3,1 ± 0,4	3,9 ± 0,3	4,6 ± 0,2
Ясность (эксперты)	3,7 ± 0,3	4,4 ± 0,2	4,8 ± 0,1
Полезность (эксперты)	3,0 ± 0,4	4,1 ± 0,3	4,7 ± 0,2

4.4. Результаты экспериментов

Сопоставление трех модификаций конвейера на одном и том же множестве событий показало устойчивое преимущество гетерогенного графа с онтологическим ограничением. По ключевой метрике Precision@1 этот вариант уверенно превосходит однородный граф и опережает «просто гетерогенный»; прирост качества сопровождается умеренным увеличением времени ответа, что приемлемо для практической эксплуатации центра мониторинга безопасности.

Сводные результаты представлены в табл. 2.

Показатель BLEU-2 растет в пользу вариантов с учетом типов связей, что отражает лучшее соответствие эталонным описаниям. Экспертные оценки пяти аналитиков SOC демонстрируют согласованное преимущество по «Корректности», «Полноте», «Ясности» и «Полезности».

GNN-RAG обеспечивает наименьшее время обработки, но склонен к «общим» пояснениям без четкой привязки к технике, а HGNN-RAG снижает долю неверных привязок, но при скудном контексте все еще может «останавливаться» на уровне тактики. Ontology-KG-HGNN-RAG повышает проверяемость выводов: каждое ключевое утверждение проходит через узел онтологии, а трассировка по идентификаторам в тексте объяснения позволяет быстро сверить вывод с первоисточником.

4.5. Анализ результатов и типичные ошибки

Базовый GNN-RAG дал корректную, но расплывчатую формулировку без привязки к ATT&CK. HGNN-RAG точно указал на T1568.002 (DGA). Ontology-KG-HGNN-RAG дополнительно связал событие с тактикой Command-and-Control и предложил проверку хоста на вредоносные DGA – эксперты признали этот вариант наиболее полезным.

Чтобы проиллюстрировать различия в поведении подходов, приведем два характерных примера.

ET SCAN SMB1. В базовом варианте встречались выводы «сверх контекста» (например, упоминание потенциального удаления файлов без опоры в графе). В онтологической версии такие домыслы исчезли, поскольку отбор контекста обязан проходить через узлы ATT&CK/CAPEC.

Сигнатура с недостаточным контекстом «ET POLICY PE EXE or DLL Windows file download». Срабатывание фиксирует загрузку PE по HTTP/HTTPS и само по себе не доказывает атаку. Однородный и гетерогенный графы при дефиците данных обычно выдают тактику без конкретной техники. Онтологически управляемый вариант принудительно проводит контекст через MITRE ATT&CK и дает T1105 («Передача инструментария») с проверяемыми признаками. Итоговое объяснение содержит технику и обоснование, пригодное для эскалации.

Основные источники ошибок. Основные источники ошибок во всех вариантах связаны с дефицитом контекста и неполнотой соответствий «сигнатура → техника». На практике это устранимо регулярной актуализацией справочника соответствий, обогащением графа проверенными связями и явным правилом «нет достаточных оснований – техника не указывается», чтобы избегать повторной интерпретации.

4.6. Ограничения и практические замечания

Качество результатов зависит от полноты и актуальности онтологического слоя. При появлении новых техник и описаний атак требуется оперативное обновление соответствий и переиндексация графа. В противном случае верхняя планка качества ограничивается отсутствием опорных узлов. Источники данных также накладывают ограничения: часть событий получена на основе открытого набора трафика и может отличаться от специфики конкретной организации. Для переноса в промышленную среду полезно дополнительно проверять систему на реальных журналах предприятия.

По ресурсам система остается пригодной для повседневной работы SOC: рост времени ответа при переходе к онтологическому варианту является умеренным, однако при масштабировании на значительно большие графы потребуется инженерная оптимизация — приближенные структуры для поиска ближайших соседей, шардирование графа и предфильтрация по приоритетам событий.

С точки зрения эксплуатации важно обеспечить трассируемость: в каждое объяснение включаются идентификаторы ATT&CK/CAPEC/SID, а интерфейс должен позволять переход к источнику соответствия. Рекомендуется регламентировать пороги включения фактов, периодичность обновления онтологии и порядок валидации изменений, а также хранить версионированные слои знаний для воспроизводимости результатов. В перспективе предполагается расширить проверку на других наборах открытого трафика и на реальных потоках из производственной сети, а также изучить влияние адаптивных порогов по классам событий.

Заключение

В работе предложен подход к объяснению оповещений IDPS (на примере Suricata) с использованием графовых знаний и больших языковых моделей.

Проведенное исследование позволяет сделать следующие выводы:

- Онтологии кибербезопасности (в частности, MITRE ATT&CK, CAPEC, STIX) содержат ценную информацию для интерпретации оповещений, генерируемых при мониторинге инцидентов безопасности. Их интеграция в процесс объяснения позволяет увязать разрозненные низкоуровневые события с общей схемой реализации атаки, что повышает осведомленность аналитика и качество решений.
- Разработан подход Ontology-KG-HGNN-RAG, сочетающий гетерогенную графовую нейронную сеть и онтологически обогащенный граф знаний для извлечения релевантного контекста, а также генеративную модель для формирования текстового объяснения. Данный подход превосходит базовый вариант (без учета типов узлов и онтологий) по полноте и корректности генерируемых объяснений примерно на 15–20 % по ключевым метрикам, а также получил более высокие оценки экспертов-практиков (на 1–1,5 балла по 5-балльной шкале).
- Показана реализуемость автоматической системы объяснений в условиях, приближенных к реальным: использованы открытые данные трафика и актуальные базы знаний. Таким образом, технология может найти применение на практике в составе систем оркестрации, автоматизации и реагирования на инциденты информационной безопасности (SOAR, Security Orchestration, Automation and Response) или систем управления событиями и информацией безопасности нового поколения (NG-SIEM, Next-Generation Security Information and Event Management).
- Ограничения метода включают зависимость от полноты базы знаний и необходимость поддержки в актуальном состоянии, а также ресурсоемкость при масштабировании на очень большие потоки. Требуется дальнейшая проработка механизмов фильтрации и реализации распределенных вычислений. Кроме того, остаются нерешенными вопросы интеграции оператора в цикл принятия решения – как максимально удобно представить объяснение и учесть обратную связь.
- Результаты свидетельствуют о перспективности объединения графовых знаний и LLM для задач кибербезопасности. Данный подход может быть расширен на смежные проблемы, включая объяснение решений систем обнаружения аномалий, обогащение уведомлений SIEM, автоматизированное написание отчетов об инцидентах на основе сырых логов и др.

В будущем планируется развивать данное направление, в том числе исследовать возможность адаптации модели к русскоязычным онтологиям, а также реализовать прототип плагина к популярным SIEM-системам, который на лету будет выводить «объяснительные подсказки» для каждого предупреждения. Предложенный подход может послужить основой для нового поколения интеллектуальных помощников аналитика SOC, совмещающих возможности ИИ и прозрачность знаний.

Благодарность. Работа выполнена при частичной финансовой поддержке бюджетной темы FFZF-2025-0016.

Рецензент: Липатников Валерий Алексеевич, доктор технических наук, профессор, заслуженный деятель науки Российской Федерации, старший научный сотрудник научно-исследовательского центра Военной академии связи имени Маршала Советского Союза С. М. Буденного, Санкт-Петербург, Россия. E-mail: lipatnikovanl@mail.ru

Литература

1. Singh A. Contextual Threat Intelligence and Alert Prioritization with Foundation-Sec-8B // International Journal of Artificial Intelligence Research and Development. 2025. Vol. 3, No. 1. P. 131–145. DOI: 10.34218/IJAIRD_03_01_009.
2. Arreche O., Guntur T., Abdallah M. XAI-IDS: Toward Proposing an Explainable Artificial Intelligence Framework for Enhancing Network Intrusion Detection Systems // Applied Sciences. 2024. Vol. 14, No. 10. Article 4170. DOI: 10.3390/app14104170.
3. Hassanin M., Moustafa N. A Comprehensive Overview of Large Language Models (LLMs) for Cyber Defences: Opportunities and Directions // arXiv preprint arXiv:2405.14487. 2024. DOI: 10.48550/arXiv.2405.14487.
4. Fan W., Ding Y., Ning L., Wang S., Li H., Yin D., Chua T.-S., Li Q. A Survey on RAG Meeting LLMs: Towards Retrieval-Augmented Large Language Models // Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'24). 2024. P. 6491–6501.
5. Mavromatis C., Karypis G. GNN-RAG: Graph Neural Retrieval for Large Language Model Reasoning // arXiv preprint arXiv:2405.20139. 2024. DOI: 10.48550/arXiv.2405.20139.
6. Wu D., Yan Y., Liu Z., Liu Z., Sun M. KG-Infused RAG: Augmenting Corpus-Based RAG with External Knowledge Graphs // arXiv preprint arXiv:2506.09542. 2025. DOI: 10.48550/arXiv.2506.09542.
7. Al-Sada B., Sadighian A., Oligeri G. MITRE ATT&CK: State of the Art and Way Forward // ACM Computing Surveys. 2024. Vol. 57, No. 1. Article 12. P. 1–37. DOI: 10.1145/3687300.
8. Li H., Shi Z., Pan C., Zhao D., Sun N. Cybersecurity Knowledge Graphs Construction and Quality Assessment // Complex & Intelligent Systems. 2024. Vol. 10. P. 1201–1217. DOI: 10.1007/s40747-023-01205-1.
9. Hemberg E., Kelly J., Shlapentokh-Rothman M., Reinstadler B., Xu K., Rutar N., O'Reilly U.-M. Linking Threat Tactics, Techniques, and Patterns with Defensive Weaknesses, Vulnerabilities and Platforms for Cyber Hunting // arXiv preprint arXiv:2010.00533. 2021. DOI: 10.48550/arXiv.2010.00533.
10. Freitas S., Gharib A. GraphWeaver: Billion-Scale Cybersecurity Incident Correlation // Proceedings of the 33rd ACM International Conference on Information and Knowledge Management (CIKM'24). 2024. P. 4479–4486. DOI: 10.1145/3627673.3680057.
11. Sikos L. F. Cybersecurity Knowledge Graphs // Knowledge and Information Systems. 2023. Vol. 65, No. 9. P. 3511–3531. DOI: 10.1007/s10115-023-01860-3.
12. Zhao X., Jiang R., Han Y., Li A., Peng Z. A Survey on Cybersecurity Knowledge Graph Construction // Computers & Security. 2024. Vol. 136. Article 103524. DOI: 10.1016/j.cose.2023.103524.
13. Xiong W., Legrand E., Åberg O., Lagerström R. Cyber Security Threat Modeling Based on the MITRE Enterprise ATT&CK Matrix // Software and Systems Modeling. 2022. Vol. 21. P. 157–177. DOI: 10.1007/s10270-021-00898-7.
14. Arikkat D. R., Abhinav M., Binu N., Rajendran N., Bhattacharjee D., Das R. K., Ajay C. R., Shiyam Sundar R., Bhavesh S. IntelBot: A Retrieval-Augmented Large Language Model Chatbot for Cyber Threat Knowledge Delivery // arXiv preprint arXiv:2411.05442. 2024. DOI: 10.48550/arXiv.2411.05442.
15. Kurniawan K., Kiesling E., Ekelhart A. CyKG-RAG: Towards Knowledge-Graph Enhanced Retrieval-Augmented Generation for Cybersecurity // CEUR Workshop Proceedings. 2024. Vol. 3950. P. 51–64.
16. Tellache A., Amara-Korba A., Mokhtari A., Moldovan H., Ghamri-Doudane Y. Advancing Autonomous Incident Response: Leveraging LLMs and Cyber Threat Intelligence // arXiv preprint arXiv:2508.10677. 2025. DOI: 10.48550/arXiv.2508.10677.
17. Qi Y., Gu Z., Li A., Zhang X., Shafiq M., Mei Y., Lin K. Cybersecurity Knowledge Graph Enabled Attack Chain Detection for Cyber-Physical Systems // Computers & Electrical Engineering. 2023. Vol. 108. Art. 108660. DOI: 10.1016/j.compeleceng.2023.108660.
18. Ekelhart A., Ekaputra F.J., Kiesling E. SLOGERT: Automated Log Knowledge Graph Construction // The Semantic Web – ESWC 2021. Lecture Notes in Computer Science. 2021. Vol. 12731. P. 219–234. DOI: 10.1007/978-3-030-77385-4_16.
19. Kurniawan K., Ekelhart A., Kiesling E., Winkler D., Quirchmayr G., Tjoa A.M. Virtual Knowledge Graphs for Federated Log Analysis // Proceedings of the 16th International Conference on Availability, Reliability and Security (ARES'21). 2021. Article 50. 10 p. DOI: 10.1145/3465481.3470077.
20. Bilot T., El-Madhoun N., Al-Agha K., Zouaoui A. Graph Neural Networks for Intrusion Detection: A Survey // IEEE Access. 2023. Vol. 11. P. 49114–49139. DOI: 10.1109/ACCESS.2023.3275789.
21. Zhong M., Lin M., Zhang C., Xu Z. A Survey on Graph Neural Networks for Intrusion Detection Systems: Methods, Trends and Challenges // Computers & Security. 2024. Vol. 141. Article 103821. DOI: 10.1016/j.cose.2024.103821.
22. Fang J., Liu W., Gao Y., Liu Z., Zhang A., Wang X., He X. Evaluating Post-hoc Explanations for Graph Neural Networks via Robustness Analysis // Advances in Neural Information Processing Systems (NeurIPS'23), Oral. 2023.
23. Lewis P., Perez E., Piktus A., Petroni F., Karpukhin V., Goyal N., Küttler H., Lewis M., Yih W.-t., Rocktäschel T., Riedel S., Kiela D. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks // Advances in Neural Information Processing Systems. 2020. Vol. 33. P. 9459–9474. DOI: 10.5555/3495724.3495881.
24. Zhu X., Xie Y., Liu Y., Li Y., Hu W. Knowledge Graph-Guided Retrieval-Augmented Generation (KG²RAG) // Proceedings of NAACL 2025 (Long Papers). 2025. P. 8912–8924. DOI: 10.18653/v1/2025.naacl-long.449.
25. Farrukh Y. A., Wali S., Khan I., Bastian N. D. Xg-nid: Dual-modality network intrusion detection using a heterogeneous graph neural network and large language model // Expert Systems with Applications. 2025. p. 128089.
26. Новикова Е. С., Бухтияров М. А., Котенко И. В., Саенко И. Б., Федорченко Е. В. Обнаружение вторжений на основе федеративного обучения: архитектура системы и эксперименты // Вопросы кибербезопасности. 2023. № 6(58). С. 50–66. DOI: 10.21681/2311-3456-2023-6-50-66.
27. Котенко И. В., Хмыров С. С. Анализ моделей и методик, используемых для атрибуции нарушителей кибербезопасности при реализации целевых атак // Вопросы кибербезопасности. 2022. № 4(50). С. 52–79. DOI: 10.21681/2311-3456-2022-4-52-79.
28. Федорченко Е. В., Котенко И. В., Федорченко А. В., Новикова Е. С., Саенко И. Б. Оценивание защищенности информационных систем на основе графовой модели эксплойтов // Вопросы кибербезопасности. 2023. № 3(57). С. 23–36. DOI: 10.21681/2311-3456-2023-3-23-36.
29. Васильев В. И., Вульфин А. М., Картак В. М., Башмаков Н. М., Кириллова А. Д. Распределенная система обнаружения сетевых атак на основе федеративного трансферного обучения // Вопросы кибербезопасности. 2024. № 6. С. 117–129. DOI: 10.21681/2311-3456-2024-6-117-129.

30. Tushkanova O., Levshun D., Branitskiy A., Fedorchenko E., Novikova E., Kotenko I. Detection of Cyberattacks and Anomalies in Cyber-Physical Systems: Approaches, Data Sources, Evaluation // Algorithms. 2023. Vol. 16, No. 2. Article 85. DOI: 10.3390/a16020085.
31. Котенко И. В., Дун Х. Обнаружение атак в Интернете вещей на основе многозадачного обучения и гибридных методов сэмплирования // Вопросы кибербезопасности. 2024. № 2(60). С. 10–21. DOI: 10.21681/2311-3456-2024-2-10-21.
32. Левшун Д. С., Веснин Д. В., Котенко И. В. Прогнозирование категорий уязвимостей в конфигурациях устройств с помощью методов искусственного интеллекта // Вопросы кибербезопасности. 2024. № 3(61). С. 33–39. DOI: 10.21681/2311-3456-2024-3-33-39.

EXPLAINABLE INTERPRETATION OF INCIDENTS BASED ON A LARGE LANGUAGE MODEL AND A RETRIEVAL-AUGMENTED GENERATION

Kotenko I. V.[7], *Abramenko G. T.*[8]

Keywords: *: cybersecurity; intrusion detection and prevention systems; threat hunting; explainable artificial intelligence; knowledge graph; deep learning: heterogeneous graph neural network; MITRE ATT&CK; LLM; RAG; Suricata.*

The purpose of the study: *to increase the reliability and explainability of interpreting Suricata IDS alerts by means of an ontologically enriched knowledge graph, heterogeneous graph representations, and Retrieval-Augmented Generation (RAG) based on a local large language model.*

Research methods: *the construction of an ontology-guided knowledge graph that links Suricata data to MITRE ATT&CK tactics/techniques is the first stage of the research. Following this, a heterogeneous graph neural network (HGNN) is trained to obtain contextual node embeddings. The retrieval of relevant context via nearest neighbours in the embedding space is then conducted. Finally, a local LLM (about 7B parameters) is generated within a RAG pipeline to generate explanations. An experimental evaluation on a corpus of approximately 25,000 Suricata alerts is then conducted using metrics of interpretation accuracy, hallucination rate, and relevance.*

Results obtained: *an ontology-guided method for interpreting Suricata alerts was developed, providing more complete and accurate explanations than a baseline approach. The utilisation of ontologies has been demonstrated to enhance the substantive content of explanations by approximately 15%, while concomitantly resulting in only a marginal increase in generation time. The integration of ontology and a heterogeneous knowledge graph has been demonstrated to improve the correct mapping of alerts to MITRE ATT&CK techniques and reduce the risk of erroneous explanations.*

Scientific novelty: *the integration of an ontology and a heterogeneous knowledge graph with RAG-based generation over a local LLM to anchor low-level IDPS events to MITRE ATT&CK techniques has been suggested; the applicability of ontologies to explainable AI in cybersecurity has been shown.*

References

1. Singh A. Contextual Threat Intelligence and Alert Prioritization with Foundation-Sec-8B // International Journal of Artificial Intelligence Research and Development. 2025. Vol. 3, No. 1. P. 131–145. DOI: 10.34218/IJAIRD_03_01_009.
2. Arreche O., Guntur T., Abdallah M. XAI-IDS: Toward Proposing an Explainable Artificial Intelligence Framework for Enhancing Network Intrusion Detection Systems // Applied Sciences. 2024. Vol. 14, No. 10. Article 4170. DOI: 10.3390/app14104170.
3. Hassanin M., Moustafa N. A Comprehensive Overview of Large Language Models (LLMs) for Cyber Defences: Opportunities and Directions // arXiv preprint arXiv:2405.14487. 2024. DOI: 10.48550/arXiv.2405.14487.
4. Fan W., Ding Y., Ning L., Wang S., Li H., Yin D., Chua T.-S., Li Q. A Survey on RAG Meeting LLMs: Towards Retrieval-Augmented Large Language Models // Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'24). 2024. P. 6491–6501.
5. Mavromatis C., Karypis G. GNN-RAG: Graph Neural Retrieval for Large Language Model Reasoning // arXiv preprint arXiv:2405.20139. 2024. DOI: 10.48550/arXiv.2405.20139.
6. Wu D., Yan Y., Liu Z., Liu Z., Sun M. KG-Infused RAG: Augmenting Corpus-Based RAG with External Knowledge Graphs // arXiv preprint arXiv:2506.09542. 2025. DOI: 10.48550/arXiv.2506.09542.
7. Al-Sada B., Sadighian A., Oligeri G. MITRE ATT&CK: State of the Art and Way Forward // ACM Computing Surveys. 2024. Vol. 57, No. 1. Article 12. P. 1–37. DOI: 10.1145/3687300.
8. Li H., Shi Z., Pan C., Zhao D., Sun N. Cybersecurity Knowledge Graphs Construction and Quality Assessment // Complex & Intelligent Systems. 2024. Vol. 10. P. 1201–1217. DOI: 10.1007/s40747-023-01205-1.
9. Hemberg E., Kelly J., Shlapentokh-Rothman M., Reinstadler B., Xu K., Rutar N., O'Reilly U.-M. Linking Threat Tactics, Techniques, and Patterns with Defensive Weaknesses, Vulnerabilities and Platforms for Cyber Hunting // arXiv preprint arXiv:2010.00533. 2021. DOI: 10.48550/arXiv.2010.00533.
10. Freitas S., Gharib A. GraphWeaver: Billion-Scale Cybersecurity Incident Correlation // Proceedings of the 33rd ACM International Conference on Information and Knowledge Management (CIKM'24). 2024. P. 4479–4486. DOI: 10.1145/3627673.3680057.

7 Igor V. Kotenko, Honored Worker of Science of the Russian Federation, Dr.Sc. of Technical Sciences, Professor, Chief Scientist and Head of Laboratory of Computer Security Problems at St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), St. Petersburg, Russia. E-mail: ivkote@comsec.spb.ru
8 Georgii T. Abramenko, Ph.D. Student, Faculty of Secure Information Technologies, ITMO University (National Research University ITMO), St Petersburg, Russia. E-mail: gtabramenko@itmo.ru

11. Sikos L. F. Cybersecurity Knowledge Graphs // Knowledge and Information Systems. 2023. Vol. 65, No. 9. P. 3511–3531. DOI: 10.1007/s10115-023-01860-3.
12. Zhao X., Jiang R., Han Y., Li A., Peng Z. A Survey on Cybersecurity Knowledge Graph Construction // Computers & Security. 2024. Vol. 136. Article 103524. DOI: 10.1016/j.cose.2023.103524.
13. Xiong W., Legrand E., Åberg O., Lagerström R. Cyber Security Threat Modeling Based on the MITRE Enterprise ATT&CK Matrix // Software and Systems Modeling. 2022. Vol. 21. P. 157–177. DOI: 10.1007/s10270-021-00898-7.
14. Arikkat D. R., Abhinav M., Binu N., Rajendran N., Bhattacharjee D., Das R. K., Ajay C. R., Shiyam Sundar R., Bhavesh S. IntelIBot: A Retrieval-Augmented Large Language Model Chatbot for Cyber Threat Knowledge Delivery // arXiv preprint arXiv:2411.05442. 2024. DOI: 10.48550/arXiv.2411.05442.
15. Kurniawan K., Kiesling E., Ekelhart A. CyKG-RAG: Towards Knowledge-Graph Enhanced Retrieval-Augmented Generation for Cybersecurity // CEUR Workshop Proceedings. 2024. Vol. 3950. P. 51–64.
16. Tellache A., Amara-Korba A., Mokhtari A., Moldovan H., Ghamri-Doudane Y. Advancing Autonomous Incident Response: Leveraging LLMs and Cyber Threat Intelligence // arXiv preprint arXiv:2508.10677. 2025. DOI: 10.48550/arXiv.2508.10677.
17. Qi Y., Gu Z., Li A., Zhang X., Shafiq M., Mei Y., Lin K. Cybersecurity Knowledge Graph Enabled Attack Chain Detection for Cyber-Physical Systems // Computers & Electrical Engineering. 2023. Vol. 108. Art. 108660. DOI: 10.1016/j.compeleceng.2023.108660.
18. Ekelhart A., Ekaputra F.J., Kiesling E. SLOGERT: Automated Log Knowledge Graph Construction // The Semantic Web – ESWC 2021. Lecture Notes in Computer Science. 2021. Vol. 12731. P. 219–234. DOI: 10.1007/978-3-030-77385-4_16.
19. Kurniawan K., Ekelhart A., Kiesling E., Winkler D., Quirchmayr G., Tjoa A.M. Virtual Knowledge Graphs for Federated Log Analysis // Proceedings of the 16th International Conference on Availability, Reliability and Security (ARES'21). 2021. Article 50. 10 p. DOI: 10.1145/3465481.3470077.
20. Bilot T., El-Madhoun N., Al-Agha K., Zouaoui A. Graph Neural Networks for Intrusion Detection: A Survey // IEEE Access. 2023. Vol. 11. P. 49114–49139. DOI: 10.1109/ACCESS.2023.3275789.
21. Zhong M., Lin M., Zhang C., Xu Z. A Survey on Graph Neural Networks for Intrusion Detection Systems: Methods, Trends and Challenges // Computers & Security. 2024. Vol. 141. Article 103821. DOI: 10.1016/j.cose.2024.103821.
22. Fang J., Liu W., Gao Y., Liu Z., Zhang A., Wang X., He X. Evaluating Post-hoc Explanations for Graph Neural Networks via Robustness Analysis // Advances in Neural Information Processing Systems (NeurIPS'23), Oral. 2023.
23. Lewis P., Perez E., Piktus A., Petroni F., Karpukhin V., Goyal N., Küttler H., Lewis M., Yih W.-t., Rocktäschel T., Riedel S., Kiela D. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks // Advances in Neural Information Processing Systems. 2020. Vol. 33. P. 9459–9474. DOI: 10.5555/3495724.3495881.
24. Zhu X., Xie Y., Liu Y., Li Y., Hu W. Knowledge Graph-Guided Retrieval-Augmented Generation (KG²RAG) // Proceedings of NAACL 2025 (Long Papers). 2025. P. 8912–8924. DOI: 10.18653/v1/2025.naacl-long.449.
25. Farrukh Y. A., Wali S., Khan I., Bastian N. D. Xg-nid: Dual-modality network intrusion detection using a heterogeneous graph neural network and large language model // Expert Systems with Applications. 2025. p. 128089.
26. Novikova, E. S.; Bukhtiyarov, M. A.; Kotenko, I. V.; Saenko, I. B.; Fedorchenko, E. V. Intrusion Detection Based on Federated Learning: System Architecture and Experiments. Voprosy kiberbezopasnosti [Cybersecurity Issues], 2023, no. 6(58), 50–66. https://doi.org/10.21681/2311-3456-2023-6-50-66.
27. Kotenko, I. V.; Khmyrov, S. S. Analysis of Models and Methods Used for Attribution of Cybersecurity Adversaries in Targeted Attacks. Voprosy kiberbezopasnosti [Cybersecurity Issues], 2022, no. 4(50), 52–79. https://doi.org/10.21681/2311-3456-2022-4-52-79.
28. Fedorchenko, E. V.; Kotenko, I. V.; Fedorchenko, A. V.; Novikova, E. S.; Saenko, I.B. Assessing the Security of Information Systems Based on a Graph Model of Exploits. Voprosy kiberbezopasnosti [Cybersecurity Issues], 2023, no. 3(57), 23–36. https://doi.org/10.21681/2311-3456-2023-3-23-36.
29. Vasiliev, V. I.; Vulfin, A. M.; Kartak, V. M.; Bashmakov, N. M.; Kirillova, A. D. Distributed Network Attack Detection System Based on Federated Transfer Learning. Voprosy kiberbezopasnosti [Cybersecurity Issues], 2024, no. 6, 117–129. https://doi.org/10.21681/2311-3456-2024-6-117-129.
30. Tushkanova, O.; Levshun, D.; Branitskiy, A.; Fedorchenko, E.; Novikova, E.; Kotenko, I. Detection of Cyberattacks and Anomalies in Cyber-Physical Systems: Approaches, Data Sources, Evaluation. Algorithms, 2023, 16(2), 85. https://doi.org/10.3390/a16020085.
31. Kotenko, I. V.; Dun, H. Detection of Attacks in the Internet of Things Based on Multitask Learning and Hybrid Sampling Methods. Voprosy kiberbezopasnosti [Cybersecurity Issues], 2024, no. 2(60), 10–21. https://doi.org/10.21681/2311-3456-2024-2-10-21.
32. Levshun, D. S.; Vesnin, D. V.; Kotenko, I. V. Predicting Categories of Vulnerabilities in Device Configurations Using Artificial Intelligence Methods. Voprosy kiberbezopasnosti [Cybersecurity Issues], 2024, no. 3(61), 33–39. https://doi.org/10.21681/2311-3456-2024-3-33-39.

КОНЦЕПТУАЛЬНАЯ МОДЕЛЬ ФУНКЦИОНИРОВАНИЯ СИСТЕМЫ ЦИФРОВОГО ДОКУМЕНТООБОРОТА В РАМКАХ ПАРАДИГМЫ «ИНДУСТРИЯ 4.0»

Тали Д. И.[1], Финько О. А.[2]

DOI: 10.21681/2311-3456-2025-5-68-77

Цель исследования: *формализация процесса функционирования системы цифрового документооборота, включающей в себя систему управления документами и системы – источники данных. Реализация предлагаемого подхода в целях формирования цифровой инфраструктуры управления информацией, отвечающей основным положениям концепции «Индустрия 4.0».*

Методы исследования: *применение методов системного анализа к условиям цифровизации структурно-сложных систем на примере электронного документооборота.*

Результат исследования: *разработана концептуальная модель функционирования системы цифрового документооборота с учетом таких характеристик перспективных цифровых систем как автономность, распределенность, интеллектуальность. Введена система показателей и критериев ее оценивания в целях повышения качества информационного взаимодействия между структурными подразделениями организаций, эксплуатирующих подобную инфраструктуру.*

Научная новизна: *представлена и обоснована концептуальная модель функционирования системы цифрового документооборота, основанная на иерархической декомпозиции ее структуры, учитывающая взаимосвязь между уровнями взаимодействия исследуемой системы. Предложенный подход в условиях цифровой трансформации позволяет обеспечить целевое предназначение (целостность) системы в течение заданного периода времени при воздействии дестабилизирующих факторов на любой из ее уровней.*

Ключевые слова: *контент, метаданные, цифровой документ, принципы цифровизации, интеллектуальные агенты, цифровая трансформация документооборота, целостность системы.*

Введение

Цифровые технологии, основанные на программно-аппаратном обеспечении и использовании возможностей информационно-коммуникационных сетей, с каждым годом совершенствуются и интегрируются во все сферы жизни, вызывая трансформацию общества и глобальной экономики, что становится завершающим этапом третьей промышленной революции. Подобные тенденции являются катализатором зарождения очередного этапа развития науки и техники, характеризующегося использованием новых подходов, позволяющих обеспечить эффективность государственного управления [1].

В условиях глобальной цифровизации и модернизации государственного управления особое внимание уделяется развитию систем электронного документооборота (СЭД). Эти системы представляют собой уникальные инструменты, направленные на оптимизацию работы государственных органов, обеспечивая эффективное управление документами, обмен информацией и взаимодействие между различными ведомствами [2, 3].

Однако в настоящее время нельзя утверждать о полноценном переходе на безбумажный документооборот в связи с тем, что в государственных структурах он приобрел смешанную форму, выражающуюся в использовании документов в формате сканированных бумажных оригиналов, что в ряде работ носит название гибридных технологий обработки электронных документов (ЭлД) [4–6]. Так, из открытых источников[3] известно, что в 2019 году только 40 % документов Правительства РФ были представлены в электронном виде, сейчас этот показатель увеличился до 90 %, но вместе с тем, говорить о переходе на полноценный «электронный документооборот» преждевременно. Об этом свидетельствует инициатива по разработке правил «цифрового» взаимодействия, единых для всех государственных органов.

Проблема перехода к цифровым технологиям вызвана рядом трудностей, связанных с отсутствием нормативной базы, а также непротиворечивого и корректного терминологического аппарата, объединяющего смежные сферы деятельности в рамках процесса управления документами.

1 Тали Дмитрий Иосифович, кандидат технических наук, докторант специальной кафедры Краснодарского высшего военного орденов Жукова и Октябрьской Революции Краснознаменного училища имени генерала армии С. М. Штеменко, г. Краснодар, Россия. E-mail: dimatali@mail.ru

2 Финько Олег Анатольевич, доктор технических наук, профессор, профессор специальной кафедры Краснодарского высшего военного орденов Жукова и Октябрьской Революции Краснознаменного училища имени генерала армии С. М. Штеменко; академический советник Российской академии ракетных и артиллерийских наук (РАРАН), г. Краснодар, Россия. E-mail: ofinko@yandex.ru. Web: http://www.mathnet.ru/ person40004

3 Информационный портал «Коммерсант». Бюрократия выходит за периметр. URL:https://www.kommersant.ru/doc/7715109 (дата обращения: 04.07.2025).

Эволюция понятия «электронный документ» в условиях цифровой трансформации

В различных нормативных актах насчитывается более двух десятков определений электронного документа [3]. Исходя из чего, можно констатировать факт отсутствия однозначного трактования термина «электронный документ», несмотря на неоднократные попытки внести уточнения и дополнения в существующие определения[4,5,6], что вызывает сложности в проектировании СЭД, отвечающих основным принципам цифровизации [7].

При этом нельзя не отметить динамику изменений данного определения, соответствующую уровню развития техники. Так, прообразом ЭлД был «документ на машинном (магнитном) носителе», определяемый как документ, созданный с использованием носителей и способов записи, обеспечивающих обработку его информации электронно-вычислительной машиной[7]. Следующая трактовка предлагала его понимать, как документированную информацию, созданную, полученную и сохраняемую организацией или частным лицом в качестве доказательства и актива для подтверждения правовых обязательств или деловой транзакции[8], что носило обобщающий характер как для бумажного, так и для ЭлД. Очередной стандарт[9] сузил понятие «электронного документа» до информации, представленной в электронной форме, но тем самым указал на техническую сущность его происхождения. Кроме того, последующий стандарт не дает конкретного определения исследуемому понятию, однако обозначает его состав как совокупность контента и метаданных, что вновь подтверждает техническую сущность ЭлД и косвенно свидетельствует о невозможности применения технологий обработки, идентичных бумажному документообороту[10]. На сегодняшний день определение «электронный документ»[11] исключено ввиду отсутствия единого мнения специалистов по трактованию этого термина[12]. Однако, там же приводится определение «документированная информация» (ДИ), под которой понимается информация, рассматриваемая как отдельная единица, обладающая свойствами документа или его части. Откуда следует вывод о том, что информационный объект, состоящий из контента и / или метаданных, может рассматриваться как документированная информация (без применения ЭП). Подтверждением тому является[13], где понятие «электронный документ» приобретает новую форму и имеет наименование «цифровой документ», определяемый как категория ЭлД, изначально созданного в цифровой форме с соблюдением правил документирования без издания его на бумажном носителе, подписанного ЭП (следуя данной логике, информационный объект не подписанный ЭП, возможно считать цифровой документированной информацией). Этот термин вносит окончательное разграничение в бумажный и электронный документооборот, тем самым подводя итог в выборе методов и средств обработки цифровых документов / цифровой документированной информации (ЦД / ЦДИ). Таким образом, понятие «электронный документ» приобретает конкретные черты, способствующие модернизации СЭД, соответствующих современному этапу развития науки и техники.

Тенденции развития систем электронного документооборота в условиях цифровизации

Совершенствование СЭД стало результатом комплекса факторов, формировавшихся на разных этапах развития общества и государственного управления. В связи с чем, эволюцию термина «электронный документ» можно соотнести с появлением СЭД, реализующих ряд функций, отвечающих требованиям конкретного этапа развития систем такого типа (рис. 1).

Учитывая вышеизложенное, исследуемые системы можно условно структурировать следующим образом:

1) 1990–2000-е: СЭД I-го поколения (базовая автоматизация документооборота);
2) 2000–2010-е: СЭД II-го поколения (смешанный документооборот);
3) 2010–2020-е: СЭД III-го поколения (гибридный документооборот);
4) 2020-е: СЭД IV-го поколения (цифровой документооборот).

СЭД I-го поколения характеризуются основной функцией, заключающейся в регистрации фактов наличия документов в подразделении. При этом такое решение было трудно реализуемым ввиду низкого уровня компьютеризации.

4 ГОСТ Р ИСО 15489-1-2019 «Система стандартов по информации, библиотечному и издательскому делу. Управление документами. Часть 1. Понятия и принципы». – М.: Стандартинформ, 2019.
5 ГОСТ Р 7.0.8-2025 «Система стандартов по информации, библиотечному и издательскому делу. Делопроизводство и архивное дело. Термины и определения». – М.: Стандартинформ, 2025.
6 ГОСТ Р 59999-2025 «Цифровой документооборот организации. Требования к эталонной модели». – М.: Российский институт стандартизации, 2025.
7 ГОСТ Р 51141-98 «Делопроизводство и архивное дело. Термины и определения» – М.: Госстандарт, 1999.
8 ГОСТ Р ИСО 15489-1-2007 «Система стандартов по информации, библиотечному и издательскому делу. Управление документами. Общие требования» – М.: Стандартинформ, 2007.
9 ГОСТ Р 7.0.8-2013 «Система стандартов по информации, библиотечному и издательскому делу. Делопроизводство и архивное дело. Термины и определения» – М.: Стандартинформ, 2014.
10 ГОСТ Р ИСО 15489-1-2019 «Система стандартов по информации, библиотечному и издательскому делу. Управление документами. Часть 1. Понятия и принципы». – М.: Стандартинформ, 2019.
11 ГОСТ Р 7.0.8-2025 «Система стандартов по информации, библиотечному и издательскому делу. Делопроизводство и архивное дело. Термины и определения». – М.: Стандартинформ, 2025.
12 Азарова С. В. 50 новых терминов. Что нужно знать о ГОСТе Р 7.0.8-2025. URL: https://ecm-journal.ru/material /50-novykh-terminov-chto-nuzhno-znat-o-goste-r-7082025-po-arkhivnomu-delu?ysclid=mav3odcxoh320574304 (дата обращения: 04.07.2025).
13 ГОСТ Р 59999-2025 «Цифровой документооборот организации. Требования к эталонной модели». – М.: Российский институт стандартизации, 2025.

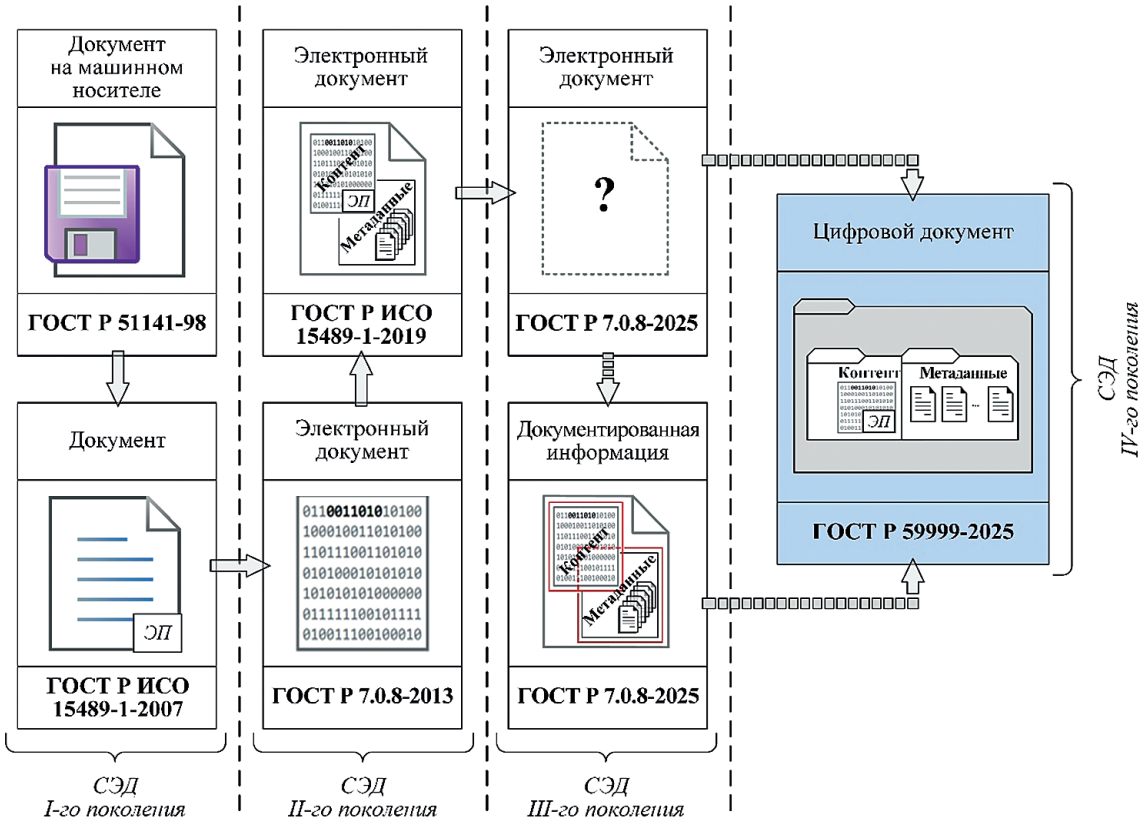


Рис. 1. Эволюция понятия «электронный документ» в соответствии с развитием АИС ЭД

С расширением доступа к сети Интернет и широким распространением документов в электронной форме СЭД начали пользоваться большим спросом. Однако, юридически значимый ЭлД рассматривался как недостоверная копия, в связи с чем бумажный оригинал хранился в обязательном порядке, что способствовало развитию смешанного документооборота.

Появление ЭП и регулирование ее применения на государственном уровне стало первым шагом к безбумажным документообороту, однако и в настоящее время переход на подобные технологии в полном объеме не завершен, что позволяет обозначить данный этап как гибридный документооборот.

Вместе с тем, развитие таких технологий как распределенные реестры, облачные технологии, а также искусственный интеллект способствуют появлению СЭД нового поколения с расширенным функционалом, позволяющим эффективнее решать возложенные на них задачи, а также отвечать требованиям безопасности обрабатываемой информации в условиях повсеместной цифровизации.

В целях определения тенденций совершенствования СЭД приведем общие результаты ретроспективного анализа их функционала в соответствии с поколениями (таблица 1).

Как отмечают специалисты [2], существующие решения с исполняемым в настоящее время функционалом не позволят перейти к принципиально новому способу управления и совершить цифровую трансформацию в организации, определяемую как процесс перехода организации на работу с ЦД, сопровождаемый изменениями в ее цифровой и организационной структуре[14].

При этом СЭД IV-го поколения (системы цифрового документооборота) приобретают принципиально новую форму, представляющую собой информационную инфраструктуру, включающую систему, обеспечивающую управление документами и доступ к ним в течение определенного времени, и системы – источники данных для создания ЦД[14]. Это указывает на высокую степень интеграции с любыми автоматизированными информационными системами (АИС), обрабатывающими как документированную информацию, так и ЭлД. Причем в качестве систем – источников данных необходимо учитывать и интеллектуальных агентов, ввиду начала внедрения таких технологий на государственном уровне, например, цифровая платформа ГосУслуги. Так, по мнению специалистов, документы и сообщения, автоматически генерируемые АИС, в скором времени будут доминировать по объему в организациях цифровой экономики [3, 8, 9].

В таком случае, подобные системы будут иметь децентрализованное управление и распределенную

14 ГОСТ Р 59999-2025 «Цифровой документооборот организации. Требования к эталонной модели». – М.: Российский институт стандартизации, 2025.

Таблица 1.

Функциональное развитие СЭД

	СЭД I-го поколения (1990-2000-е)	СЭД II-го поколения (2000-2010-е)	СЭД III-го поколения (2010-2020-е)
Исполняемый функционал	создание и хранение документов;	автоматизация процессов;	автоматизация процессов;
	учет и регистрация документов;	централизованное управление;	централизованное управление;
	поиск и сортировка документов;	поиск и сортировка документов;	поиск и сортировка документов;
	ручное согласование документов	сокращение времени на обработку и распределение документов;	сокращение времени на обработку и распределение документов;
		мониторинг и учет;	мониторинг и учет;
		снижение ошибок	снижение ошибок;
			использование ЭП;
			повышение безопасности обрабатываемой информации;
			контроль версий документов;
			удаленный доступ;
			совместная работа с документами;
			электронное хранение.

структуру, что вызывает необходимость переосмысления подходов к формированию, обработке, хранению ЦД / ЦДИ, существующих только в цифровой форме, а также их защите в процессе жизненного цикла. Далее такие высокоинтегрированные системы будем называть системами цифрового документооборота (СЦД). Введем понятие «СЦД» – комплекс технических средств, программного обеспечения, методов обработки информации и организационных процессов, обеспечивающий автоматизированное управление, создание, хранение, поиск и обработку ЦД, а также поддерживающий доступ к ЦДИ на протяжении жизненного цикла[14,15].

В условиях повсеместной цифровизации возникает проблемная ситуация, состоящая в необходимости разработки моделей и методов функционирования СЦД, основывающихся на принципах автономности, интеллектуальности, распределенности, в целях повышения качества информационного взаимодействия. При этом существующий научно-методический аппарат не учитывает интегрированную структуру подобных систем, что не позволяет расширить их функционал до уровня, отвечающего требованиям современного этапа цифровой трансформации.

15 ГОСТ 15971-90 «Системы обработки информации. Термины и определения». – М.: Государственный комитет СССР по управлению качеством продукции и стандартам, 1992.

Вместе с тем, внедрение новых технологий порождает и новые угрозы, вызываемые рядом уязвимостей, сопутствующих этим решениям [10–12]. В связи с чем, модели и методы функционирования СЦД должны учитывать подобные дестабилизирующие факторы, в целях обеспечения стабильности ее параметров.

Таким образом, целью исследования является формирование концептуальной модели функционирования СЦД, основанной на применении принципов теории систем и системного анализа к условиям цифровизации структурно-сложных систем. При этом под формированием концептуальной модели функционирования СЦД понимается процесс установления взаимосвязей между ее структурными уровнями в интересах формализации и последующей оценки математическими методами каждого из них, при воздействии дестабилизирующих факторов [13, 14].

Постановка задачи

Обозначим СЦД следующим образом:

$$Sd = \langle A, L, Y, X, D, T, R \rangle, \quad (1)$$

где A – множество технических средств (аппаратных комплексов, источников информации, серверов и т.п.), обеспечивающих поддержку системы;

L – множество программных средств, реализующих функционал обработки и управления ЦД / ЦДИ; Y – множество методов обработки информации, включающие алгоритмы обработки, хранения, поиска и передачи ЦД / ЦДИ; X – множество агентов и их действий, связанных с управлением ЦД / ЦДИ (под агентами понимаются, как уполномоченные пользователи, так интеллектуальные агенты, причем их действия могут рассматриваться как дестабилизирующий фактор); D – множество ЦД и связанной с ними ЦДИ, которая обрабатывается и хранится системой; T – множество интервалов времени, в которые наблюдается СЦД, $t_i \in T$, $i = \overline{0,n}$; t_0 – начало эксплуатации, t_n – конец эксплуатации; R – множество результатов обработки ЦД / ЦДИ; $F{:}A\times L\times Y\times X\times D\times R\longrightarrow C$ – функция, характеризующая выполнение операций обработки ЦД / ЦДИ, где C – множество сценариев функционирования СЦД [7].

$$C = \left\{C_{(+)}^{(x_j)},\ C_{(-)}^{(x_j)}\right\},$$

где $C_{(+)}^{(x_j)}$ – сценарий функционирования СЦД, при котором целостность ЦД / ЦДИ под влиянием $x_j \in X$ обеспечена (при $j = 1,\ldots,z$), $C_{(-)}^{(x_j)}$ – сценарий функционирования СЦД, при котором целостность ЦД / ЦДИ под влиянием $x_j \in X$ утрачена или обеспечение указанного свойства не доступно в текущий момент функционирования СЦД.

Исходя из определения[16] признаком «целостности системы» является реализация ею своего целевого предназначения. Целевым предназначением рассматриваемых систем является обеспечение целостности ЦД / ЦДИ. Введем допущение: нахождение ЦД / ЦДИ в состоянии $C_{(+)}^{(x_j)}$ характеризует состояние целостности СЦД, а переход ЦД / ЦДИ в состояние $C_{(-)}^{(x_j)}$ – нарушение целостности СЦД. Справедливо утверждать, что в первом случае это приведет к нормальному сценарию функционирования (штатный режим работы) СЦД, а во втором – к его нарушению.

В таком случае, общая задача исследования может быть сформулирована так:

$$Q{:}F(A,L,Y^*,X,D,T,R) \longrightarrow \min\left|C_{(-)}^{(x_j)}\right| \mid t_i \in T,\ i = \overline{0,n}, \quad (2)$$

где Q – комплекс моделей и методов обработки ЦД / ЦДИ, подлежащих разработке; $\left|C_{(-)}^{(x_j)}\right|$ – количество (мощность множества) сценариев нарушения функционирования СЦД.

Необходимо разработать такой комплекс моделей и методов обработки ЦД / ЦДИ, который при заданных ограничениях и допущениях позволит минимизировать количество сценариев нарушения функционирования СЦД в условиях воздействия дестабилизирующих факторов.

16 ГОСТ Р 59341-2021 «Системная инженерия. Защита информации в процессе управления информацией системы». – М.: Стандартинформ, 2021.

При этом целевая функция обеспечения устойчивости СЦД примет вид:

$$Q{:}F(A,L,Y^*,X,D,T) \longrightarrow \max \mathrm{J} \mid t_i \in T,\ i = \overline{0,n}, \quad (3)$$

где J – показатель устойчивого функционирования СЦД, определяемый как:

$$J = \int_{t_0}^{t_n}\left[W_s S(t) - W_x \sum_{j=1}^{z} \beta_j \chi_j(t)\right]dt, \quad (4)$$

где W_s и W_x – веса, отражающие важность стабилизирующих параметров и дестабилизирующих факторов (при этом $W_s \geq 0$ – вес стабилизирующих параметров, $W_x \geq 0$ – вес дестабилизирующих факторов); $S(t)$ – показатель стабильности СЦД; β_j – коэффициент важности каждого дестабилизирующего фактора; $\chi_j(t)$ – уровень воздействия дестабилизирующих факторов на СЦД при $j = 1,\ldots,z$.

Исходя из (3) необходимо максимизировать совокупный показатель J устойчивого функционирования СЦД за период $[t_0, \ldots, t_n]$, где $W_s S(t)$ – вклад в стабильность системы; $-W_x \sum_{j=1} \beta_j \chi_j(t)$ – отрицательный вклад в стабильность системы от воздействия дестабилизирующих факторов, которые необходимо минимизировать [15, 16].

Для достижения целевой функции необходимо выбрать управляющее воздействие $U(t)$, в качестве которого выступает комплекс Q моделей и методов обработки ЦД / ЦДИ, который влияет на состояние СЦД и уровень воздействия дестабилизирующих факторов, то есть $S(t)$ и $\chi_j(t)$ зависят от реализуемых мер.

Тогда целевая функция СЦД, направленная на обеспечение ее устойчивости в условиях воздействия дестабилизирующих факторов примет вид:

$$J = \max_{U(t)} \int_{t_0}^{t_n}\left[W_s S(t)U(t) - W_x \sum_{j=1}^{z} \beta_j \chi_j(t)U(t)\right]dt. \quad (5)$$

Концептуальная модель функционирования системы цифрового документооборота

Для достижения (5) воспользуемся принципами теории систем и системного анализа [17–19]. Для чего используем иерархический подход в описании структуры СЦД как структурно-сложного технического объекта:

1) микроуровень, или уровень внутренних проявлений элементов, составляющих исследуемую систему;
2) макроуровень, или уровень индивидуальных проявлений исследуемой системы в ее взаимосвязи со своим локальным окружением;
3) метауровень, или уровень коллективных проявлений систем исследуемого вида, входящих в состав некоторой более общей метасистемы.

Обозначим структуру СЦД, как ρ_{micro}, ρ_{macro}, ρ_{meta}, где ρ_{micro} – микроуровень, ρ_{macro} – макроуровень, ρ_{meta} – метауровень исследуемого объекта (рис. 2).

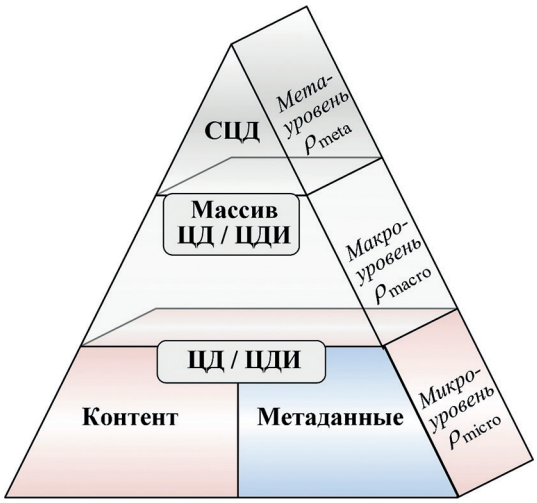


Рис. 2. Структура СЦД как структурно-сложного технического объекта

Подробнее опишем ρ_{micro} СЦД, состоящий из ЦД / ЦДИ, который является системным объектом $D_{micro}^{(t_i)}$, представляющим собой в момент времени t_i совокупность контента $K_{micro}^{(t_i)}$ и множества метаданных $M_{micro}^{(t_i)}$:

$$Sd_{micro}:\ D_{micro}^{(t_i)} = \{K_{micro}^{(t_i)}, M_{micro}^{(t_i)}\} \mid t_i \in T,\ i = \overline{0,n}. \quad (6)$$

В таком случае, состояние целостности СЦД можно представить в виде функции, описывающей результаты обработки ЦД / ЦДИ:

$$F_{micro}:\ D_{micro}^{(t_i)} \longrightarrow R_{micro} \mid t_i \in T,\ i = \overline{0,n}, \quad (7)$$

при этом $R_{micro} = \{R_{micro}^{(+)}, R_{micro}^{(-)}\}$, где $R_{micro}^{(+)}$ – положительный «1» результат обработки ЦД / ЦДИ (целостность обеспечена), $R_{micro}^{(-)}$ – отрицательный «0» результат обработки ЦД / ЦДИ (целостность нарушена).

В силу того, что в настоящее время $D_{micro}^{(t_i)}$ представляет собой суммативный объект, зависящий от двух компонент – контента $K_{micro}^{(t_i)}$ и метаданных $M_{micro}^{(t_i)}$, возникает уязвимость в обеспечении его целостности [7, 20]. Иными словами, $D_{micro}^{(t_i)}$ перестает существовать, как функционально целостный объект, при недопустимом изменении или уничтожении любой из его компонент. То есть, нахождение СЦД в состоянии $R_{micro}^{(-)}$ (нарушение функционирования) увеличивается до трех вариантов: (0, 0), (0, 1), (1, 0).

Таким образом, в целях обеспечения целостности СЦД на микроуровне среды функционирования требуется:

$$Sd_{micro}:\ P_{micro} = \frac{1}{n}\sum_{i=1}^{n} F_{micro}(D_{micro}^{(t_i)}) \longrightarrow 1 \mid t_i \in T,\ i = \overline{1,n}, \quad (8)$$

где P_m – вероятность обеспечения целостности СЦД на микроуровне среды функционирования.

Предполагается, что выполнить (8) возможно за счет формирования структуры ЦД / ЦДИ, обладающей свойствами целостности и интегративности.

В качестве критерия оценки уровня устойчивости СЦД на микроуровне среды функционирования примем следующее условие:

$$P_{micro}^{\text{разр.}} > P_{micro}^{\text{сущ.}}, \quad (9)$$

где $P_{micro}^{\text{разр.}}$ и $P_{micro}^{\text{сущ.}}$ – вероятность обеспечения целостности СЦД на микроуровне среды функционирования при разработанной и существующей структуре ЦД / ЦДИ в условиях воздействия дестабилизирующих факторов [7, 21].

На макроуровне ρ_{macro} СЦД представлена обеспечивающей инфраструктурой, что может быть записано в виде:

$$Sd_{macro} = \langle A_{macro}, L_{macro}, Y_{macro}, X_{macro}, D_{macro}, T \rangle, \quad (10)$$

где A_{macro} – множество технических средств (аппаратных комплексов, источников информации, серверов и т.п.), обеспечивающих функционирование системы; L_{macro} – множество программных компонентов (программные модули, программное обеспечение и т.п.), реализующих функционал обработки и управления ЦД / ЦДИ; Y_{macro} – множество методов обработки ЦД / ЦДИ; X_{macro} – множество агентов и их действий, связанных с управлением ЦД / ЦДИ (под агентами понимаются, как уполномоченные пользователи, так интеллектуальные агенты, причем их действия могут рассматриваться как дестабилизирующий фактор); D_{macro} – множество (массив) ЦД / ЦДИ, как системных объектов, обрабатываемых на макроуровне ρ_{macro} СЦД; T – множество интервалов времени, в которые наблюдается СЦД, $t_i \in T,\ i = \overline{0,n}$; t_0 – начало эксплуатации, t_n – конец эксплуатации.

Исходя из того, что ЭлД в известных системах хранятся в виде баз данных, а их изменения имеют динамический характер, состояние целостности СЦД на макроуровне ρ_{macro} среды функционирования представим в виде функции:

$$F_{macro}:\ D_{macro}^{(t_i)} \times Y_{macro} \times X_{macro} \longrightarrow R_{macro}, \mid t_i \in T,\ i = \overline{0,n}, \quad (11)$$

где $R_{macro} = \{R_{macro}^{(+)}, R_{macro}^{(-)}\}$, при этом $R_{macro}^{(+)}$ – положительный результат («1») обработки массива ЦД / ЦДИ (целостность обеспечена), $R_{macro}^{(-)}$ – отрицательный результат («0») обработки массива ЦД / ЦДИ (целостность нарушена).

Аналогично выражению (8) запишем требование по обеспечению целостности СЦД на макроуровне среды функционирования:

$$Sd_{macro}:\ P_{macro} = \frac{1}{n}\sum_{i=1}^{n} F_{macro}(D_{macro}^{(t_i)}, Y_{macro}, X_{macro}) \longrightarrow$$
$$\longrightarrow 1 \mid t_i \in T,\ i = \overline{1,n}, \quad (12)$$

где P_{macro} – вероятность обеспечения целостности СЦД на макроуровне среды функционирования.

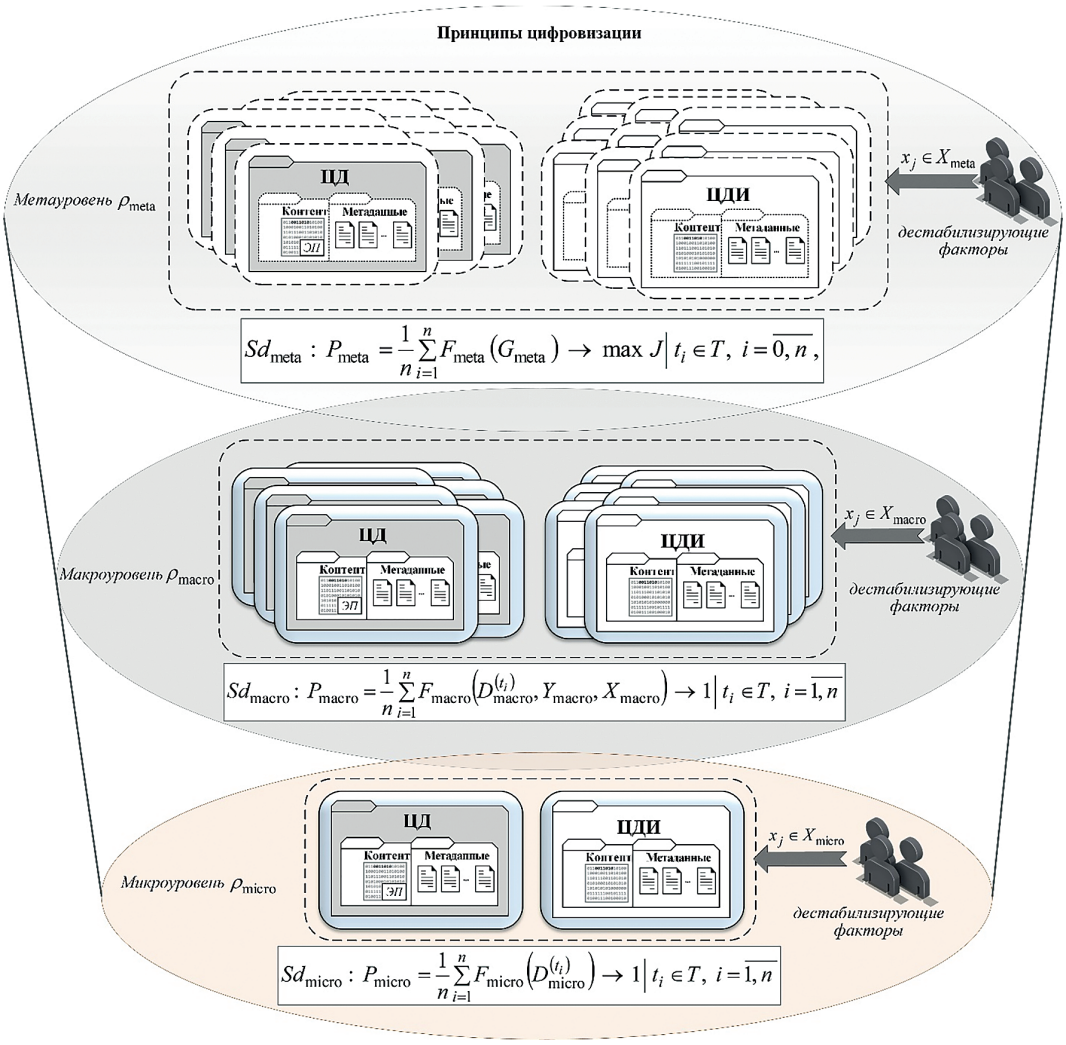


Рис. 3. Структура концептуальной модели функционирования СЦД в рамках парадигмы «Индустрия 4.0»

В таком случае, в качестве критерия оценки уровня устойчивости СЦД на макроуровне среды функционирования примем условие:

$$P_{macro}^{\text{разр.}} > P_{macro}^{\text{сущ.}}, \qquad (13)$$

где $P_{macro}^{\text{разр.}}$ и $P_{macro}^{\text{сущ.}}$ – вероятность обеспечения целостности СЦД на макроуровне среды функционирования при разработанных и существующих методах обработки ЦД / ЦДИ в условиях воздействия дестабилизирующих факторов [22].

На метауровне ρ_{meta} среды функционирования СЦД представляет собой распределенную структуру под управлением технологии цифровых двойников:

$$Sd_{meta} = \langle A_{meta}, L_{meta}, G_{meta}, \kappa_{meta}, E_{meta} \rangle, \qquad (14)$$

где A_{meta} – множество распределенных технических средств (аппаратных комплексов, источников информации, серверов и т.п.), обеспечивающих функционирование системы; L_{meta} – множество программных компонентов распределенных технических средств (программные модули, программное обеспечение и т.п.), реализующих функционал обработки и управления ЦД / ЦДИ; G_{meta} – множество состояний СЦД, отражающих ее текущие параметры; κ_{meta} – входные параметры, отражающие состояние микро – ρ_{micro} и макроуровня ρ_{macro} СЦД; E_{meta} – множество механизмов обнаружения нарушения целостности, обрабатываемой информации на всех уровнях среды функционирования СЦД.

Ввиду того, что СЦД на метауровне ρ_{meta} среды функционирования имеет распределенную структуру состояние ее целостности в формализованном виде можно представить в виде функции:

$$F_{meta} : G_{meta} \to C, \qquad (15)$$

где $C = \{C_{(+)}^{(x_j)}, C_{(-)}^{(x_j)}\}$, при этом $C_{(+)}^{(x_j)}$ – нормальный сценарий функционирования (целостность СЦД обеспечена), $C_{(-)}^{(x_j)}$ – сценарий нарушения функционирования (целостность СЦД нарушена).

Запишем требование по обеспечению целостности СЦД на метауровне среды функционирования в виде:

$$Sd_{meta}: P_{meta} = \frac{1}{n}\sum_{i=1}^{n} F_{meta} G_{meta} \longrightarrow max J \mid t_i \in T,\ i = \overline{0,n}, \quad (16)$$

где P_{meta} – показатель обеспечения целостности СЦД на метауровне среды функционирования.

В качестве критерия оценки уровня устойчивости СЦД на метауровне среды функционирования примем условие:

$$P_{meta}^{\text{разр.}} > P_{meta}^{\text{сущ.}}, \quad (17)$$

где $P_{meta}^{\text{разр.}}$ и $P_{meta}^{\text{сущ.}}$ – вероятность обеспечения целостности СЦД на метауровне среды функционирования при использовании разработанных и существующих методов обработки ЦД / ЦДИ в условиях воздействия дестабилизирующих факторов [23].

Исходя из приведенной формализации структуры СЦД взаимосвязи между уровнями определяются следующим образом:

$$\begin{cases} \rho_{micro}: D_{micro}^{(t_i)} = K_{micro}^{(t_i)}, M_{micro}^{(t_i)} \mid t_i \in T,\ i = \overline{0,n}; \\ \rho_{macro}: D_{macro}^{(t_i)} = \Lambda_{macro}(D_{micro}^{(t_i)}) \mid t_i \in T,\ i = \overline{0,n}; \\ \rho_{meta}: G_{meta} = \Theta_{meta}(\kappa_{meta}, D_{macro}^{(t_i)}) \mid t_i \in T,\ i = \overline{0,n}; \end{cases} \quad (18)$$

где Λ_{macro}, Θ_{meta} – функции взаимодействия для описания связей между уровнями.

Предлагаемая формализация представлена в графическом виде как структура концептуальной модели функционирования СЦД в условиях цифровой трансформации (рис. 3).

При этом в целях качественной оценки устойчивости СЦД будем использовать критерий (3), а количественной – критерии (9), (13), (17).

Выводы

Таким образом, авторами была сделана попытка формализации процесса функционирования перспективных СЦД на основе методов теории систем и системного анализа с учетом современных принципов цифровизации (автономность, распределенность, интеллектуальность), что отвечает современным запросам общества и производства в рамках концепции «Индустрия 4.0». Предлагаемый подход к формализованному представлению систем этого класса позволяет ввести систему показателей и критериев оценивания, что упростит их проектирование и практическое внедрение. Предполагается, что использование такого подхода позволит повысить качество информационного взаимодействия между структурными подразделениями организаций, эксплуатирующих подобную инфраструктуру.

Литература

1. Шваб К. Четвертая промышленная революция. – М: Эксмо, 2021. 208 с.
2. Управление документами в цифровой экономике: организация, регламентация, реализация / М. В. Ларин, Н. Г. Суровцева, Е. В. Терентьева, В. Ф. Янковая / Под ред. М. В. Ларина – М.: РГГУ, 2021. 242 с.
3. Ларин М. В. Электронные документы: теоретические аспекты // Самарский архивист. 2021. № 2. С. 3–9.
4. Елисеев Н. И., Тали Д. И. Проблемы и перспективы развития систем юридически значимого электронного документооборота // В сборнике: Информационная безопасность. Сборник статей конференции. 2019. С. 61–66.
5. Иванов А. И., Безяев А. В., Качайкин Е. И., Елфимов А. В. Искусственный интеллект: автоматизированный нейросетевой анализ «мертвой» подписи под документами на бумажных носителях // В сборнике: Безопасность информационных технологий. Сборник научных статей по материалам II Всероссийской научно-технической конференции. Пенза, 2020. С. 90–96.
6. Соловьев А. В. Проблема определения электронного документа долговременного хранения // Информационные технологии и вычислительные системы. 2022. № 1. С. 47–54.
7. Тали Д. И. Модели электронного документа в рамках парадигмы «Индустрия 4.0» // Управление большими системами. 2025. № 115. С. 66–99.
8. Ульянова Н. Д. Чат-боты в системах электронного документооборота // Вестник образовательного консорциума Среднерусский университет. Информационные технологии. 2023. № 2(22). С. 14–19.
9. Ковалева Н. Н., Ересько П. В., Изотова В. Ф. Проблемы и перспективы использования искусственного интеллекта в системах электронного документооборота // Вестник Воронежского государственного университета. Серия: Право. 2023. № 4(55). С. 87–92.
10. Язов Ю. К., Авсентьев А. О. Проблемные вопросы создания многоагентных систем защиты информации от утечки по техническим каналам // Вестник Воронежского института МВД России. 2024. № 3. С. 86–97.
11. Шамсутдинов Р. Р., Васильев В. И., Вульфин А. М. Интеллектуальная система мониторинга информационной безопасности промышленного интернета вещей с использованием механизмов искусственных иммунных систем // Системная инженерия и информационные технологии. 2024. Т. 6. № 4(19). С. 14–31.
12. Боговик А. В., Сафиулов Д. М. Предложения по модернизации протокола мониторинга телекоммуникационного оборудования узла связи специального назначения // Телекоммуникации и связь. 2025. № 2(5). С. 53–64.
13. Макаренко С. И. Информационный конфликт системы связи с системой дестабилизирующих воздействий. Часть. I: Концептуальная модель конфликта с учетом ведения разведки, физического, радиоэлектронного и информационного поражения средств связи // Техника радиосвязи. 2020. Выпуск 2 (45). С. 104–117.
14. Гончаров В. В., Мишенина О. В. Защита информации в автоматизированных системах: концептуально-математические аспекты // Правовая информатика. 2024. № 3. С. 43–57.
15. Махов Д. С. Повышение устойчивости управления параметрами функционирования пространственно - распределенных радиотехнических систем робототехнических комплексов на основе нечетких множеств // Вопросы оборонной техники. Серия 16: Технические средства противодействия терроризму. 2020. № 5–6 (143–144). С. 36–44.

16. Лепешкин О. М., Остроумов О. А., Синюк А. Д., Черных И. С. Проблема обеспечения функциональной устойчивости и непрерывности функционирования системы связи // Вестник компьютерных и информационных технологий. 2023. Т. 20. № 4(226). С. 16–26.
17. Волкова В. Н., Логинова А. В., Леонова А. Е., Черный А. Ю. Закономерности теории систем: состояние исследований и применения // В сборнике: Системный анализ в проектировании и управлении. Сборник научных трудов XXVI Международной научно-практической конференции. В 3-х частях. Санкт-Петербург. 2023. С. 65–74.
18. Калинин В. И., Юсупов Р. М., Соколов Б. В. Междисциплинарное взаимодействие и развитие теории систем, кибернетики и информатики // В сборнике: Системный анализ в проектировании и управлении. Сборник научных трудов XXVI Международной научно-практической конференции. В 3-х частях. Санкт-Петербург. 2023. С. 7–13.
19. Новиков Д. А. Принцип декомпозиции в задачах управления организационно-техническими системами // В сборнике: Математическая теория управления и ее приложения (МТУиП-2020). Материалы конференции. Государственный научный центр Российской Федерации АО «Концерн «ЦНИИ «Электроприбор». Санкт-Петербург. 2020. С. 256–259.
20. Тали Д. И. Принцип целостности и интегративности в формировании электронного документа // Правовая информатика. 2022. № 3. С. 72–83.
21. Тали Д. И., Финько О. А. и др. Способ обеспечения интегративной целостности электронного документа // Патент на изобретение RU 2812304, опубл. 29.01.2024, бюл. № 4.
22. Тали Д. И., Финько О. А., Диченко С. А. Способ формирования и контроля целостности многомерной структуры электронных документов // Патент на изобретение RU 2840783, опубл. 28.05.2025, бюл. № 16.
23. Тали Д. И., Финько О. А. Способ и система распределенного контроля целостности электронных документов при вероятной компрометации ключей подписи // Патент на изобретение RU 2844401, опубл. 29.07.2025, бюл. № 22.

CONCEPTUAL MODEL OF FUNCTIONING DIGITAL DOCUMENT MANAGEMENT SYSTEMS WITHIN THE FRAMEWORK OF THE «INDUSTRY 4.0» PARADIGM

Tali D. I.[17], Finko O. A.[18]

Keywords: *content, metadata, digital document, principles of digitalization, intelligent agents, digital transformation of document flow, system integrity.*

The purpose of the study *is to formalize the process of functioning of a digital document management system, which includes a document management system and data source systems. Implementation of the proposed approach in order to form a digital information management infrastructure that meets the basic provisions of the «Industry 4.0» concept.*

Research methods: *application of the methodology of system analysis to the conditions of digitalization of structurally complex systems using the example of electronic document management.*

The result of the research: *a conceptual model of the functioning of a digital document management system has been developed, taking into account such characteristics of promising digital systems as autonomy, distribution, and intelligence. A system of indicators and evaluation criteria has been introduced in order to improve the quality of information interaction between the structural divisions of organizations operating such an infrastructure.*

Scientific novelty: *a conceptual model of the functioning of a digital document management system is presented and substantiated, based on the hierarchical decomposition of its structure, taking into account the relationship between the levels of interaction of the system under study. The proposed approach in the context of digital transformation makes it possible to ensure the intended purpose (integrity) of the system for a given period of time when exposed to destabilizing factors at any of its levels.*

References

1. Shvab K. Chetvertaya promyshlennaya revolyutsiya. – M: Eksmo, 2021. 208 s.
2. Upravleniye dokumentami v tsifrovoy ekonomike: organizatsiya, reglamentatsiya, realizatsiya / M. V. Larin, N. G. Surovtseva, Ye. V. Terent'yeva, V. F. Yankovaya / Pod red. M.V. Larina – M.: RGGU, 2021. 242 s.
3. Larin M. V. Elektronnyye dokumenty: teoreticheskiye aspekty // Samarskiy arkhivist. 2021. № 2. S. 3–9.
4. Yeliseyev N. I., Tali D. I. Problemy i perspektivy razvitiya sistem yuridicheski znachimogo elektronnogo dokumentooborota // V sbornike: Informatsionnaya bezopasnost'. Sbornik statey konferentsii. 2019. S. 61–66.
5. Ivanov A. I., Bezyayev A. V., Kachaykin Ye. I., Yelfimov A. V. Iskusstvennyy intellekt: avtomatizirovannyy neyrosetevoy analiz «mertvoy» podpisi pod dokumentami na bumazhnykh nositelyakh // V sbornike: Bezopasnost' informatsionnykh tekhnologiy. Sbornik nauchnykh statey po materialam II Vserossiyskoy nauchno-tekhnicheskoy konferentsii. Penza, 2020. S. 90–96.
6. Solov'yev A. V. Problema opredeleniya elektronnogo dokumenta dolgovremennogo khraneniya // Informatsionnyye tekhnologii i vychislitel'nyye sistemy. 2022. № 1. S. 47–54.

17 Dmitry Tali, candidate of technical sciences, doctoral candidate of a special department, Krasnodar Higher Military Orders of Zhukov and the October Revolution Red Banner School named after General of the Army S. M. Shtemenko, Krasnodar, Russia. E-mail: dimatali@mail.ru

18 Oleg Finko, Dr.Sc., Professor, Professor of a special department, Krasnodar Higher Military Orders of Zhukov and the October Revolution Red Banner School named after General of the Army S.M. Shtemenko, Academic Advisor of the Russian Academy of Rocket and Artillery Sciences (RARAN), Krasnodar, Russia. E-mail: ofinko@yandex.ru. Web: http://www.mathnet.ru/person40004. ORCID 0000-0002-7376-271

7. Tali D. I. Modeli elektronnogo dokumenta v ramkakh paradigmy «Industriya 4.0» // Upravleniye bol'shimi sistemami. 2025. № 115. S. 66-99.
8. Ul'yanova N. D. Chat-boty v sistemakh elektronnogo dokumentooborota // Vestnik obrazovatel'nogo konsortsiuma Srednerusskiy universitet. Informatsionnyye tekhnologii. 2023. № 2(22). S. 14–19.
9. Kovaleva N. N., Yeres'ko P. V., Izotova V. F. Problemy i perspektivy ispol'zovaniya iskusstvennogo intellekta v sistemakh elektronnogo dokumentooborota // Vestnik Voronezhskogo gosudarstvennogo universiteta. Seriya: Pravo. 2023. № 4(55). S. 87–92.
10. Yazov Yu. K., Avsent'yev A. O. Problemnyye voprosy sozdaniya mnogoagentnykh sistem zashchity informatsii ot utechki po tekhnicheskim kanalam // Vestnik Voronezhskogo instituta MVD Rossii. 2024. № 3. S. 86–97.
11. Shamsutdinov R. R., Vasil'yev V. I., Vul'fin A. M. Intellektual'naya sistema monitoringa informatsionnoy bezopasnosti promyshlennogo interneta veshchey s ispol'zovaniyem mekhanizmov iskusstvennykh immunnykh sistem // Sistemnaya inzheneriya i informatsionnyye tekhnologii. 2024. T. 6. № 4(19). S. 14–31.
12. Bogovik A. V., Safiulov D. M. Predlozheniya po modernizatsii protokola monitoringa telekommunikatsionnogo oborudovaniya uzla svyazi spetsial'nogo naznacheniya // Telekommunikatsii i svyaz'. 2025. № 2(5). S. 53–64.
13. Makarenko S. I. Informatsionnyy konflikt sistemy svyazi s sistemoy destabiliziruyushchikh vozdeystviy. Chast'. I: Kontseptual'naya model' konflikta s uchetom vedeniya razvedki, fizicheskogo, radioelektronnogo i informatsionnogo porazheniya sredstv svyazi // Tekhnika radiosvyazi. 2020. Vypusk 2(45). S. 104–117.
14. Goncharov V. V., Mishenina O. V. Zashchita informatsii v avtomatizirovannykh sistemakh: kontseptual'no-matematicheskiye aspekty // Pravovaya informatika. 2024. № 3. S. 43–57.
15. Makhov D. S. Povysheniye ustoychivosti upravleniya parametrami funktsionirovaniya prostranstvenno – raspredelennykh radiotekhnicheskikh sistem robototekhnicheskikh kompleksov na osnove nechetkikh mnozhestv // Voprosy oboronnoy tekhniki. Seriya 16: Tekhnicheskiye sredstva protivodeystviya terrorizmu. 2020. № 5-6 (143-144). S. 36–44.
16. Lepeshkin O. M., Ostroumov O. A., Sinyuk A. D., Chernykh I. S. Problema obespecheniya funktsional'noy ustoychivosti i nepreryvnosti funktsionirovaniya sistemy svyazi // Vestnik komp'yuternykh i informatsionnykh tekhnologiy. 2023. T. 20. № 4(226). S. 16–26.
17. Volkova V. N., Loginova A. V., Leonova A. Ye., Chernyy A. Yu. Zakonomernosti teorii sistem: sostoyaniye issledovaniy i primeneniya // V sbornike: Sistemnyy analiz v proyektirovanii i upravlenii. Sbornik nauchnykh trudov XXVI Mezhdunarodnoy nauchno-prakticheskoy konferentsii. V 3-kh chastyakh. Sankt-Peterburg. 2023. S. 65–74.
18. Kalinin V. I., Yusupov R. M., Sokolov B. V. Mezhdistsiplinarnoye vzaimodeystviye i razvitiye teorii sistem, kibernetiki i informatiki // V sbornike: Sistemnyy analiz v proyektirovanii i upravlenii. Sbornik nauchnykh trudov XXVI Mezhdunarodnoy nauchno-prakticheskoy konferentsii. V 3-kh chastyakh. Sankt-Peterburg. 2023. S. 7–13.
19. Novikov D. A. Printsip dekompozitsii v zadachakh upravleniya organizatsionno-tekhnicheskimi sistemami // V sbornike: Matematicheskaya teoriya upravleniya i yeye prilozheniya (MTUiP-2020). Materialy konferentsii. Gosudarstvennyy nauchnyy tsentr Rossiyskoy Federatsii AO «Kontsern «TSNII «Elektropribor». Sankt-Peterburg. 2020. S. 256–259.
20. Tali D. I. Printsip tselostnosti i integrativnosti v formirovanii elektronnogo dokumenta // Pravovaya informatika. 2022. № 3. S. 72–83.
21. Tali D. I., Fin'ko O. A. i dr. Sposob obespecheniya integrativnoy tselostnosti elektronnogo dokumenta // Patent na izobreteniye RU 2812304, opubl. 29.01.2024, byul. № 4.
22. Tali D. I., Fin'ko O. A., Dichenko S. A. Sposob formirovaniya i kontrolya tselostnosti mnogomernoy struktury elektronnykh dokumentov // Patent na izobreteniye RU 2840783, opubl. 28.05.2025, byul. № 16.
23. Tali D. I., Fin'ko O. A. Sposob i sistema raspredelennogo kontrolya tselostnosti elektronnykh dokumentov pri veroyatnoy komprometatsii klyuchey podpisi // Patent na izobreteniye RU 2844401, opubl. 29.07.2025, byul. № 22.

ПОСТКВАНТОВЫЙ АЛГЕБРАИЧЕСКИЙ АЛГОРИТМ ЭЦП С ТРЕМЯ СКРЫТЫМИ ГРУППАМИ

Молдовян А. А.[1]

DOI: 10.21681/2311-3456-2025-5-78-87

Цель работы: *создание дополнительных предпосылок для разработки постквантового стандарта на алгоритмы ЭЦП, основанные на вычислительной сложности решения больших систем нелинейных уравнений (БСНУ) в конечном поле.*

Метод исследования: *применение трех скрытых коммутативных групп, элементы каждой из которых некоммутативны с элементами другой, для реализации усиленной рандомизации подписи в алгебраических схемах ЭЦП, стойкость которых базируется на вычислительной трудности решения БСНУ в простом конечном поле $GF(p)$. Вычисление подгоночного элемента подписи в виде матрицы* $\mathbf{S}$ *в зависимости от трех попарно некоммутативных матриц, выбираемых случайным образом из скрытых групп. Применение конечных алгебры матриц 3×3 и 5×5, заданных над полем $GF(p)$ с 64-битным и 40-битным простым порядком p.*

Результаты исследования: *предложен новый механизм усиленной рандомизации подгоночного элемента подписи и разработан алгебраический алгоритм ЭЦП, перспективный в качестве прототипа практичного постквантового стандарта ЭЦП. Выбор случайных матриц из скрытых групп задается посредством возведения генераторов соответствующих скрытых групп в случайные степени, вычисляемые в зависимости от параметров рандомизации и рандомизирующего элемента ЭЦП. Впервые для повышения уровня стойкости к потенциальным атакам на основе альтернативных секретных ключей указанные степени вычисляются как решение системы линейных уравнений. Представлены оценки стойкости к прямой атаке и атаке на основе известных подписей. Приведено сравнение параметров разработанного алгоритма ЭЦП с известными алгоритмами, использующими трудность решения БСНУ, и обсуждаются способы повышения его производительности.*

Научная и практическая значимость *полученных результатов заключается в создании новой предпосылки для обоснования выбора алгебраических алгоритмов ЭЦП со скрытыми группами в качестве основы для разработки практичного постквантового стандарта ЭЦП, состоящей в повышении уровня стойкости к потенциальным атакам на основе эквивалентных ключей.*

Ключевые слова: *конечная алгебра матриц; ассоциативная алгебра; постквантовая криптография; вычислительно трудная задача; скрытая коммутативная группа; цифровая подпись; рандомизация подписи.*

Введение

В ходе технического прогресса возрастает значение электронной цифровой подписи (ЭЦП), которая имеет важную роль в современных информационных технологиях, применяемых в различных сферах общественной деятельности. Существующие стандарты на алгоритмы ЭЦП основаны на вычислительной трудности задачи факторизации (ЗФ) и задачи дискретного логарифмирования (ЗДЛ). Последние результаты в области квантовых вычислений дают основание для прогнозирования появления в ближайшем будущем практически доступного квантового компьютера, на котором ЗФ и ЗДЛ будут решаться за полиномиальное время. С этого момента криптографические алгоритмы с открытым ключом, стойкость которых ограничена сверху вычислительной сложностью ЗФ и ЗДЛ, включая действующие стандарты ЭЦП, перестают быть безопасными. Такая ситуация делает актуальной задачу разработки постквантовых двухключевых криптоалгоритмов и стандартов на их основе, которые являются стойкими к атакам с использованием квантовых компьютеров. Исследования в этом направлении относятся к постквантовой криптографии, тематика которой привлекает внимание и усилия многих исследователей [1–3]. Для обеспечения стойкости к атакам с использованием квантовых компьютеров разрабатываемые криптоалгоритмы должны базироваться на вычислительно трудных задачах, отличных от ЗФ и ЗДЛ. Переход к задачам другого типа обусловил существенный рост размеров открытого ключа и подписи, что поднимает вопрос о практичности использования постквантовых криптоалгоритмов с открытым ключом (в смысле снижения объемов оперативной памяти и количества вычислительных операций, т.е. издержек их практического применения).

В области постквантовой криптографии значительное внимание уделяется разработке алгоритмов открытого шифрования и ЭЦП на трудно обратимых отображениях с секретной лазейкой [4–8]. Стойкость таких алгоритмов основана на вычислительной сложности решения больших систем нелинейных уравнений (БСНУ), заданных над конечным полем

1 Молдовян Александр Андреевич, доктор технических наук, главный научный сотрудник лаборатории проблем компьютерной безопасности Санкт-Петербургского федерального исследовательского центра Российской академии наук. ORCID: https://orcid.org/0000-0001-5480-6016. E-mail: maa1305@yandex.ru

сравнительно малого порядка. Для нахождения решений БСНУ квантовый компьютер не является эффективным, что обеспечивает постквантовую стойкость разрабатываемых криптоалгоритмов. Несмотря на достаточно большое число разработанных алгоритмов ЭЦП такого типа, не удалось решить проблему чрезвычайно большого размера открытого ключа (от десятков килобайт до нескольких мегабайт в зависимости от уровня стойкости). Даже сравнительно новая парадигма построения криптоалгоритмов на трудно обратимых отображениях с секретной лазейкой дает только ограниченное решение данной проблеме [9–11].

Принципиально новый подход к использованию вычислительной сложности решения БСНУ связан с разработкой алгебраических алгоритмов со скрытой коммутативной группой [12–15]. Эти алгоритмы относятся к рандомизированным схемам ЭЦП, в которых подпись $(e, \mathbf{S})$ генерируется в зависимости от предварительно генерируемых случайных натуральных чисел и включает два элемента: 1) случайное число e (рандомизирующий элемент ЭЦП) и 2) элемент конечной алгебры $\mathbf{S}$, используемой в качестве алгебраического носителя (подгоночный элемент ЭЦП). При этом проверочное уравнение включает $\mathbf{S}$ два или более раза в качестве множителя. Формируемое в процессе генерации подписи значение e является уникальным, обусловливая уникальность значения $\mathbf{S}$. Однако в отличии от рандомизированных схем ЭЦП, стойкость которых основана на вычислительной сложности ЗДЛ или скрытой ЗДЛ [16–18], для рассматриваемых алгоритмов рандомизация ЭЦП является ограниченной (в смысле того, что $\mathbf{S}$ принимает значения из подмножества Ψ обратимых элементов алгебры, используемой в качестве алгебраического носителя, мощность которого ψ намного меньше порядка Ω мультипликативной группы алгебры).

Как показано в статьях [19, 20], ограниченность рандомизации создает предпосылки для осуществления атаки на основе известных подписей (атака A1), которая состоит в вычислении элементов секретного ключа по некоторому набору известных подписей путем решения некоторой БСНУ, включающие уравнения, записываемые по формуле для вычисления подгоночного значения $\mathbf{S}$. При этом прямая атака на алгебраические алгоритмы со скрытой группой состоит в решении БСНУ, которая составляется по формулам генерации элементов открытого ключа в зависимости от элементов секретного ключа (атака A2). Атака A1 позволяет вычислить только часть секретного ключа, однако это позволяет существенно уменьшить вычислительную сложность атаки A2, т.е. снизить уровень стойкости алгоритма ЭЦП. Если сложность атаки A1 превышает сложность атаки A2, то используемый механизм рандомизации может рассматриваться приемлемым (достаточным) для данного конкретного алгебраического алгоритма ЭЦП (критерий достаточной полноты рандомизации по [21, 22]). В качестве количественной меры достигаемого уровня рандомизации в заданном алгоритме можно принять значение отношения $\log_2\Omega$ к разности $\log_2\Omega - \log_2\Psi$, а термин «усиление рандомизации» – трактовать как повышение уровня рандомизации при разработке нового механизма рандомизации ЭЦП. Следует отметить, что выполнение критерия достаточной полноты рандомизации не привязано к некоторому пороговому значению уровня рандомизации и требует рассмотрения вычислительной сложности атак A1 и A2 для конкретного алгебраического алгоритма ЭЦП. Тем не менее, усиление рандомизации можно трактовать как «приближение» к выполнению критерия полноты рандомизации.

Способ получения оценки стойкости к атаке A1 детально рассматривается в работах [21, 22], в которых в частности показано, что обеспечение полноты рандомизации подгоночного элемента $\mathbf{S}$ путем его вычисления в зависимости от случайного элемента $\mathbf{V}$, выбираемого из мультипликативной группы используемого алгебраического носителя, не приводит автоматически к достаточной полноте рандомизации ЭЦП, поскольку для корректности схемы ЭЦН элемент $\mathbf{V}$ должен входить в качестве множителя также и в формулу для вычисления вектора-фиксатора $\mathbf{R}$ при генерации ЭЦП. При этом значение $\mathbf{R}$ вычисляется в процессе верификации ЭЦП (т.е. $\mathbf{V}$ для данной известной подписи входит в качестве неизвестной в два независимых векторных уравнения). Для усиления рандомизации в схемах ЭЦП с удвоенным проверочным уравнением в работе [21] впервые применены две скрытые группы и вычисление элемента $\mathbf{S}$ в зависимости от двух случайных элементов, выбираемых из разных скрытых групп. В работе [22] эта идея использована для разработки алгоритма ЭЦП с одним проверочным уравнением, удовлетворяющего упомянутому критерию достаточной полноты рандомизации.

Механизм рандомизации, включающий вычисление элемента $\mathbf{S}$ в зависимости от случайного значения $\mathbf{V}$, при разработке алгоритма ЭЦП требует использования удвоенного проверочного уравнения, приводящего к существенному снижению производительности процедур генерации и верификации ЭЦП. По этой причине с практической точки зрения больший интерес представляют механизмы рандомизации [22], реализуемые в рамках схем ЭЦП с одним поверочным уравнением, в которое элемент $\mathbf{S}$ входит многократно (два и более раза).

В работах предложены механизмы рандомизации, использующие две скрытые циклические группы, элементы одной из которых некоммутативны с элементами другой. Данный механизм может быть представлен следующей формулой для вычисления подгоночного элемента подписи $\mathbf{S}$:

$$\mathbf{S} = \mathbf{D}\mathbf{P}^{b}\mathbf{G}^{n}\mathbf{F}, \quad (1)$$

где $\mathbf{D}$ и $\mathbf{F}$ – элементы секретного ключа; $\mathbf{P}$ и $\mathbf{G}$ – генераторы скрытых групп; b и n – случайные натуральные степени. Использование случайного выбора элементов из двух скрытых групп обеспечивает существенное увеличение параметра ψ, за счет чего повышается уровень рандомизации. В разработанных на основе формулы (1) алгоритмах ЭЦП [22], использующих четырехмерные конечные некоммутативные ассоциативные алгебры в качестве алгебраического носителя, выполняется критерий достаточной полноты рандомизации при обеспечении уровня стойкости 2^{100}. Разработка алгоритмов ЭЦП с двумя скрытыми группами потребовала использования одного или двух дополнительных подгоночных элементов в виде натуральных чисел, входящих в проверочное уравнение как степени при некоторых элементах открытого ключа, присутствующих в уравнении верификации ЭЦП.

Логическим расширение такого приема усиления рандомизации является переход к механизму, включающему случайный выбор элементов из трех скрытых циклических групп и описываемому следующей формулой:

$$\mathbf{S} = \mathbf{D}\mathbf{P}^{b}\mathbf{G}^{n}\mathbf{J}^{d}\mathbf{F}, \quad (2)$$

где $\mathbf{D}$ и $\mathbf{F}$ – элементы секретного ключа; $\mathbf{P}$, $\mathbf{G}$ и $\mathbf{J}$ – генераторы скрытых групп, таких, что элементы каждой из них некоммутативны с элементами двух других; b, n и d – случайные натуральные степени. Достаточно очевидно, что это расширение дает существенное увеличение уровня рандомизации и потенциальную возможность разработки алгоритмов ЭЦП, удовлетворяющих критерию достаточной полноты рандомизации подписи при обеспечиваемом уровне стойкости 2^{192} и более и сохраняющих высокую производительность процедур генерации и верификации ЭЦП. Однако, разработка алгоритмов ЭЦП на основе формулы (2) требует применения новых конструктивных приемов.

Целью настоящей работы является создание дополнительных предпосылок, дающих обоснование целесообразности разработки проекта постквантового стандарта ЭЦП на основе алгебраических алгоритмов ЭЦП, основанных на вычислительной сложности решения БСНУ, что потенциально позволит принять практичный с точки зрения размеров открытого ключа и подписи и производительности стандарт ЭЦП для применения в постквантовую эру. В качестве таких дополнительных предпосылок предполагается разработка алгоритма, обладающего стойкость 2^{192} и 2^{256} в различных его модификациях, и расширение множества приемов построения алгоритмов ЭЦП со скрытыми группами.

Для достижения этой цели применяется механизм рандомизации, описываемый формулой (2), и применяются новые конструктивные приемы разработки алгоритма ЭЦП, реализующего этот механизм при выполнении критерия достаточной полноты рандомизации.

Основные приемы, используемые при разработке алгоритма ЭЦП

1. В качестве основного механизма для предотвращения подделки подписи на основе решения проверочного уравнения относительно неизвестного значения $\mathbf{S}$ предполагается использовать значение хеш-функции Φ от $\mathbf{S}$ в качестве степени одной из операций экспоненцирования, присутствующих в проверочном уравнении. Данный прием ранее использован в работах [21, 22]. Его алгоритмическая реализация предполагает использование вспомогательного подгоночного элемента подписи в виде натурального числа s, поскольку значение степени $\rho = \Phi(\mathbf{S})$ является псевдослучайным и становится известным только после фиксирования значения подгоночного элемента $\mathbf{S}$.

2. Задается двухкратное вхождение в проверочное уравнение элемента $\mathbf{S}$ как множителя первой степени. Многократное вхождение в уравнение верификации элемента $\mathbf{S}$ без его возведения в большую степень, не может обеспечить приемлемую стойкость к подделке подписи, поэтому использование второго конструктивного приема становится возможным только одновременно с первым.

3. В используемой для вычисления значения $\mathbf{S}$ формуле (2) в качестве генератора циклической группы $\mathbf{G}$ выбирается произведение $\mathbf{J}^{x}\mathbf{P}^{z}$ (где $\mathbf{P}$ и $\mathbf{J}$ – обратимые матрицы, удовлетворяющие условию $\mathbf{JP} \neq \mathbf{PJ}$) с секретными степенями x и z. Третий прием требует выполнения переборной процедуры, состоящей в подборе таких степеней x и z, при которых порядок матрицы $\mathbf{G} = \mathbf{J}^{x}\mathbf{P}^{z}$ имеет достаточно большой размер. Благодаря тому, что доля матриц большого порядка является преобладающей, указанная процедура незначительно увеличивает вычислительную сложность процесса формирования открытого и секретного ключей. Выполнимость требуемых условий $\mathbf{GP} \neq \mathbf{PG}$ и $\mathbf{GJ} \neq \mathbf{JG}$ является достаточно очевидной.

Используемый алгебраический носитель

В известных алгоритмах ЭЦП со скрытой группой в качестве алгебраического носителя используются m-мерные конечные некоммутативные ассоциативные алгебры (КНАА), обычно заданные над конечным

полем нечетной характеристики $GF(p)$, а в отдельных случаях характеристики два [23]. При этом более детальный анализ стойкости к атакам А1 и А2 выполнен для случая использования четырехмерных КНАА, поскольку в этом анализе существенно используется возможность описания координат векторов из скрытой группы по координатам некоторого представителя скрытой группы и значениям η ($\eta < m$) скалярных переменных. Для задания КНАА произвольных четных размерностей известны унифицированные способы их задания [24, 25], причем с ростом значения размерности m ожидается существенное повышение стойкости алгебраических алгоритмов, основанных на вычислительной трудности решения БСНУ. Однако, для размерностей $m > 4$ строение КНАА, заданных по способам [24, 25], является мало изученным, что не позволяет получить приемлемых оценок стойкости.

В работе [26] возможность разработки алгоритма ЭЦП на КНАА большой размерности $m = 9$ и получения приемлемых оценок стойкости, в качестве КНАА используется конечная алгебра матриц 3×3, заданная над простым полем $GF(p)$ с 80-битной характеристикой p. Действительно, легко видно, что умножение матриц размера $\mu\times\mu$ может быть представлено как умножение векторов размерности $m = \mu^2$, заданное по таблице умножения базисных векторов специального вида (см. [26] для случая $\mu = 3$).

В настоящей работе в качестве алгебраического носителя разрабатываемого постквантового алгоритма ЭЦП используется конечная алгебра матриц, заданная над полем $GF(p)$ с 64-битной характеристикой p и алгоритмы генерации матриц требуемого порядка, описанные в [26]. Для задания более высокого уровня стойкости предполагается реализация разработанного алгоритма на алгебре матриц 5×5, заданной над полем $GF(p)$ с 40-битной характеристикой p. Более высокий уровень стойкости обеспечивается существенным возрастанием числа уравнений в БСНУ, задача решения которой возникает в атаках А1 и А2.

Далее в статье вместо термина «элементы матриц» будем использовать термин «координаты матриц», резервируя слово «элементы» для использования терминов «элементы (конечного) поля» и «элементы открытого (секретного) ключа».

1. Разработанный алгоритм ЭЦП

Формирование секретного и открытого ключей

Предлагается разработка двух вариантов алгоритма ЭЦП, отличающихся тем, что в первом варианте используется в качестве алгебраического носителя конечная алгебра матриц 3×3, заданная над полем $GF(p)$ с 64-битной характеристикой p, а во втором – алгебра матриц 5×5, заданная над полем $GF(p)$ с 40-битным простым числом p. Для первого варианта используется такое простое число p, что число $q = p^2 + p + 1$ также является, а для второго варианта – простое число p, при котором имеем простое значение $q = p^4 + p^3 + p^2 + p + 1$. Учитывая формулу для порядка Ω мультипликативной группы алгебры матриц размера $n\times n$, выражающую Ω через значения p и n (см., например, формулу (5) в [26]), легко показать существование матриц порядка q первого и второго случаев. Из указанной формулы легко установить возможные значения порядков матриц. При этом в [26] показано, что вычислительная сложность процедур генерации нужного простого значения p и матрицы порядка q для первого случая является сравнительно низкой. По аналогии с рассуждениями [26] легко показать, что во втором случае вычислительная сложность упомянутых процедур также является приемлемой для их использования в алгоритмах генерации секретного и открытого ключей. Обе версии разработанного алгоритма описываются одинаковыми математическими формулами.

Для генерации секретного ключа выбираются случайные натуральные числа $u < q$, $v < q$, $w < q$, $x < q$, $y < q$ и $z < q$ и случайные обратимые (невырожденные) нескалярные матрицы $\mathbf{A}$, $\mathbf{B}$, $\mathbf{D}$, $\mathbf{F}$, $\mathbf{J}$ и $\mathbf{P}$, которые являются попарно некоммутативными, а матрицы $\mathbf{J}$ и $\mathbf{P}$ имеют одинаковый порядок q. Размер секретного ключа равен 528 (870) байт для первой (второй) версии алгоритма ЭЦП. При генерации ЭЦП используется вспомогательная секретная матрица $\mathbf{G}$, которая вычисляется по следующей формуле

$$\mathbf{G} = \mathbf{J}^x\mathbf{P}^z. \quad (3)$$

Формула (3) учитывается при генерации случайных положительных значений $x < q$ и $z < q$. Легко показать, что для любой пары значений $x > 0$ и $z > 0$ вычисляемая матрица $\mathbf{G}$ будет некоммутативным с каждой из матриц $\mathbf{J}$ и $\mathbf{P}$. При генерации различных пар значений (x, z) определяется значение порядка $\omega(\mathbf{G})$ матрицы $\mathbf{G}$ и в качестве элементов секретного ключа фиксируются значения x и z, при которых $\omega(\mathbf{G})$ является нечетным числом.

Открытый ключ формируется в виде набора из семи матриц ($\mathbf{Q}$, $\mathbf{U}$, $\mathbf{Y}$, $\mathbf{Z}$, $\mathbf{T}_1$, $\mathbf{T}_2$, $\mathbf{T}_3$), элементы которого вычисляются по следующим формулам:

$$\mathbf{Q} = \mathbf{B}\mathbf{J}^z\mathbf{B}^{-1};\ \mathbf{U} = \mathbf{F}^{-1}\mathbf{J}\mathbf{F};\ \mathbf{Y} = \mathbf{A}\mathbf{P}\mathbf{A}^{-1};\ \mathbf{Z} = \mathbf{D}\mathbf{P}^x\mathbf{D}^{-1}; \quad (4)$$

$$\mathbf{T}_1 = \mathbf{A}^{-1}\mathbf{P}^u\mathbf{D}^{-1};\ \mathbf{T}_2 = \mathbf{F}^{-1}\mathbf{J}^u\mathbf{P}^w\mathbf{D}^{-1};\ \mathbf{T}_3 = \mathbf{F}^{-1}\mathbf{1}\mathbf{J}^v\mathbf{B}^{-1}. \quad (5)$$

Суммарный размер открытого ключа равен 504 (875) байт для первой (второй) версии алгоритма ЭЦП. В формулах (3) и (4) можно было бы использовать различные значения степеней x и z, однако этот вариант несколько увеличивает размер секретного ключа, но не влияет на оценку стойкости алгоритма.

В процедурах генерации и верификации ЭЦП используется некоторая специфицированная коллизионно стойкая 256-битная хеш-функции Φ, которая является частью рассматриваемого алгоритма ЭЦП.

Алгоритм генерации ЭЦП

Процедура генерации ЭЦП к документу M включает следующие шаги:

1. Сгенерировать случайные натуральные числа $k < q$, $t < q$ и $r < q$ и вычислить значение рандомизирующей матрицы $\mathbf{R}$ по формуле:

$$\mathbf{R} = \mathbf{A}\mathbf{P}^{t}\mathbf{G}^{k}\mathbf{J}^{r}\mathbf{B}^{-1}. \quad (6)$$

2. Вычислить хеш-значение от матрицы $\mathbf{R}$ с присоединенным к нему документом M: $e = e_1||e_2 = \Phi(M, \mathbf{R})$, где 256-битное хеш-значение e представлено в виде конкатенации двух 128-битных натуральных чисел e_1 и e_2.

3. Вычислить натуральное число d:

$$d = x - e_2 - u \bmod q. \quad (7)$$

4. Вычислить натуральные числа σ и b (удовлетворяющие одновременно уравнениям $\sigma + u + xe_1 + b = t$ и $w + x\sigma + b = z \bmod q$):

$$\sigma = (1 - x)^{-1}(t - xe_1 - u - z + w) \bmod q; \quad (8)$$

$$b = z - w - x\sigma \bmod q. \quad (9)$$

5. Вычислить натуральное число n:

$$n = 2^{-1}(k - 1) \bmod \omega(\mathbf{G}). \quad (10)$$

6. Вычислить матрицу $\mathbf{S} = \mathbf{D}\mathbf{P}^{b}\mathbf{G}^{n}\mathbf{J}^{d}\mathbf{F}$.

7. Вычислить хеш-значение $\rho = \rho_1||\rho_2 = \Phi(\mathbf{S}, e)$, где значение ρ представлено в виде конкатенации двух 128-битных натуральных чисел ρ_1 и ρ_2.

8. Вычислить вспомогательный подгоночный элемент ЭЦП в виде натурального числа s:

$$s = z^{-1}(r - d - v - \rho_1) - \rho_2 \bmod q. \quad (11)$$

Сгенерированная подпись представляет собой четверку значений $(e, \sigma, s, \mathbf{S})$ с суммарным размером 136 (197) байт для первой (второй) версии алгоритма ЭЦП. Вычислительную трудность алгоритма генерации ЭЦП можно оценить как три операции возведения в степень в конечной алгебре матриц (вычисление матриц $\mathbf{P}^{t}$, $\mathbf{G}^{k}$, $\mathbf{J}^{d}$), что составляет ≈15500 (90000) операций умножения в поле $GF(p)$ для первой (второй) версии алгоритма. Учитывая, что в первой версии алгоритма используется 64-битное простое число p, а во второй – 40-битое, легко видеть, что производительность второй версии примерно в 2,3 раза меньше.

Алгоритм верификации ЭЦП

Верификация подписи $(e, \sigma, s, \mathbf{S})$ к документу M выполняется по открытому ключу $(\mathbf{Q}, \mathbf{U}, \mathbf{Y}, \mathbf{Z}, \mathbf{T}_1, \mathbf{T}_2, \mathbf{T}_3)$ с использованием следующего алгоритма:

1. Вычислить хеш-значение $\rho = \rho_1||\rho_2 = \Phi(\mathbf{S}, e)$.

2. Вычислить матрицу $\mathbf{R}'$ по следующему проверочному уравнению:

$$\mathbf{R}' = \mathbf{Y}^{\sigma}\mathbf{T}_1\mathbf{Z}^{e_1}\mathbf{S}\mathbf{U}^{e_2}\mathbf{T}_2\mathbf{Z}^{\sigma}\mathbf{S}\mathbf{U}^{\rho_1}\mathbf{T}_3\mathbf{Q}^{s+\rho_2}. \quad (12)$$

3. Вычислить хеш-функцию от матрицы $\mathbf{R}'$ с присоединенным к нему документом M: $\varepsilon = \varepsilon_1||\varepsilon_2 = \Phi(\mathbf{R}', M)$, где 256-битное хеш-значение представлено в виде конкатенации двух 128-битных чисел ε_1 и ε_2.

4. При справедливости равенств $\varepsilon_1 = e_1$ и $\varepsilon_2 = e_2$ ЭЦП делается вывод о подлинности подписи, в противном случае – о ложности.

Вычислительную сложность процедуры верификации ЭЦП можно оценить как шесть операций экспоненцирования в конечной алгебре матриц. что составляет ≈31000 (180000) операций умножения в поле $GF(p)$ для первой (второй) версии алгоритма.

Корректность разработанного алгоритма

Корректность предложенного алгоритма ЭЦП можно показать, выполняя подстановку в уравнении (12) матриц $\mathbf{Q}$, $\mathbf{U}$, $\mathbf{Y}$, $\mathbf{Z}$, $\mathbf{T}_1$, $\mathbf{T}_2$ и $\mathbf{T}_3$, выраженных через элементы секретного ключа по формулам (4) и (5):

$$\mathbf{R}' = (\mathbf{A}\mathbf{P}\mathbf{A}^{-1})^{\sigma}\,\mathbf{A}^{-1}\mathbf{P}^{u}\mathbf{D}^{-1}(\mathbf{D}\mathbf{P}^{x}\mathbf{D}^{-1})^{e_1} \times$$
$$\times (\mathbf{D}\mathbf{P}^{b}\mathbf{G}^{n}\mathbf{J}^{d}\mathbf{F})\,(\mathbf{F}^{-1}\mathbf{J}\mathbf{F})^{e_2}\,\mathbf{F}^{-1}\mathbf{J}^{u}\mathbf{P}^{w}\mathbf{D}^{-1}(\mathbf{D}\mathbf{P}^{x}\mathbf{D}^{-1})^{\sigma} \times$$
$$\times (\mathbf{D}\mathbf{P}^{b}\mathbf{G}^{n}\mathbf{J}^{d}\mathbf{F})(\mathbf{F}^{-1}\mathbf{J}\mathbf{F})^{\rho_2}\,\mathbf{F}^{-1}\mathbf{J}^{v}\mathbf{B}^{-1}(\mathbf{B}\mathbf{J}^{z}\mathbf{B}^{-1})^{s+\rho_2} =$$
$$= \mathbf{A}\mathbf{P}^{\sigma+u+xe_1+b}\mathbf{G}^{n}\mathbf{J}^{d+e_1+u} \times \mathbf{P}^{w+x\sigma+b}\mathbf{G}^{n}\mathbf{J}^{d+x+\rho_1+z(s+\rho_2)}\mathbf{B}^{-1}.$$

Заменяя в последнем уравнении значения d, σ, b, n и s правой частью соответствующих формул (7)–(11) и учитывая равенство (3), получаем:

$$\mathbf{R}' = \mathbf{A}\mathbf{P}^{t}\mathbf{G}^{n}\mathbf{J}^{x}\mathbf{P}^{z}\mathbf{G}^{n}\mathbf{J}^{r}\mathbf{B}^{-1} = \mathbf{A}\mathbf{P}^{t}\mathbf{G}^{2n+1}\mathbf{J}^{r}\mathbf{B}^{-1} =$$
$$= \mathbf{A}\mathbf{P}^{t}\mathbf{G}^{k}\mathbf{J}^{r}\mathbf{B}^{-1} = \mathbf{R}.$$

С учетом равенства $\mathbf{R} = \mathbf{R}'$ имеем $\varepsilon_1||\varepsilon_2 = \Phi(\mathbf{R}', M) = \Phi(\mathbf{R}, M) = e_1||e_2$, т. е. корректно сгенерированная подпись проходит процедуру верификации как подлинная подпись, что означает корректность разработанного алгоритма ЭЦП.

2. Обсуждение стойкости

В основе стойкости разработанного алгоритма лежит вычислительная сложность решения БСНУ. В случае прямой атаки (атака А2) БСНУ записывается по формулам (4) и (5), связывающем матрицы секретного ключа (неизвестные) с элементами открытого ключа. Если степени u, v, w, x, y и z принять как неизвестные, то будем иметь систему степенных и экспоненциальных матричных уравнений. Чтобы свести атаку А1 к решению системы, включающих только степенные матричные уравнения, в формулах (4) и (5) значения $\mathbf{J}^{z}$, $\mathbf{J}^{u}$ и $\mathbf{J}^{v}$ следует рассматривать как неизвестные матрицы, выбираемые из коммутативной скрытой группы, генерируемой матрицей $\mathbf{J}$,

а значения $\mathbf{P}^x$, $\mathbf{P}^u$ и $\mathbf{P}^w$ – как неизвестные матрицы, выбираемые из коммутативной скрытой группы, генерируемой матрицей $\mathbf{P}$. В статье [26] приводится обоснование, того, что координаты матриц 3×3, выбираемых из коммутативной группы могут быть описаны по координатам известной нескалярной матрицы (представителя коммутативной группы), содержащейся в этой группе, и $\eta = \mu$ ($\mu = 3$) скалярным переменным, т.е. толкование каждой из матриц $\mathbf{J}^z$, $\mathbf{J}^u$, $\mathbf{J}^v$. $\mathbf{P}^x$, $\mathbf{P}^u$ и $\mathbf{P}^w$ в качестве матричной неизвестной дает μ уникальных скалярных неизвестных при сведении решения системы матричных уравнений к системе скалярных уравнений (уравнений в поле $GF(p)$). По аналогии со случаем алгебры матриц 3×3 можно показать, что координаты матриц, содержащихся в заданной коммутативной группе конечной алгебры матриц 5×5 описываются по координатам нескалярного представителя этой группы и $\eta = \mu = 5$ скалярным переменным.

При таком толковании повышается степень матричных уравнений в БСНУ, однако повышением вычислительной сложности решения БСНУ за счет этого момента не будем принимать во внимание при выполнении оценки стойкости разработанного алгоритма ЭЦП (сложность решения БСНУ несущественно увеличивается за счет увеличения степени некоторых уравнений). Каждая из матричных неизвестных $\mathbf{A}$, $\mathbf{B}$, $\mathbf{D}$, $\mathbf{F}$, $\mathbf{J}$ и $\mathbf{P}$ дает μ^2 скалярных неизвестных, а каждая из матричных неизвестных $\mathbf{J}^z$, $\mathbf{J}^u$, $\mathbf{J}^v$. $\mathbf{P}^x$, $\mathbf{P}^u$ и $\mathbf{P}^w$ – μ скалярных неизвестных. Всего в скалярной БСНУ получаем $6\mu^2 + 6\mu$ неизвестных и $7\mu^2$ уравнений. Для случая, когда число уравнений превышает число неизвестных, сложность решения БСНУ может быть оценена по числу степенных уравнений, содержащихся в БСНУ. В этом случае с учетом оценок [4] получаем значение уровня стойкости разработанного алгоритма к атаке A2, равное 2^{192} для первой версии ($\mu = 3$) и 2^{256} – для второй версии ($\mu = 5$) разработанного алгоритма.

Применяя способ оценивания стойкости к атаке на основе известных подписей (атака A1), описанный в [21, 22], получаем следующее. В рамках атаки A1 уравнения, записываемые по формуле (2), и уравнения, записываемые по формуле (6), образуют две независимые и похожие системы степенных уравнений. Поэтому для оценки сложности атаки A1 достаточно рассмотреть только одну из них, например, относящуюся к формуле (2). Последняя задает для первой из N известных подписей пять фиксированных матричных неизвестных $\mathbf{D}$, $\mathbf{P}^{b_1}$, $\mathbf{G}^{n_1}$, $\mathbf{J}^{d_1}$ и $\mathbf{F}$. Действительно, неизвестные $\mathbf{D}$ и $\mathbf{F}$ присутствуют в уравнении, составляемом по формуле (2) для каждой известной подписи, а неизвестные $\mathbf{P}^{b_1}$, $\mathbf{G}^{n_1}$ и $\mathbf{J}^{d_1}$ фиксируют представителей трех неизвестных коммутативных групп. Для каждой i-й подписи ($i = 2, 3, \ldots, N$) выбор случайных значений $\mathbf{P}^{b_1}$, $\mathbf{G}^{n_1}$ и $\mathbf{J}^{d_1}$ из соответствующих скрытых групп описывается через координаты соответствующих представителей $\mathbf{P}^{b_1}$, $\mathbf{G}^{n_1}$ и $\mathbf{J}^{d_1}$ и μ уникальных скалярных неизвестных.

Таким образом, первая подпись задает $5\mu^2$ фиксированных скалярных неизвестных, а каждая из остальных – 3μ уникальных скалярных неизвестных. Для N известных подписей получаем систему из $N\mu^2$ скалярных уравнений, включающую $5\mu^2 + 3\mu(N-1)$ скалярных неизвестных. Из условия равенства числа уравнений и числа неизвестных получаем формулу для вычисления числа N_0 известных подписей, нужных для выполнения атаки A1:

$$N_0 = (\mu - 3)^{-1}(5\mu - 3). \qquad (13)$$

Для второй версии алгоритма $\mu = 5$ и $N_0 = 11$, что соответствует БСНУ, включающей 275 степенных уравнений в поле $GF(p)$ 40-битного порядка, и уровню стойкости $>2^{256}$. Для первой версии алгоритма $\mu = 3$ при любом значении N_0 условие равенства числа уравнений и неизвестных является недостижимым. По причине нелинейности уравнений, входящих в БСНУ, можно предположить, что при достаточно большом значении N_0 в процессе решения БСНУ ($N_0 > 11$) в принципе могут быть найдены значения секретных матриц $\mathbf{D}$ и $\mathbf{F}$, однако оцениваемая вычислительная сложность решения этой системы $>>2^{256}$ битовых операций. Таким образом, для обеих версий разработанного алгоритма стойкость к атаке A1 превышает стойкость к атаке A2, т.е. критерий достаточности рандомизации ЭЦП выполняется.

В связи с использованием 256-битой хеш-функции возникает вопрос об обеспечении достаточного уровня стойкости к атаке на основе парадокса о днях рождения (превышающего стойкость к атаке A2). Сценарий этой атаки включает генерацию большого числа документов, получение к им подлинных подписей и нахождением одинаковых подписей вида (e, σ, s, $\mathbf{S}$) к двум различным документам. Из алгоритма генерации ЭЦП видно, что для получения одинакового значения элемента $\mathbf{S}$ в двух различных подписях, в ходе в ходе вычисления каждой из них должны быть использованы 1) одинаковые значения e и 2) одинаковые значения случайных натуральных чисел $k < q$, $t < q$ и $r < q$ (поскольку вычисляется как в зависимости от e, так и в зависимости от k, t и r). Вероятность первого события равна 2^{-256}, а второго – $q-3 < 2^{-384}$ (для каждой из версий разработанного алгоритма). Поскольку данные события независимы (значение e зависит не только от значений k, t и r, но и от документа M), то вероятность того, что для двух различных подписей значения всех соответствующих четырех параметров совпадают,

Таблица 1.

Сравнение двух версий разработанного алгоритма ЭЦП с известными аналогами

Алгоритм	Размер открытого ключа, байт	Размер подписи, байт	Скорость генерации ЭЦП, отн.ед.	Скорость верификации ЭЦП, отн.ед.	Уровень стойкости
Версия 1	504	136	25,2	12,6	2^{192}
Версия 2	875	197	11,1	5,6	2^{256}
[26]	630	170	9,6	7,6	2^{192}
[22]	512	144	10,6	7,1	2^{100}
[14]	387	97	7,9	16,0	2^{100}
[15]	256	113	7,9	10,6	2^{80}
[13]	768	160	2,0	7,1	2^{80}

равна $<2^{-640}$. В соответствии с парадоксом о днях рождения для получения вероятности существования двух различных документов с одинаковыми значениями подписи должны быть сгенерированы 2^{320} подписей к различным документам, что определяет уровень стойкости 2^{320} к атаке на основе парадокса о днях рождения.

Таким образом, в рамках обеих версий разработанного алгоритма ЭЦП использование 256-битной хеш-функции является достаточным и не требуется увеличения разрядности функции Ф. Более того, можно применить 192-битную хеш-функцию Ф. При этом для первой (второй) версии алгоритма можно использовать алгебру матриц 3×3 (5×5), заданную над простым полем, имеющим 96-битный (32-битный) порядок p, обеспечивая более высокий уровень стойкости к атаке на основе парадокса о днях рождения по сравнению с атакой A2, значение которой при таком модифицировании не изменяется, поскольку не изменяется число степенных уравнений в БСНУ, решение которой выполняется в ходе атаки A2. Последнее замечание относится также и к атаке A1, поэтому указанное модифицирование сохраняет выполнимость критерия полноты рандомизации ЭЦП, приводя к повышению производительности процедур генерации и верификации подписи в 1,8 (1,6) раза для первой (второй) версии разработанного алгоритма.

Приводимое в табл. 1 сравнение разработанного алгоритма с известными алгебраическими алгоритмами ЭЦП, основанными на вычислительной трудности решения БСНУ, обладает более привлекательным сочетанием значений основных параметров. С учетом указанной возможности повышения производительности каждой из двух версий алгоритма можно сделать вывод о его достаточной практичности.

Выводы

Впервые реализован алгебраический алгоритм ЭЦП, основанный на вычислительной сложности решения БСНУ и использующий три скрытые коммутативные группы. В алгоритме предусмотрены две версии реализации с различным уровнем стойкости (2^{192} и 2^{256}). Обе версии обладают относительно малыми размерами подписи и открытого ключа (в сравнении с известными постквантовыми алгоритмами ЭЦП для заданного уровня стойкости) и достаточно высокой производительностью. Выполненная алгоритмическая разработка и использованные приемы расширяют предпосылки для решения о целесообразности приложения усилий по разработке постквантового стандарта ЭЦП на конечных некоммутативных алгебрах.

Исследование выполнено за счет гранта Российского научного фонда № 24-41-04006, https://rscf.ru/project/24-41-04006/

Литература

1. Post-Quantum Cryptography. 15th International Conference, PQCrypto 2024, Oxford, UK, June 12–14, 2024, Proceedings // Lecture Notes in Computer Science. 2024. V. 14771–14772. Springer, Cham.
2. Post-Quantum Cryptography. 14th International Conference, PQCrypto 2023, College Park, MD, USA, August 16–18, 2023, Proceedings // Lecture Notes in Computer Science. 2023. V. 14154. Springer, Cham.
3. Ding J., Petzoldt A., Schmidt D. S. Multivariate Cryptography // In: Multivariate Public Key Cryptosystems. Advances in Information Security. 2020. V. 80. Springer, New York, NY. https://doi.org/10.1007/978-1-0716-0987-3_2.

4. J. Ding, A. Petzoldt. Current State of Multivariate Cryptography // IEEE Security and Privacy Magazine. 2017. V. 15. N. 4. P. 28–36.
5. Hashimoto Y. Recent Developments in Multivariate Public Key Cryptosystems // In: Takagi, T., Wakayama, M., Tanaka, K., Kunihiro, N., Kimoto, K., Ikematsu, Y. (eds). International Symposium on Mathematics, Quantum Theory, and Cryptography. Mathematics for Industry, 2021. V. 33. P. 209–229. Springer, Singapore. https://doi.org/10.1007/978-981-15-5191-8_16.
6. Ikematsu Y., Nakamura S., Takagi T. Recent progress in the security evaluation of multivariate public-key cryptography // IET Information Security. 2022. P. 1–17. DOI: 10.1049/ise2.12092.
7. Øygarden M., Felke P., Raddum H. Analysis of multivariate encryption schemes: Application to Dob and C* // Journal of Cryptology. 2024. V. 37. N. 3. Article 20. DOI: 10,1007/s00145-024-09501-w.
8. Omar S., Padhye S., Dey D. Cryptanalysis of multivariate threshold ring signature schemes // Information Processing Letters.2023. V. 181. Article 106357. DOI: 10.1016/j.ipl.2022.106357.
9. Moldovyan N. A. Finite algebras in the design of multivariate cryptography algorithms // Bulletin of Academy of Sciences of Moldova. Mathematics. 2023. No. 3 (103). P. 80–89. DOI: https://doi.org/10.56415/basm.y2023.i3.p80.
10. Moldovyan N. A. Parameterized method for specifying vector finite fields of arbitrary dimensions // Quasigroups and related systems. 2024. Vol. 32. N. 2. P.299–312. DOI: 10.56415/qrs.v32.21.
11. Moldovyan A. A., Moldovyan N. A. Vector finite fields of characteristic two as algebraic support of multivariate cryptography // Computer Science Journal of Moldova. 2024. V. 32. N. 1(94). P. 46–60. DOI: 10.56415/csjm.v32.04.
12. Moldovyan N. A. Algebraic signature algorithms with a hidden group, based on hardness of solving systems of quadratic equations // Quasigroups and Related Systems. 2022. Vol. 30. N. 2(48). P. 287–298. DOI: 10.56415/qrs.v30.24.
13. Moldovyan D. N. A new type of digital signature algorithms with a hidden group // Computer Science Journal of Moldova. 2023. V. 31. N. 1(91). P. 111–124. doi:10.56415/csjm.v31.06.
14. Duong M. T., Moldovyan D. N., Do B. V., Nguyen M. H. Post-quantum signature algorithms on noncommutative algebras, using difficulty of solving systems of quadratic equations // Computer Standards & Interfaces. 2023. V. 86. Article 103740. DOI: 10.1016/j.csi.2023.103740.
15. Молдовян Д. Н., Молдовян А. А. Алгебраические алгоритмы ЭЦП, основанные на трудности решения систем уравнений // Вопросы кибербезопасности. 2022. № 2(48). С. 7–17. DOI: 10.21681/2311-3456-2022-2-7-17.
16. Moldovyan D. N. New Form of the Hidden Logarithm Problem and Its Algebraic Support // Bulletin of Academy of Sciences of Moldova. Mathematics. 2020. No. 2(93). P. 3–10.
17. Moldovyan D. N. A practical digital signature scheme based on the hidden logarithm problem // Computer Science Journal of Moldova. 2021. Vol. 29. N. 2(86). P. 206–226.
18. Moldovyan N. A. Signature Schemes on Algebras, Satisfying Enhanced Criterion of Post-quantum Security // Buletinul Academiei de Stiinte a Republicii Moldova. Matematica. 2020. N. 2(93). P. 62–67.
19. Moldovyan A. A. Complete signature randomization in an algebraic cryptoscheme with a hidden group // Quasigroups and related systems. 2024. Vol. 32. N. 1. P. 95–108. DOI: DOI: 10.56415/qrs.v32.08.
20. Молдовян А. А., Молдовян Д. Н., Костина А. А. Алгебраические алгоритмы ЭЦП с полной рандомизацией подписи // Вопросы кибербезопасности. 2024. № 2(60). С. 95–102. DOI: 10.21681/2311-3456-2024-2-95-102.
21. Молдовян Д. Н., Костина А. А. Способ усиления рандомизации подписи в алгоритмах ЭЦП на некоммутативных алгебрах // Вопросы кибербезопасности. 2024. № 4(62). С. 71–81. DOI: 10.21681/2311-3456-2024-4-71-81.
22. Молдовян Н. А, Петренко А. С. Алгебраический алгоритм ЭЦП с двумя скрытыми группами // Вопросы кибербезопасности. 2024. № 6(64). С. 98–107. DOI: 10.21681/2311-3456-2024-6-98-107.
23. Duong M. T., Do B. T., Nguyen M. H., Kurysheva A. A., Kostina A. A., Moldovyan D. N. Signature Algorithms on Non-commutative Algebras Over Finite Fields of Characteristic Two // Future Data and Security Engineering. Big Data, Security and Privacy, Smart City and Industry 4.0 Applications. Springer Nature Singapore, 2022. P. 273–284, DOI: 10.1007/978-981-19-8069-5-18.
24. Moldovyan D. N. A unified method for setting finite non-commutative associative algebras and their properties // Quasigroups and Related Systems. 2019. V. 27. N. 2, pp. 293–308.
25. Moldovyan N. A. Unified method for defining finite associative algebras of arbitrary even dimensions, Quasigroups and Related Systems. 2018. V. 26. N. 2. P. 263–270.
26. Захаров Д. В., Костина А. А., Морозова Е. В., Молдовян Д. Н. Алгоритм ЭЦП на алгебре матриц 3x3, использующий две скрытые группы // Вопросы кибербезопасности. 2025. № 3(67). С. 45–54. DOI: 10.21681/2311-3456-2025-3-45-54.

POST-QUANTUM ALGEBRAIC SIGNATURE ALGORITHM WITH THREE HIDDEN GROUPS

Moldovyan A. A.[2]

Keywords: *finite matrix algebra; associative algebra; computationally hard problem; post-quantum cryptography; hidden commutative group; digital signature; signature randomization.*

Purpose of work *is the creation of additional prerequisites for the development of a post-quantum standard for digital signature algorithms based on the computational complexity of solving large systems of nonlinear equations (LSNE) in a finite field.*

2 Alexander A. Moldovyan, Dr. Sc. (in Tech.), chief researcher of laboratory of computer security problems, St. Petersburg Federal Research Center of the Russian Academy of Sciences, St. Petersburg, Russia. ORCID: https://orcid.org/0000-0001-5480-6016. E-mail: maa1305@yandex.ru

Research methods: *application of three hidden commutative groups, the elements of each of which are non-commutative with the elements of another one, for the implementation of enhanced signature randomization in algebraic digital signature schemes, the security of which is based on the computational difficulty of solving the LSNE in the ground finite field* $GF(p)$*. Calculation of the fitting signature element in the form of a matrix* $\mathbf{S}$ *depending on three pairwise non-commutative matrices selected randomly from the hidden groups. Application of the finite algebra of 3×3 and 5×5 matrices defined over the field* $GF(p)$ *with 64-bit and 40-bit prime order* p*.*

Results of the study: *a new mechanism for enhanced randomization of the signature fitting element is proposed and an algebraic algorithm for digital signature is developed, which is promising as a prototype of a practical post-quantum digital signature standard. The selection of random matrices from hidden groups is specified by exponentiating the generators of the corresponding hidden groups to random powers, calculated depending on the randomization parameters and the randomizing element of the digital signature. For the first time, to increase the security level to potential attacks based on alternative secret keys, the specified degrees are calculated as a solution to a system of linear equations. Estimates of the security to a direct attack and an attack based on known signatures are presented. A comparison of the parameters of the developed digital signature algorithm with known algorithms using the difficulty of solving the LSNE is given, and ways to improve its performance are discussed.*

Practical relevance: *the obtained results consist in creating a new premise for substantiating the choice of algebraic digital signature algorithms with hidden groups as a basis for developing a practical post-quantum digital signature standard, which consists in increasing the level of resistance to potential attacks based on equivalent keys.*

References

1. Post-Quantum Cryptography. 15th International Conference, PQCrypto 2024, Oxford, UK, June 12–14, 2024, Proceedings // Lecture Notes in Computer Science. 2024, vol. 14771–14772. Springer, Cham.
2. Post-Quantum Cryptography. 14th International Conference, PQCrypto 2023, College Park, MD, USA, August 16–18, 2023, Proceedings // Lecture Notes in Computer Science. 2023, vol. 14154. Springer, Cham.
3. Ding J., Petzoldt A., Schmidt D. S. Multivariate Cryptography // In: Multivariate Public Key Cryptosystems. Advances in Information Security. 2020. V. 80. Springer, New York, NY. https://doi.org/10.1007/978-1-0716-0987-3_2.
4. J. Ding, A. Petzoldt Current State of Multivariate Cryptography // IEEE Security and Privacy Magazine. 2017, vol. 15, no. 4, pp. 28–36.
5. Hashimoto Y. Recent Developments in Multivariate Public Key Cryptosystems // In: Takagi, T., Wakayama, M., Tanaka, K., Kunihiro, N., Kimoto, K., Ikematsu, Y. (eds). International Symposium on Mathematics, Quantum Theory, and Cryptography. Mathematics for Industry. 2021, vol. 33, pp. 209–229. Springer, Singapore. https://doi.org/10.1007/978-981-15-5191-8_16.
6. Ikematsu Y., Nakamura S., Takagi T. Recent progress in the security evaluation of multivariate public-key cryptography // IET Information Security. 2022, pp. 1–17. DOI: 10.1049/ise2.12092.
7. Øygarden M., Felke P., Raddum H. Analysis of multivariate encryption schemes: Application to Dob and C* // Journal of Cryptology. 2024, vol. 37, no. 3, article 20. DOI: 10,1007/s00145-024-09501-w.
8. Omar S., Padhye S., Dey D. Cryptanalysis of multivariate threshold ring signature schemes // Information Processing Letters.2023, vol. 181, article 106357. DOI: 10.1016/j.ipl.2022.106357.
9. Moldovyan N. A. Finite algebras in the design of multivariate cryptography algorithms // Bulletin of Academy of Sciences of Moldova. Mathematics. 2023, no. 3(103), pp. 80–89. DOI: https://doi.org/10.56415/basm.y2023.i3.p80.
10. Moldovyan N. A. Parameterized method for specifying vector finite fields of arbitrary dimensions // Quasigroups and related systems. 2024, vol. 32, no. 2, pp. 299–312. DOI: 10.56415/qrs.v32.21.
11. Moldovyan A. A., Moldovyan N. A. Vector finite fields of characteristic two as algebraic support of multivariate cryptography // Computer Science Journal of Moldova. 2024, vol. 32, no. 1(94), pp. 46–60. DOI: 10.56415/csjm.v32.04.
12. Moldovyan N. A. Algebraic signature algorithms with a hidden group, based on hardness of solving systems of quadratic equations // Quasigroups and Related Systems. 2022, vol. 30 no. 2(48), pp. 287–298. DOI: 10.56415/qrs.v30.24.
13. Moldovyan D. N. A new type of digital signature algorithms with a hidden group // Computer Science Journal of Moldova. 2023, vol. 31, no. 1(91), pp. 111–124. doi:10.56415/csjm.v31.06.
14. Duong M. T., Moldovyan D. N., Do B. V., Nguyen M. H. Post-quantum signature algorithms on noncommutative algebras, using difficulty of solving systems of quadratic equations // Computer Standards & Interfaces. 2023, vol. 86, article 103740. DOI: 10.1016/j.csi.2023.103740.
15. Moldovyan D. N., Moldovyan A. A. Algebraic signature algorithms based on difficulty of solving systems of equations. Voprosy kiberbezopasnosti [Cybersecurity issues]. 2022, no. 2(48), pp. 7–17. DOI: 10.21681/2311-3456-2022-2-7-17.
16. Moldovyan D. N. New Form of the Hidden Logarithm Problem and Its Algebraic Support // Bulletin of Academy of Sciences of Moldova. Mathematics. 2020, no. 2(93), pp. 3–10.
17. Moldovyan D. N. A practical digital signature scheme based on the hidden logarithm problem // Computer Science Journal of Moldova. 2021, vol. 29, no. 2(86), pp. 206–226.
18. Moldovyan N. A. Signature Schemes on Algebras, Satisfying Enhanced Criterion of Post-quantum Security // Buletinul Academiei de Stiinte a Republicii Moldova. Matematica. 2020, no. 2(93), pp. 62–67.
19. Moldovyan A. A. Complete signature randomization in an algebraic cryptoscheme with a hidden group // Quasigroups and related systems. 2024, vol. 32, no. 1, pp. 95–108. DOI: DOI: 10.56415/qrs.v32.08.
20. Moldovyan A. A., Moldovyan D. N., Kostina A. A. Algebraic signature algorithms with complete signature randomization. Voprosy kiberbezopasnosti [Cybersecurity issues]. 2024, no. 2(60), pp. 95–102. DOI: 10.21681/2311-3456-2024-2-95-102.
21. Moldovyan D. N., Kostina A. A. A method for strengthening signature randomization in algebraic signature algorithms on noncommutative algebras. Voprosy kiberbezopasnosti [Cybersecurity issues]. 2024, no. 4(62), pp. 71–81. DOI: 10.21681/2311-3456-2024-4-71-81.

22. Moldovyan N. A., Petrenko A. S. Algebraic signature algorithm with two hidden groups. Voprosy kiberbezopasnosti [Cybersecurity issues]. 2024, no. 6(64), pp. 98–107. DOI: 10.21681/2311-3456-2024-6-98-107.
23. Duong M. T.,, Do B. T., Nguyen M. H., Kurysheva A. A., Kostina A. A., Moldovyan D. N. Signature Algorithms on Non-commutative Algebras Over Finite Fields of Characteristic Two // Future Data and Security Engineering. Big Data, Security and Privacy, Smart City and Industry 4.0 Applications. Springer Nature Singapore, 2022, pp. 273–284, DOI: 10.1007/978-981-19-8069-5-18.
24. Moldovyan D. N. A unified method for setting finite non-commutative associative algebras and their properties // Quasigroups and Related Systems. 2019, vol. 27, no. 2, pp. 293–308.
25. Moldovyan N. A. Unified method for defining finite associative algebras of arbitrary even dimensions, Quasigroups and Related Systems. 2018, vol. 26, no. 2, pp. 263–270.
26. Zakharov D. V., Moldovyan D. N., Kostina A. A., Morozova E. V., Moldovyan D. N. A digital signature algorithm on the algebra of 3×3 matrices, which uses two hidden groups. Voprosy kiberbezopasnosti [Cybersecurity issues]. 2025, no. 3(67), pp. 45–54. DOI: 10.21681/2311-3456-2025-3-45-54.

ПАРАМЕТРИЗАЦИЯ ПОСТКВАНТОВОЙ ЭЛЕКТРОННОЙ ПОДПИСИ КНАА-2-ЭЦП

Петренко А. С.[1]

DOI: 10.21681/2311-3456-2025-5-88-95

Цель работы: *формализовать выбор параметров и определение требований к КНАА-2-ЭЦП, обеспечивающих постквантовую устойчивость (не ниже 2^{200} для Cat 3 и 2^{256} для Cat 5) при сокращении длины подписи до 112 байт (Cat 3) и 128 байт (Cat 5).*

Метод исследования: *обоснование соответствия КНАА-2-ЭЦП нормативам FIPS 204/205/206, SP 800-57-1 Rev.5, ISO/IEC 14888-4:2024 и ГОСТ Р 34.10-2018 с точным определением форматов ASN.1/DER и OID, разработка двух наборов параметров для категорий стойкости 3 и 5 по шкале NIST, задание количественных целевых показателей, формулирование обязательных требований к реализации, в частности константного времени, корректной рандомизации и контрмер.*

Результаты исследования: *установлены два набора параметров, обеспечивающие стойкость $\geq 2^{200}$ (Cat 3) и $\geq 2^{256}$ (Cat 5) и компактность ключей и подписей, определены требования к размеру ключей и подписи, заданы метрики производительности и безопасности, установлены критерии проверки корректности реализации без полного цикла ACVP, регламентированы форматы и процедуры валидации, определены контрмеры против побочных атак и требования к постоянному времени выполнения.*

Научная и практическая значимость *результатов статьи заключается в том, что найденные параметры, требования и профили использования впервые демонстрирует реализацию усиленной рандомизации подписи без удвоения проверочного уравнения, обеспечивая компактность ключей и подписей при сохранении высокой стойкости к квантовым атакам и к атакам по побочным каналам. Это позволяет значительно снизить объем транзакционных данных и требования к пропускной способности сетей и хранилищ, что критически важно для масштабируемых блокчейн-систем, аппаратных криптопроцессоров и энергоэффективных IoT-устройств.*

Ключевые слова: *конечная некоммутативная ассоциативная алгебра; форматы ASN.1/DER; категории стойкости NIST; константное время выполнения; рандомизация подписи; контрмеры безопасности; постквантовая электронная подпись.*

Введение. Рекомендации нормативных документов

Структурно выбор параметров и определение требований основываются на шести документах, напрямую определяющих современные требования к постквантовым электронным подписям и их валидации. Первые три – это стандарты Национального института стандартов и технологий США, целиком посвящённые подписи. Так, FIPS 204 описывает решётчатую схему ML DSA (прежнее название Dilithium) и задаёт образцовую структуру постквантового стандарта от административных пунктов до приложений с контрольными примерами [1]. FIPS 205 регламентирует хэш подпись без сохранения состояния SLH DSA (производную от SPHINCS+) [2]. FIPS 206, находящийся на завершающем этапе утверждения, посвящён подписи FN DSA (Falcon) и примечателен минимальным из финалистов NIST объёмом подписи в 666 байт [3].

Также немаловажно рассмотреть NIST SP 800 57 1 (в 5 редакции), в котором введена шкала категорий криптографической стойкости и приведено сопоставление этих категорий с длинами симметричных ключей [4]. На основании этого документа в текущей статье определены два уровня безопасности – категория 3 ($\approx 2^{200}$ операций перебора) и категория 5 ($\approx 2^{256}$).

Международный стандарт ISO/IEC 14888 4 (2024 г.) посвящён хэш подписям с сохранением состояния [5], его ценность для данной работы состоит в едином формате представления открытых ключей и подписей на языке ASN.1. Формат основывается на структурах SEQUENCE, где модуль p и базисные векторы кодируются как INTEGER и OCTET STRING соответственно, аналогично разделу A.1 FIPS 204. КНАА-2 ЭЦП оформлена в том же формате ASN.1 с кодированием DER (Distinguished Encoding Rules, канонические двоичные правила кодирования ASN.1), что обеспечивает легкое внедрение в существующие инфраструктуры открытых ключей [6].

Дополнительно целесообразно ориентироваться на ГОСТ Р 34.10-2018 «Информационные технологии. Криптографическая защита информации. Процессы формирования и проверки электронной подписи». Хотя данный стандарт и регламентирует классическую подпись на эллиптических кривых,

1 Петренко Алексей Сергеевич, младший научный сотрудник лаборатории проблем компьютерной безопасности Санкт-Петербургского федерального исследовательского центра Российской академии наук, Санкт-Петербург, Россия. ORCID: https://orcid.org/0000-0002-9954-4643. Scopus Author ID: 57200260915. E mail: a.petrenko1999@rambler.ru

он подробно описывает форматы представления ключей и подписей, процедуры одиночного контрольного примера (KAT), отрицательного теста и парной согласованности ключа. Наконец, ГОСТ 19.105 78 и ГОСТ 19.301 79 определяют требования к оформлению программной документации и методик испытаний.

Перечисленные документы в совокупности образуют достаточную нормативную базу для валидации постквантовой электронной подписи КНАА-2-ЭЦП [7]. Они определяют диапазоны параметров, форму представления результатов и минимальный набор контрольных испытаний, благодаря чему характеристики КНАА-2 ЭЦП можно прямо сопоставлять с утверждёнными и проектируемыми постквантовыми стандартами ЭЦП [8-11].

В статье используются следующие обозначения:

- σ – электронная подпись (результат алгоритма Sign);
- pk / sk – открытый и секретный ключ соответственно;
- v – вектор фиксатор (компонента подписи);
- r, s – скалярные компоненты подписи;
- p – модуль конечного поля $GF(p)$;
- mul – одна операция умножения в $GF(p)$ с последующей редукцией по модулю p.

Применяемые наборы параметров

Используются две целевые категории, определённые руководством NIST SP 800 57 1: категория 3 и категория 5. Категория 3 сопоставима с полной подборкой, требующей порядка 2^{200} элементарных операций, что эквивалентно симметричному ключу длиной не менее 192 бит. Категория 5 соответствует трудоёмкости 2^{256} и соответствует уровню безопасности симметричного ключа длиной 256 бит, тем самым обеспечивая максимально возможный запас прочности в открытой публикации NIST. Ниже в таблице 1 приведено краткое соответствие между категориями, минимальными длинами симметричных ключей и наиболее близкими параметрами уже утверждённых NIST схем подписей (σ – обозначение результирующей электронной подписи).

Таблица 1.

Сопоставление категорий стойкости NIST, эквивалентных симметричных ключей и размеров подписей ML DSA и SLH DSA

Категория NIST	Симметричный ключ (экв.)	Подпись FIPS 204 (ML DSA)	Подпись FIPS 205 (SLH DSA)
3	≥ 192 бит	ML DSA 65 (кат. 3): $\sigma \approx 3{,}23$ КБ	SPHINCS+ 192s: $\sigma \approx 15{,}9$ КБ
5	≥ 256 бит	ML DSA 87 (кат. 5): $\sigma \approx 4{,}52$ КБ	SPHINCS+ 256s: $\sigma \approx 29{,}1$ КБ

Переход от теоретических категорий к конкретным параметрам схемы КНАА-2 ЭЦП основан на следующем результате. Задача восстановления секретного вектора сводится к решению уравнения высокой степени в конечной ассоциативной алгебре, а её сложность оценивается снизу сведением к задаче поиска короткого вектора (SVP) размерности 4. Для параметров, приведённых далее, эта сложность не меньше 2^{200} и 2^{256} операций соответственно, что надёжно покрывает требования двух выбранных категорий.

Практически же категории разделяют области применения, так уровень 3 достаточен для массовых транзакций и пользовательских электронных удостоверений с жизненным циклом до 15 лет, когда как уровень 5 предназначен для долгосрочного нотариального хранения, критических государственных регистров и инфраструктур, рассчитанных на горизонты более тридцати лет. В обеих ситуациях сокращение подписи до 112–128 байт позволяет не менять форматы блоков и сохранять пропускную способность систем даже при полном переходе на постквантовую криптографию.

В данной статье в качестве объектов сравнения используются стандартизованные или находящиеся в финальной стадии стандартизации алгоритмы постквантовой ЭЦП (Dilithium, Falcon, SPHINCS+), а также разработанная схема на основе КНАА. При этом схемы XMSS и LMS, равно как и отечественные ЭЦП (Шиповник, Крыжовник и Гиперикум), в сравнительный анализ не включены ввиду отсутствия их формализованной стандартизации, ограниченности распространённых реализаций и других прикладных барьеров.

Для каждой из двух целевых категорий стойкости установлены конкретные величины открытого ключа, подписи, секретного ключа и трудоёмкости операций подписи и проверки. Таблица 2 сводит полученные значения. При этом mul – это одна операция умножения в поле $GF(p)$ с редукцией по модулю p. Число mul (напр. 9 200 при Verify) показывает трудоёмкость независимо от платформы.

Данные в таблице являются нормативной «верхней границей», реализация же, предъявляемая к испытаниям, должна удовлетворять указанным размерам и не выходить за пределы трудоёмкости операций.

Определяющим конкурентным параметром КНАА-2 ЭЦП является объём результирующей подписи σ. Для выбранных наборов параметров (табл. 1) длина подписи равна 112 байт в категории 3 и 128 байт в категории 5. Ниже в таблице 3 приведено сравнение с действующими и проектируемыми постквантовыми стандартами (данные взяты из текстов FIPS и технических отчётов NIST).

DOI: 10.21681/2311-3456-2025-5-88-95

Таблица 2.

Основные параметры KHAA 2 ЭЦП и трудоёмкость операций для категорий 3 и 5

Показатель	Категория 3	Категория 5
Модуль поля p (двоичных разрядов)	64 бит (реком. $p = 2^{64} - 59$)	128 бит (реком. $p = 2^{128} - 159$)
Размер открытого ключа pk	512 байт (прототип)	704 байт (прототип)
Размер подписи σ	112 байт (прототип)	128 байт (прототип)
Размер секретного ключа sk	640 байт	768 байт
Операций умножения $GF(p)$ при Sign (процедуре подписи)	6 600 mul	6 530 mul
Операций умножения $GF(p)$ при Verify (процедуре проверки)	9 200 mul	9 200 mul

Таблица 3.

Длины подписей постквантовых схем

Алгоритм (категория стойкости)	Длина подписи, байт
KNAA-2 ЭЦП Cat 3	112
KNAA-2 ЭЦП Cat 5	128
FN DSA (Falcon 512, Cat 1)	666
FN DSA (Falcon 1024, Cat 5)	1 280
ML DSA (Dilithium 44, Cat 3)	2 420
ML DSA (Dilithium 87, Cat 5)	4 595
SLH DSA (SPHINCS+ 128s)	7 856
SLH DSA (SPHINCS+ 256s)	41 472

Сокращение подписи до сотен байт имеет два практических следствия. Во-первых, объём блока в распределённом реестре не увеличивается при переходе на постквантовую криптографию, что позволяет сохранять прежний уровень пропускной способности сети и прежнюю стоимость транзакций. Во-вторых, сужается полоса пропускания каналов связи и объём долговременного хранилища: при средней нагрузке 10 000 транзакций в минуту годовой выигрыш объёма относительно схемы Falcon 1024 составляет порядка 10 ТБ.

Показатель длины подписи выбран ключевым критерием эффективности, а следовательно любые оптимизации реализации не должны приводить к увеличению значения σ выше норм, зафиксированных в таблице 2.

Выбор числовых параметров KHAA-2 ЭЦП осуществлялся исходя из достаточной криптографической стойкости и технологической реализуемости на массовых аппаратных платформах.

С точки зрения стойкости фиксированы два предельных модуля поля: $p_{64} = 2^{64} - 59$ для категории 3 и $p_{128} = 2^{128} - 159$ для категории 5. Оба значения являются простыми и удовлетворяют условию $r = p^2 + p + 1$ – простое, что обеспечивает отсутствие малых множителей в порядке мультипликативной подгруппы и равномерное распределение коэффициентов векторов. При таких модулях плотность перебора возможных значений вектора фиксатора превышает 2^{200} и 2^{256} соответственно, что надёжно перекрывает минимальные требования категорий стойкости.

Вторым ограничением служит норма секретных векторов. Для категории 3 установлено значение $\|s\| \le 2^{31}$, при этом среднее число ненулевых коэффициентов в процедуре поиска короткого вектора не превышает 72, а полный перебор оказывается эквивалентным 2^{200} шагам. Для категории 5 норма сокращена до $\|s\| \le 2^{17}$, при расширении поля до 128 бит, что увеличивает сопротивляемость атаке перебора коэффициентов до уровня 2^{256}.

Операционное ограничение связано с возможностями современных 64 и 128 битовых процессоров, так умножение координат модуля p_{64} выполняется одной инструкцией, а умножение по p_{128} – двумя последовательными операциям $mul\ lo$ / $mul\ hi$ без привлечения арифметики с плавающей запятой. $mul\ lo$ / mul hi это пара 64 битных инструкций, которые по очереди выдают младшие и старшие 64 бита из 128 битного произведения двух слов. На x86 64 это MUL (низ) и MULH (верх), а на ARM – UMULH для старшей части. В схеме категории 5 одна операция умножения $GF(p)$ реализуется именно этой парой – сначала $mul\ lo$, затем $mul\ hi$, после чего выполняется редукция по модулю p. Тем самым достигается заявленное число умножений (6 600 / 9 200), а время проверки подписи ориентировочно 0,5 μs (микросекунд) на ядро 3 ГГц (при компилируемой реализации).

Наконец, размеры открытого и секретного ключей (512/704 байт и 640/768 байт в зависимости

от категории) выбраны так, чтобы полностью помещаться в память типового защищённого микроконтроллера и в регистр хранилища аппаратного модуля HSM без фрагментации. В совокупности предложенные параметры гарантируют необходимый уровень стойкости и соответствуют практическим ограничениям вычислительных и коммуникационных ресурсов.

Требования безопасности

Корректная реализация КНАА-2 ЭЦП должна обеспечивать постоянное (равномерное) время выполнения всех операций формирования и проверки подписи. Продолжительность вычислений не может зависеть от содержимого секретных данных, иначе возникает возможность побочных атак по времени или по электромагнитному излучению. Настоящее требование согласуется с разделом 9 стандартов FIPS 204 и FIPS 205, где прямо предписывается исключать любые условные переходы (ветвление) кода, условно связанные с закрытой информацией.

В практическом исполнении это означает следующее. Все арифметические действия над элементами поля $GF(p)$ – умножения, возведения в степень, сложения и вычитания – выполняются по фиксированному числу машинных инструкций, а использование таблиц переходов, где адрес ячейки определяется секретным индексом, не допускается. Переходы между функциями и выбор алгоритмических ветвей разрешается основывать лишь на открытых параметрах, известных ещё до начала криптографического процесса, таких как размер модуля p, категория стойкости или вариант компиляции.

Равномерность времени при этом оценивается экспериментально, на серии из десяти тысяч однотипных тестовых примеров измеряется медиана длительности цикла «подпись – проверка». Разброс между минимальным и максимальным значениями не должен превышать одного процента от медианы. Измерения могут проводиться как средствами программного профилирования (например, perf stat с подсчётом тактов), так и аппаратным методом – прямым наблюдением сигнала питания с временным разрешением не хуже одной наносекунды. При превышении указанного порога реализация считается не отвечающей требованиям и подлежит доработке.

Следует отметить, что алгоритм КНАА-2 ЭЦП требует случайных данных только при генерации секретных ключей, они же создаются на базе логистического хаотического отображения:

$$x_{\ell+1} = r\, x_\ell\, (1 - x_\ell), \quad 3{,}999 \le r \le 4, \quad 0 < x_0 < 1.$$

При указанном значении параметра r траектория отображения демонстрирует режим полностью развитого хаоса ($r \approx 4$): малейшие изменения начального условия x_0 приводят к непредсказуемому расхождению последовательностей. Числовые выходы x_ℓ после преобразования в двоичную форму (берутся младшие 32 бита каждого результата умножения) служат источником случайности для детерминированного генератора псевдослучайных чисел (DRBG), реализованного по рекомендациям ГОСТ Р 50.1.111-2016 и NIST SP 800-90A Rev.1 (DRBG). DRBG реализован как Hash_DRBG на SHA 256 (NIST SP 800 90A Rev.1) vtnjlb100] и, альтернативно, как ГОСТ-совместимый генератор на хэше «Стрибог-256» (ГОСТ Р 34.11-2012). В реализации выбирается один из механизмов, но тесты качества (NIST STS, автотест DRBG) выполняются для выбранного варианта.

Процедура инициализации включает три шага.

1. Из встроенного стохастического генератора TRNG (True Random Number Generator) [12] считывается 256 битовое начальное значение (seed), которое интерпретируется как вещественное x_0 в интервале (0,1).
2. Логистическое отображение выполняется в 1024 итерациях (1000 начальных + 24 для выборки), чтобы исключить корреляцию с x_0. Результаты x_{1000}, … , x_{1023} переводятся в 32 битовые слова и подаются на вход DRBG.
3. DRBG выдаёт требуемый объём случайных байтов для наполнения секретных векторов схемы.

Качество полученного потока проверяют по трём критериям:

- прохождение стандартного статистического набора NIST STS с p-value (уровень значимости) каждой процедуры не ниже 0,01;
- отсутствие значимых автокорреляций на интервале до 128 символов (показатель взаимной корреляции не выше 0,05);
- положительный результат встроенного автотеста DRBG при каждом запуске.

Совместное использование трёх источников неопределённости в виде аппаратного TRNG, хаотического логистического отображения и крипто-стойкого DRBG способно обеспечить непрерывность случайного потока и исключает предсказуемость секретных параметров генерации ключей.

OID-идентификация и форматы ASN.1

Идентификаторы объектов.

Организации разработчику выделяется номер <OrgID> в ветви 1.3.6.1.4.1. Для схемы КНАА-2 ЭЦП резервируются два объекта (формальный синтаксис ASN.1) (см. вставку 1).

Все целые поля записываются без знака, в формате big endian. Эти структуры можно прямо вкладывать в сертификат X.509 (стандартный контейнер открытого ключа (RFC 5280)) [13], контейнер CMS (Cryptographic Message Syntax, формат электронных подписей и зашифрованных сообщений (RFC 5652))

```
knaadsaCat3OID  OBJECT IDENTIFIER ::=   { iso(1) org(3) dod(6) internet(1)
                                        private(4) enterprise(1)
                                        <OrgID> 1 }
knaadsaCat5OID  OBJECT IDENTIFIER ::=   { iso(1) org(3) dod(6) internet(1)
                                        private(4) enterprise(1)
                                        <OrgID> 2 }
Первый обозначает параметры категории стойкости 3, второй параметры категории 5.
Форматы данных (формальный синтаксис ASN.1).
KNAADSA-Types DEFINITIONS ::= BEGIN
KnaadsaPublicKey ::= SEQUENCE {
       modulus       INTEGER,             -- простое p (64 или 128 бит)
       basis         OCTET STRING         -- 4×4 коэффициентов (32 байта)
}
KnaadsaPrivateKey ::= SEQUENCE {
       publicPart  KnaadsaPublicKey,
       seed          OCTET STRING,        -- начальное x0 логистического отображения
       secretVecs    OCTET STRING         -- s1 || s2 (по 4 элемента)
}
KnaadsaSignature ::= SEQUENCE {
       vVector       OCTET STRING,        -- вектор-фиксатор (32 байта)
       rScalar       INTEGER,
       sScalar       INTEGER
}
END
```

Вставка 1

[14] или объект PKCS #11 (запись ключа или подписи внутри аппаратного токена HSM), для этого достаточно указать OID knaadsaCat3OID либо knaadsaCat5OID.

Профили применения алгоритма

Для различных эксплуатационных сценариев задаются три профиля, различающиеся только внешними параметрами вызовов и способом упаковки данных, внутренняя математика алгоритма при этом остаётся неизменной.

Профиль P-SINGLE (одиночная подпись транзакции).

Используется в клиентских приложениях и лёгких кошельках. На вход функции подписи передаются сообщение M произвольной длины и секретный ключ категории 3 или 5. Полученная подпись добавляется непосредственно в поле транзакции (приведено в виде псевдокода):

```
function sign_single(sk, M):
        (v, r, s) ← Sign(M, sk)
        return (v, r, s)
```

Вставка 2

Верификация выполняется узлами сети, открытый ключ хранится в структуре аккаунта.

Профиль P-AGG (агрегированная подпись узлов валидаторов).

Предназначен для протоколов консенсуса, где множество валидаторов подтверждает один блок. Подписи формируются независимо, затем компонента v агрегируется путём поэлементного умножения, после чего совместно вычисляется скаляр r_{sum}. Это позволяет держать размер агрегированной подписи фиксированным в виде трех элементов, независимо от количества участников (приведено в виде псевдокода):

```
function aggregate_signatures (list_of_
signatures):
    v_prod ← 1
    r_sum  ← 0
    s_sum  ← 0
    for (v, r, s) in list_of_signatures:
        v_prod ← v_prod * v
        r_sum  ← r_sum  + r
        s_sum  ← s_sum  + s
    return (v_prod, r_sum mod p, s_sum mod p)
```

Вставка 3

Корректность проверяется единожды для объединённой тройки.

Профиль P-PRECOMP (предкомпилированная проверка в виртуальной машине). Используется в исполнителях смарт контрактов. Функция проверки переносится во встроенный модуль виртуальной машины (EVM core) как предкомпилированный контракт (precompile) для ускорения. Формат вызова (приведен в виде псевдокода):

```
precompile id = 0xD3...
input  = pk ‖ v ‖ r ‖ s ‖ hash(M)
output = 0x01 (valid) | 0x00 (invalid)
```

Вставка 4

Газ (gas) – расчётная единица, которой EVM-совместимые сети измеряют стоимость вычислений, так каждая операция интерпретатора имеет фиксированную цену в газ, а плата пользователя равна произведению объёма газ на текущую цену токена. Для предкомпилированной проверки КНАА-2-ЭЦП (идентификатор 0xD3) принимается тариф: 8 000 газ базово + 3 газ за каждые 100 операций mul в $GF(p)$. Для 9 200 mul это 35 600 газ. Тариф выбран по аналогии с EIP 198 (modexp) и существующими предкомпиляциями EVM. Такое соотношение стандартной и переменной частей делает новую проверку сопоставимой по стоимости с уже существующими криптографическими предкомпилированными контрактами и упрощает экономическую модель смарт-контрактов.

Особенности реализации КНАА-2 ЭЦП

Также для корректной реализации важны следующие контрмеры, представленные на рисунке 1:

1. Постоянное время (Timing SAF).
   - Единственный путь исполнения, отсутствие if/else, зависящих от секретных значений.
   - Таблицы допускаются только с фиксированным, независящим от секрета индексом, перебор осуществляется по всем элементам, если требуется выбор.
   - Допустимое расхождение полного цикла «подпись → проверка» не более ±1 % от медианы по 10 000 измерений (измеряется утилитой perf stat -e cycles с подсчётом тактов процессора).
2. Маскирование секретных векторов (Add Mask).
   - Перед вычислением формируется 256 битная маска m через DRBG.
   - Вектор фиксатор и скаляры обновляются $v \leftarrow v + m$, $r \leftarrow r + m \pmod p$.
   - Маска уничтожается вызовом explicit_bzero (или эквивалентной функцией безопасного обнуления памяти) сразу после шага, без задержки хранения в ОЗУ.
3. Контроль диапазона (Range Check).
   - Каждая координата после операции проверяется на условие $0 \le x_i < p$.
   - Нарушение диапазона интерпретируется как возможный сбой питания или же как лазерная инъекция, в таком случае реализация обнуляет ключи и выходит в аварийный режим.
4. Лимит использования ключа (Key Reuse Cap).
   - Категория 3: максимум 10 000 подписей на пару pk,sk.
   - Категория 5: максимум 100 000 подписей.
   - Таймер повторной функции генерации ключей (keygen) ведётся в неизменяемой памяти, при достижении порога ключ помечается недействительным, генерируется новый с обновлённым начальным значением (seed) из TRNG.
   - Реализации на интерпретируемых языках (Python, JavaScript) допустимы в лабораторной фазе, но на промышленном этапе должны заменяться компилируемым модулем, чтобы выдержать лимиты времени проверки (< 10 μs на Cat 5)

Перечисленные меры реализуются на уровне библиотечного кода и не влияют на размеры подписи или ключей, обеспечивая защиту от побочных временных, корреляционных и сбойных атак при сохранении целевых показателей производительности.

Заключение

В работе сформулированы и достигнуты основные цели по параметризации и валидации постквантовой электронной подписи КНАА-2-ЭЦП. Выбраны и обоснованы два набора параметров для категорий стойкости 3 и 5 по шкале NIST, обеспечивающие требуемый запас криптографической стойкости. Достигнуты целевые показатели эффективности, длина подписи ≤ 112 байт (Cat 3) и ≤ 128 байт (Cat 5) при числе операций проверки ≤ 9 200 mul и времени Verify ≤ 0,5 мкс в компилируемой реализации.

Установлены строгие требования к реализации, включая постоянное время исполнения, корректное

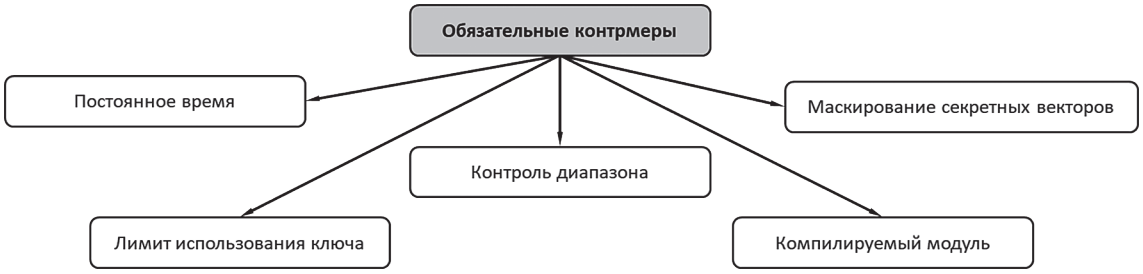


Рис.1. Обязательные контрмеры при программной реализации

использование DRBG и обязательные контрмеры. Что в свою очередь гарантирует устойчивость к атакам по побочным каналам без ухудшения целевых характеристик производительности.

Определены форматы данных ASN.1/DER с соответствующими OID идентификаторами и профили применения алгоритма для различных сценариев использования в блокчейн-сетях, HSM и IoT-устройствах. Полученные результаты создают основу для тематических испытаний и исследований КНАА-2-ЭЦП и его практического внедрения в квантово-устойчивые криптографические системы в дальнейшем.

Исследование выполнено за счет гранта Российского научного фонда № 24-41-04006.
https://rscf.ru/project/24-41-04006/

Литература

1. FIPS PUB 204. Module-Lattice-Based Digital Signature Algorithm (ML-DSA). – Gaithersburg, MD: National Institute of Standards and Technology, 2024. – 65 p. – DOI: 10.6028/NIST.FIPS.204.
2. FIPS PUB 205. Stateless Hash-Based Digital Signature Algorithm (SLH-DSA). – Gaithersburg, MD: National Institute of Standards and Technology, 2024. – 76 p. – DOI: 10.6028/NIST.FIPS.205.
3. FIPS PUB 206. Falcon Digital Signature Algorithm. – Gaithersburg, MD: National Institute of Standards and Technology, 2025. – 72 p.
4. Barker E., Chen L., Roginsky A., Mani A., Smid M., Polk T. Recommendation for Key Management, Part 1: General (Revision 5). – NIST Special Publication SP 800-57 Part 1 Rev.5. – Gaithersburg, MD: National Institute of Standards and Technology, 2020. – DOI:10.6028/NIST.SP.800-57pt1r5.
5. ISO/IEC 14888-4:2024. Information security – Cryptographic techniques – Digital signatures with appendix – Part 4: Stateful hash-based mechanisms. – Geneva: ISO/IEC, 2024.
6. ITU-T Recommendation X.680 (08/2021). Information technology – Abstract Syntax Notation One (ASN.1): Specification of basic notation. — Geneva: ITU-T, 2021.
7. Молдовян Н. А., Петренко А. С. Алгебраический алгоритм ЭЦП с двумя скрытыми группами // Вопросы кибербезопасности. 2024. № 6(64). С. 98–107. DOI: 10.21681/2311-3456-2024-6-98-107.
8. Markov, A. S., Varenitca, V. V., Arustamyan, S. S. Topical issues in the implementation of secure software development processes. (2023). In the collection: Proceedings of the International Conference on Information Processes and Systems Development and Quality Assurance. IPSQDA-2023. P. 48–53.
9. Балябин А. А., Петренко С. А. Модель блокчейн-платформы с кибериммунитетом в условиях квантовых атак // Вопросы кибербезопасности. 2025. № 3(67). С. 72–82. DOI: 10.21681/2311-3456-2025-3-72-82.
10. Петренко А. С., Петренко С. А. Основные алгоритмы квантового криптоанализа // Вопросы кибербезопасности. 2023. № 1(53). С. 100–115. DOI: 10.21681/2311-3456-2023-1-100-115.
11. Petrenko, A. S. Applied Quantum Cryptanalysis (scientific monograph). (2023). River Publishers. 256 p. ISBN 9788770227933. DOI: 10.1201/9781003392873.
12. NIST CSRC. Automated Cryptographic Validation Testing System (ACVTS) [Электронный ресурс]. — Gaithersburg, MD: National Institute of Standards and Technology, 2020–2025. – URL: https://csrc.nist.gov/projects/cryptographic-algorithm-validation-program/how-to-access-acvts (дата обращения: 22.09.2025).
13. Housley R., Fluhrer S., Kampanakis P., Westerbaan B. Use of the SLH-DSA Signature Algorithm in the Cryptographic Message Syntax (CMS). – RFC 9814. – IETF, July 2025. – DOI:10.17487/RFC9814.
14. Housley R. Update to the Cryptographic Message Syntax (CMS) for Algorithm Identifier Protection. – RFC 8933. – IETF, October 2020. – DOI:10.17487/RFC8933.

PARAMETERIZATION OF THE POST-QUANTUM ELECTRONIC SIGNATURE KNAA-2-EDS

Petrenko A. S.[2]

Keywords: *finite noncommutative associative algebra; ASN.1/DER formats; NIST persistence categories; constant runtime; signature randomization; security countermeasures; post-quantum electronic signature.*

Purpose of work *to formalize the choice of parameters and the definition of requirements for the KNAA-2-EDS, ensuring post-quantum stability (at least 2200 for Cat 3 and 2256 for Cat 5) while reducing the signature length to 112 bytes (Cat 3) and 128 bytes (Cat 5).*

Research methods: *substantiation of compliance of KNAA-2-EDS with FIPS 204/205/206, SP 800-57-1 Rev.5, ISO/IEC 14888-4:2024 and GOST R 34.10-2018 standards with precise definition of ASN.1/DER and OID formats, development of two sets of parameters for resistance categories 3 and 5 on the NIST scale, setting quantitative targets, formulation of mandatory implementation requirements, in particular constant time, correct randomization and countermeasures.*

2 Alexey S. Petrenko, Junior Researcher, Laboratory of Computer Security Problems , St. Petersburg Federal Research Center of the Russian Academy of Sciences, St. Petersburg, Russia. ORCID: https://orcid.org/0000-0002-9954-4643. Scopus Author ID: 57200260915. E-mail: a.petrenko1999@rambler.ru

Results of the study: *two sets of parameters have been established to ensure the durability of* $\geq 2^{200}$ *(Cat 3) and* $\geq 2^{256}$ *(Cat 5) and the compactness of keys and signatures, requirements for the size of keys and signatures have been defined, performance and security metrics have been set, criteria for verifying the correctness of implementation without a full ACVP cycle have been established, formats and validation procedures have been regulated, countermeasures against side attacks and constant execution time requirements are defined.*

Practical relevance: *The result of the article is that the found parameters, requirements and usage profiles demonstrate for the first time the implementation of enhanced signature randomization without doubling the verification equation, ensuring the compactness of keys and signatures while maintaining high resistance to quantum attacks and side-channel attacks. This significantly reduces the amount of transactional data and bandwidth requirements for networks and storage, which is critical for scalable blockchain systems, hardware cryptoprocessors, and energy-efficient IoT devices.*

References

1. FIPS PUB 204. Module-Lattice-Based Digital Signature Algorithm (ML-DSA). – Gaithersburg, MD: National Institute of Standards and Technology, 2024. – 65 p. – DOI:10.6028/NIST.FIPS.204.
2. FIPS PUB 205. Stateless Hash-Based Digital Signature Algorithm (SLH-DSA). – Gaithersburg, MD: National Institute of Standards and Technology, 2024. – 76 p. – DOI:10.6028/NIST.FIPS.205.
3. FIPS PUB 206. Falcon Digital Signature Algorithm. – Gaithersburg, MD: National Institute of Standards and Technology, 2025. – 72 p.
4. Barker E., Chen L., Roginsky A., Mani A., Smid M., Polk T. Recommendation for Key Management, Part 1: General (Revision 5). – NIST Special Publication SP 800-57 Part 1 Rev. 5. – Gaithersburg, MD: National Institute of Standards and Technology, 2020. – DOI:10.6028/NIST.SP.800-57pt1r5.
5. ISO/IEC 14888-4:2024. Information security – Cryptographic techniques – Digital signatures with appendix – Part 4: Stateful hash-based mechanisms. – Geneva: ISO/IEC, 2024.
6. ITU-T Recommendation X.680 (08/2021). Information technology – Abstract Syntax Notation One (ASN.1): Specification of basic notation. – Geneva: ITU-T, 2021.
7. Moldovyan, N. A., Petrenko, A. S. Algebraic signature algorithm with two hidden groups (2024), Voprosy kiberbezopasnosti [Cibersecurity issues], no. 6(64), pp. 98–107, 2024. DOI: 10.21681/2311-3456-2024-6-98-107.
8. Markov, A. S., Varenitca, V. V., Arustamyan, S. S. Topical issues in the implementation of secure software development processes. (2023). In the collection: Proceedings of the International Conference on Information Processes and Systems Development and Quality Assurance. IPSQDA-2023. P. 48–53.
9. Balyabin, A. A., Petrenko, S. A. Model of a blockchain platform with cyber-immunity under quantum attacks. (2025). Question Kiberbezopasnosti [Cybersecurity issues]. No. 3(67). P. 72–82. DOI: 10.21681/2311-3456-2025-3-72-82 (Russian Text).
10. Petrenko, A. S., Petrenko, S. A. Basic Algorithms Quantum Cryptanalysis. (2023). Question Kiberbezopasnosti [Cybersecurity issue]. No.1(53), pp. 100–115. DOI:10.21681/2311-3456-2023-1-100-115 (Russian Text).
11. Petrenko, A. S. Applied Quantum Cryptanalysis (scientific monograph). (2023). River Publishers. 256p. ISBN9788770227933. DOI:10.1201/9781003392873.
12. NIST CSRC. Automated Cryptographic Validation Testing System (ACVTS) [Electronic resource]. – Gaithersburg, MD: National Institute of Standards and Technology, 2020–2025. – URL: https://csrc.nist.gov/projects/cryptographic-algorithm-validation-program/how-to-access-acvts (access date: 09/22/2025).
13. Housley R., Fluhrer S., Kampanakis P., Westerbaan B. Use of the SLH-DSA Signature Algorithm in the Cryptographic Message Syntax (CMS). – RFC 9814. – IETF, July 2025. – DOI:10.17487/RFC9814.
14. Housley R. Update to the Cryptographic Message Syntax (CMS) for Algorithm Identifier Protection. – RFC 8933. – IETF, October 2020. – DOI:10.17487/RFC8933.

ИССЛЕДОВАНИЕ ПОДХОДОВ К РЕАЛИЗАЦИИ КВАНТОВОГО ПОВТОРИТЕЛЯ

Гончаров Р. К.[1], Киселев А. Д.[2], Егоров В. И.[3]

DOI: 10.21681/2311-3456-2025-5-96-102

Цель исследования: *провести систематизацию и критический анализ существующих подходов к подавлению ошибок в квантовых повторителях, а также в оценке их преимуществ и ограничений для реализации масштабируемых квантовых сетей и квантового интернета.*

Методы исследования: *в работе проведен детальный анализ современной литературы, включающий сопоставление различных схем квантовых повторителей, а также оценку ресурсных затрат и производительности.*

Результаты исследования: *квантовые повторители можно условно разделить на три поколения, каждое из которых демонстрирует оптимальную эффективность в определенных условиях. Первое поколение, реализующее вероятностное подавление ошибок посредством генерации объявленной запутанности и двустороннего очищения запутанности, проста в реализации и обеспечивает базовую функциональность квантовой сети, однако требует значительных временных затрат из-за необходимости обмена классическими сигналами и длительного хранения квантовых состояний. Второе поколение сочетает вероятностную генерацию запутанности с детерминированным исправлением ошибок операций посредством квантовых кодов исправления ошибок, что снижает требования к долговременной квантовой памяти и ускоряет процесс распределения запутанности, хотя обмен классическими сигналами между соседними узлами остается обязательным. Третье поколение полностью полагается на детерминированное подавление ошибок с использованием односторонней передачи классической информации, что позволяет существенно сократить временные задержки и достичь высоких скоростей генерации запутанных состояний, несмотря на необходимость более плотного расположения повторителей и высококачественных локальных вентилей. Кроме того, обзор охватывает новые направления, такие как повторители без памяти и полностью фотонные повторители. Проведенный сравнительный анализ ресурсных затрат демонстрирует, что оптимизация параметров работы повторителей является ключевым фактором для реализации масштабируемых квантовых сетей, что имеет непосредственное значение для квантового распределения ключей, квантовой метрологии и распределенных квантовых вычислений.*

Научная новизна: *научная новизна заключается в интеграции разрозненных подходов к реализации квантовых повторителей в единое целостное представление, что позволяет объективно оценить их эффективность по ключевым параметрам. Обзор подчеркивает перспективы применения новых классов повторителей, таких как повторители без памяти и полностью фотонные схемы, которые могут стать важным элементом в развитии квантового интернета.*

Ключевые слова: *квантовая сеть, запутанность, элементарное звено, классификация.*

Введение

В области квантовых коммуникаций передача квантовых состояний и генерация запутанности между удаленными сторонами существенно важны для таких приложений, как квантовое распределение ключей (КРК), квантовая метрология, распределенные вычисления и телепортация [1–3]. Обмен квантовой информации с высокой точностью воспроизведения (от англ. *fidelity*) исходного квантового состояния на больших расстояниях ограничен оптическими потерями и шумом. Фундаментальная граница PLOB[4] определяет пределы передачи в системах «точка-точка» (~200 км для оптоволокна). Невозможность клонирования квантовых состояний[5] требует альтернативных подходов: установка доверенных узлов между приемником и передатчиком, по свойствам повторяющие легитимных пользователей; применение недоверенного промежуточного узла, например MDI (от англ. *measurement-device-independent*)[6]; или же применение квантовых повторителей [4]. Последние разделяют канал передачи на сегменты (генерация запутанности) и распределяют запутанность между этими сегментами (переброс запутанности), такой подход

1 Гончаров Роман Константинович, младший научный сотрудник лаборатории квантовых коммуникаций, федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО», Санкт-Петербург. E-mail: rkgoncharov@itmo.ru

2 Киселев Алексей Донославович, доктор физико-математических наук, профессор научно-образовательного центра фотоники и оптоинформатики, федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО», Санкт-Петербург. E-mail: alexei.d.kiselev@gmail.com

3 Егоров Владимир Ильич, кандидат физико-математических наук, доцент научно-образовательного центра фотоники и оптоинформатики, федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО», Санкт-Петербург. E-mail: viegorov@itmo.ru

4 Pirandola, S. Fundamental limits of repeaterless quantum communications / S. Pirandola, R. Laurenza, C. Ottaviani, L. Banchi // Nature Communications. – 2017. – Vol. 8. – № 1. – P. 15043. DOI: 10.1038/ncomms15043.

5 Wootters, W.K. A single quantum cannot be cloned / W.K. Wootters, W.H. Zurek // Nature. – 1982. – Vol. 299. – № 5886. – P. 802-803. DOI: 10.1038/299802a0.

6 Lo, H.-K. Measurement-Device-Independent Quantum Key Distribution / H.-K. Lo, M. Curty, B. Qi // Physical Review Letters. – 2012. – Vol. 108. – № 13. – P. 130503. DOI: 10.1103/PhysRevLett.108.130503.

позволяет распределять запутанность на большие расстояния без прямой отправки состояний.

Несмотря на то, что доступные квантовые повторители пока нереализуемы широко, практические подходы уже исследуются [5–8], и появляются первые интеграции в городских квантовых сетях [9–11].

В настоящей работе обсуждается классификация квантовых повторителей в соответствии с методами подавления ошибок, проведено сравнение соответствующих схем между собой и выявляются их преимущества, а также обсуждаются перспективы развития квантовых сетей на запутанности.

Методы подавления ошибок и классификация квантовых повторителей

Квантовые коммуникации сталкиваются с двумя основными проблемами при попытке распределения запутанных состояний: потери в канале и на оборудовании и ошибки операций, вносимые каналом, квантовыми вентилями и памятью. В этом отношении квантовые повторители можно классифицировать в зависимости от используемых методов подавления ошибок: вероятностное подавление ошибок и детерминированное подавление ошибок.

Вероятностные протоколы требуют двусторонней классической связи для информирования соответствующих узлов о том, следует ли переходить к следующему шагу протокола; тем самым снижая скорость передачи данных. Для подавления ошибок потерь применяется протокол генерации объявленной (от англ. *heralded*) запутанности, а популярной схемой обнаружения ошибок операций является протокол двустороннего очищения (от англ. *distillation*) запутанности[7].

Детерминированные протоколы используют квантовые коды исправления ошибок или одностороннее очищение запутанности. Логический кубит кодируется в блок физических кубитов, которые отправляются по каналу с потерями, и далее восстанавливается при помощи квантового исправления ошибок. Для этого требуется односторонняя классическая передача информации, не влияющая на пропускную способность квантового канала. Однако такой подход может исправить ошибки только вплоть до 3 дБ потерь в канале из-за теоремы о запрете клонирования.

Направления исследований квантовых повторителей можно разделить на три поколения на основе их методов подавления ошибок потерь и ошибок операций[8]. Каждое поколение работает лучше всего для определенного набора значений рабочих параметров, таких как локальная скорость работы вентиля, точность воспроизведения вентиля и эффективность соединения элементов в схеме.

Первое поколение квантовых повторителей и протокол DLCZ

Первое поколение квантовых повторителей [12; 13] использует вероятностное подавление ошибок. Среди подобных протоколов выделяется протокол DLCZ, где запутывание осуществляется между удаленными атомными ансамблями. Этот протокол специально разработан для стимуляции квантовой памяти с помощью лазерных импульсов.

Пусть легитимные стороны разделены средой с потерями и некоторым расстоянием L. При этом в случае первого поколения требуется $N_{\mathrm{qm}} \sim 4\frac{L}{L_0} - 2$ устройств квантовой памяти [4], где L_0 — расстояние между соседними узлами. Отметим, что значение N_{qm} далее будет разниться от поколения к поколению. Среди работ о последних экспериментальных разработках в области развития квантовой памяти можно выделить следующие [14–23].

Цель протокола DLCZ состоит в запутывании возбужденных состояний посредством интерференции излучаемых соседними звеньями фотонов на светоделителе в центральном реле. Однако этот процесс может потребовать нескольких попыток, что требует от повторителей возможности длительного хранения квантовой информации из инициирующего импульса, что типично для протоколов, основанных на объявленной запутанности.

Первая реализация DLCZ включала в себя охлажденные лазером облака захваченных щелочных атомов, известных своими устойчивыми оптическими переходами[9]. Впоследствии протокол был улучшен и дополнен[10]. Появились и другие реализации, например, с использованием источника пар фотонов, а также схожие по принципу работы протоколы на базе однофотонных состояний[11] и состояний кота Шредингера [13].

Оценка производительности описанных протоколов, т.е. оценка среднего времени ожидания K (среднее количество попыток для генерации запутанности между конечными узлами) в разветвленной сети является нетривиальной задачей. Для повторителя из двух сегментов точное выражение было получено в работе[12], в то время как для произвольного количества звеньев применяются либо подход

7 Briegel, H.-J.H.-J.H.-J. Quantum repeaters: the role of imperfect local operations in quantum communication / H.-J.H.-J.H.-J. Briegel, W. Dür, J. I. Cirac, P. Zoller // Physical Review Letters. – 1998. – Vol. 81. – № 26. – P. 5932. DOI: 10.1103/PhysRevLett.81.5932.

8 Muralidharan, S. Optimal architectures for long distance quantum communication / S. Muralidharan, L. Li, J. Kim et al. // Scientific Reports. – 2016. – Vol. 6. – № 1. – P. 20463. DOI: 10.1038/srep20463.

9 Liu, C. Observation of coherent optical information storage in an atomic medium using halted light pulses / C. Liu, Z. Dutton, C. H. Behroozi, L. V. Hau // Nature. – 2001. – Vol. 409. – № 6819. – P. 490–493. DOI: 10.1038/35054017.

10 Sangouard, N. Quantum repeaters based on atomic ensembles and linear optics / N. Sangouard, C. Simon, H. De Riedmatten, N. Gisin // Reviews of Modern Physics. – 2011. – Vol. 83. – № 1. – P. 33–80. DOI: 10.1103/RevModPhys.83.33.

11 Sangouard, N. Long-distance entanglement distribution with single-photon sources / N. Sangouard, C. Simon, J. Minář et al. // Physical Review A. – 2007. – Vol. 76. – № 5. – P. 050301. DOI: 10.1103/PhysRevA.76.050301.

12 Collins, O. A. Multiplexed Memory-Insensitive Quantum Repeaters / O. A. Collins, S. D. Jenkins, A. Kuzmich, T. A. B. Kennedy // Phys. Rev. Lett. – 2007. – Vol. 98. – № 6. – P. 60502. DOI: 10.1103/PhysRevLett.98.060502.

марковских цепочек[13], требующий больший вычислительных мощностей при росте числа сегментов, либо более эффективный рекурсивный подход [24] для числа звеньев 2^k, где k — уровень вложенности. Часто используют простое приближение для случая удвоения числа элементарных звеньев:

$$K_n \approx \left(\frac{3}{2p_{\text{swap}}}\right)^k \cdot \frac{1}{p_{\text{gen}}}, \quad (1)$$

где p_{gen} — вероятность успешной генерации запутанности и p_{swap} — вероятность переброса запутанности.

Такое приближение, хоть и остается достаточно точным при высоких вероятностях генерации и переброса запутанности, несколько переоценивает время ожидания [24].

Второе поколение квантовых повторителей

Второе поколение повторителей[14] использует вероятностный подход для ошибок потерь и детерминированный — для ошибок операций. В этой архитектуре запутанность между соседними узлами устанавливается посредством генерации объявленной запутанности, а далее выполняется кодирование логического кубита с использованием квантового исправления ошибок. Скорость генерации в таком случае ограничена временной задержкой, связанной с двусторонней классической связью между соседними узлами и работой локальных вентилей. При этом, если совокупная вероятность ошибок операций остается достаточно малой, допускается реализация схемы без дополнительного кодирования, как, например, для повторителей на основе одиночных ионных кубитов [25].

Физические ресурсы, затрачиваемые в повторителе второго поколения, зависят от размера n используемого кодового блока. Для каждого повторителя требуется число кубитов, равное удвоенному размеру блока, для хранения закодированных состояний, а общее кубитов, сохраняемых в памяти, масштабируется как $N_{\text{qm}} \sim n\frac{L}{L_0}$. При этом размер кодового блока растет полилогарифмически с расстоянием, что позволяет при достижении достаточной точности локальных вентилей и малых операционных ошибок добиться более высокой скорости передачи данных по сравнению с первым поколением.

Применение мультиплексирования (наличие нескольких передатчиков и приемников на узле) может повысить вероятность успешного установления запутанности между соседними узлами, что способствует сокращению общего времени генерации запутанности [6; 26]. Однако использование двусторонней классической связи для подтверждения успешного выполнения процедур остается ограничивающим фактором по скорости, что стимулирует дальнейшее развитие схем перехода к односторонней передаче сигналов, характерной для третьего поколения.

Третье поколение квантовых повторителей

Третье поколение[15] полностью полагается на детерминированный подход, исправляя как ошибки потерь, так и ошибки операций посредством квантового исправления ошибок и одностороннего протокола хеширования. Квантовая информация сначала кодируется в блок физических кубитов, которые передаются через квантовый канал. На этапе детектирования, если уровень ошибок остается ниже допустимого порога, применяется квантовый код для восстановления исходного логического кубита, после чего блок повторно передается следующему узлу повторителя.

Отсутствие двусторонней классической связи (требуется лишь односторонняя передача сигналов для очищения запутанности) сокращает временные задержки и позволяет достигать скоростей генерации запутанности, ограниченных только временем локальных операций.

Квантовые коды исправления ошибок для третьего поколения включают коды четности[16], поверхностные коды [27], коды с использованием многоуровневых систем [28]; биноминальные коды или GKP коды[17,18]. При этом общее число квантовых ячеек определяется размером кодового блока и масштабируется аналогично второму поколению. Однако протоколы требуют плотного расположения узлов (и меньших L_0), поскольку исправление ошибок эффективно вплоть до потерь 3 дБ.

Совместное развитие квантовых повторителей различных поколений

Для количественной оценки эффективности квантовых повторителей вводится функция затрат, учитывающая как временные, так и физические ресурсы. Временной ресурс определяется задержками двусторонней классической связи (характерной для первого и второго поколений) и временем локальных операций (доминирующим для второго и третьего

13 Shchukin, E. Waiting time in quantum repeaters with probabilistic entanglement swapping / E. Shchukin, F. Schmidt, P. van Loock // Physical Review A. – 2019. – Vol. 100. – № 3. – P. 032322. DOI: 10.1103/PhysRevA.100.032322.

14 Mazurek, P. Long-distance quantum communication over noisy networks without long-time quantum memory / P. Mazurek, A. Grudka, M. Horodecki et al. // Physical Review A. – 2014. – Vol. 90. – № 6. – P. 062311. DOI: 10.1103/PhysRevA.90.062311.

15 Muralidharan, S. Ultrafast and Fault-Tolerant Quantum Communication across Long Distances / S. Muralidharan, J. Kim, N. Lütkenhaus et al. // Physical Review Letters. – 2014. – Vol. 112. – № 25. DOI: 10.1103/PhysRevLett.112.250501.

16 Ralph, T. C. Loss-Tolerant Optical Qubits / T. C. Ralph, A. J. F. Hayes, A. Gilchrist // Physical Review Letters. – 2005. – Vol. 95. – № 10. – P. 100501. DOI: 10.1103/PhysRevLett.95.100501.

17 Albert, V. V. Performance and structure of single-mode bosonic codes / V. V. Albert, K. Noh, K. Duivenvoorden et al. // Physical Review A. – 2018. – Vol. 97. – № 3. – P. 032346. DOI: 10.1103/PhysRevA.97.032346.

18 Noh, K. Quantum Capacity Bounds of Gaussian Thermal Loss Channels and Achievable Rates With Gottesman-Kitaev-Preskill Codes / K. Noh, V. V. Albert, L. Jiang // IEEE Transactions on Information Theory. – 2019. – Vol. 65. – № 4. – P. 2563–2582. DOI: 10.1109/TIT.2018.2873764.

поколений). Физические затраты отражаются в общем числе ячеек памяти, необходимых для реализации объявленной запутанности или квантовых кодов исправления ошибок. При этом функция затрат определяется как [4; 29]:

$$C(L) = \frac{N_{qm}}{r} = \frac{N}{r} \times \frac{L}{L_0}, \quad (2)$$

где r — скорость генерации потенциального ключа в бит/с; N — число кубитов, необходимое для каждой установки квантового повторителя.

Анализ показывает, что при высоких ошибках операций, доминируют схемы первого поколения. При промежуточных значениях ошибок и относительно высоких потерях на оборудовании повторители второго поколения оказываются более выгодными, а при низких потерях на оборудовании, малых ошибках в работе вентилей и высокой скорости локальных операций оптимальным является третье поколение. Таким образом, параллельно прорабатываются решения для всех трех поколений. Многообещающим подходом является использование гибридных схем, сочетающих различные типы квантовых повторителей [30] для соответствующих диапазонов длин каналов.

Существует ряд обзоров, посвященных различным аспектам квантовых повторителей [4; 26; 31; 32]. Например, обзор[19] посвящен протоколу DLCZ и его улучшениям. Классификация протоколов по трем поколениям более подробно рассматриваются[20] в [4]. В более позднем обзоре [4] также обсуждаются недавно появившиеся повторители без памяти [33] и полностью фотонные повторители [34]. Суть последних заключается в том, что они реализуются исключительно на фотонных ресурсах. При этом ожидается возможность совместной интеграции повторителей без памяти и с памятью [34]. В работе [4] особое внимание все более важной роли квантовых повторителей для развития квантового интернета [32; 35–37].

Выводы

В данной работе были рассмотрены ключевые аспекты квантовых коммуникаций при помощи квантовых повторителей, что позволило:

- обосновать преимущества и недостатки использования вероятностных и детерминированных схем в контексте оптимизации скорости и эффективности распределения запутанности;
- определить применимость различных поколений повторителей, что имеет прямое значение для развития масштабируемых квантовых сетей.

Рассмотренные подходы позволяют повысить эффективность передачи квантовой информации, расширяя допустимые расстояния вплоть до тысяч километров. Это имеет важное значение в таких научных областях, как КРК, квантовая метрология и распределенные квантовые вычисления, где устойчивость и скорость обмена квантовой информацией являются ключевыми факторами.

Отдельное внимание в обзоре уделено современным экспериментальным результатам и интеграции протоколов квантовых повторителей в существующие городские квантовые сети. Ряд исследований продемонстрировал, что предложенные схемы подтверждают свою практическую применимость. Следует также отметить, что внедрение квантовых повторителей в критическую инфраструктуру связи потребует не только решения технических задач подавления ошибок и масштабирования, но и обеспечения кибербезопасности всего программного стека, управляющего этими системами, что является предметом отдельного исследования [38].

Благодарность
Исследование выполнено за счет гранта Российского научного фонда (проект № 24-21-00484).

Литература

1. Pirandola, S. Advances in quantum cryptography / S. Pirandola, U. L. Andersen, L. Banchi et al. // Advances in Optics and Photonics. – 2020. – Vol. 12. – № 4. – P. 1012. DOI: 10.1364/AOP.361502.
2. Аверьянов, В. С. О первичных технических устройствах и требованиях к ключам безопасности квантовых систем / В. С. Аверьянов, И. Н. Карцан // Вопросы кибербезопасности. – 2023. – № 2(54). – С. 65–72. DOI: 10.21681/2311-3456-2023-2-65-72.
3. Петренко, С. А. Модель квантовых угроз безопасности информации для национальных блокчейн-экосистем и платформ / С. А. Петренко, А. А. Балябин // Вопросы кибербезопасности. – 2025. – № 1(65). – С. 7–17. DOI: 10.21681/2311-3456-2025-1-7-17.
4. Azuma, K. Quantum repeaters: From quantum networks to the quantum internet / K. Azuma, S.E. Economou, D. Elkouss et al. // Reviews of Modern Physics. – 2023. – Vol. 95. – № 4. – P. 045006. DOI: 10.1103/RevModPhys.95.045006.
5. Wallnöfer, J. Faithfully Simulating Near-Term Quantum Repeaters / J. Wallnöfer, F. Hahn, F. Wiesner et al. // PRX Quantum. – 2024. – Vol. 5. – № 1. – P. 010351. DOI: 10.1103/PRXQuantum.5.010351.

19 Sangouard, N. Quantum repeaters based on atomic ensembles and linear optics / N. Sangouard, C. Simon, H. De Riedmatten, N. Gisin // Reviews of Modern Physics. – 2011. – Vol. 83. – № 1. – P. 33–80. DOI: 10.1103/RevModPhys.83.33.

20 Muralidharan, S. Optimal architectures for long distance quantum communication / S. Muralidharan, L. Li, J. Kim et al. // Scientific Reports. – 2016. – Vol. 6. – № 1. – P. 20463. DOI: 10.1038/srep20463.

6. Chakraborty, T. Towards a spectrally multiplexed quantum repeater / T. Chakraborty, A. Das, H. van Brug et al. // npj Quantum Information. – 2025. – Vol. 11. – № 1. – P. 3. DOI: 10.1038/s41534-024-00946-2.
7. Avis, G. Analysis of multipartite entanglement distribution using a central quantum-network node / G. Avis, F. Rozpędek, S. Wehner // Phys. Rev. A. – 2023. – Vol. 107. – № 1. – P. 12609. DOI: 10.1103/PhysRevA.107.012609.
8. Krutyanskiy, V. Telecom-Wavelength Quantum Repeater Node Based on a Trapped-Ion Processor / V. Krutyanskiy, M. Canteri, M. Meraner et al. // Physical Review Letters. – 2023. – Vol. 130. – № 21. – P. 213601. DOI: 10.1103/PhysRevLett.130.213601.
9. Kucera, S. Demonstration of quantum network protocols over a 14-km urban fiber link / S. Kucera, C. Haen, E. Arenskötter et al. // npj Quantum Information. – 2024. – Vol. 10. – № 1. – P. 88. DOI: 10.1038/s41534-024-00886-x.
10. Liu, J.-L. Creation of memory–memory entanglement in a metropolitan quantum network / J.-L. Liu, X.-Y. Luo, Y. Yu et al. // Nature. – 2024. – Vol. 629. – № 8012. – P. 579–585. DOI: 10.1038/s41586-024-07308-0.
11. Knaut, C. M. Entanglement of nanophotonic quantum memory nodes in a telecom network / C. M. Knaut, A. Suleymanzade, Y. C. Wei et al. // Nature. – 2024. – Vol. 629. – № 8012. – P. 573–578. DOI: 10.1038/s41586-024-07252-z.
12. Goncharov, R. Performance of Quantum Repeaters Using Multimode Schrödinger Cat States / R. Goncharov, A. D. Kiselev, V. Egorov // Bulletin of the Russian Academy of Sciences: Physics. – 2024. – Vol. 88. – № 6. – P. 901–908. DOI: 10.1134/S1062873824706809.
13. Goncharov, R. Quantum repeaters and teleportation via entangled phase-modulated multimode coherent states / R. Goncharov, A. D. Kiselev, E. S. Moiseev et al. // Physical Review Applied. – 2023. – Vol. 20. – № 4. – P. 044030. DOI: 10.1103/PhysRevApplied.20.044030.
14. Davidson, J. H. Improved light-matter interaction for storage of quantum states of light in a thulium-doped crystal cavity / J. H. Davidson, P. Lefebvre, J. Zhang et al. // Physical Review A. – 2020. – Vol. 101. – № 4. – P. 042333. DOI: 10.1103/PhysRevA.101.042333.
15. Moiseev, E. S. Broadband quantum memory in a cavity via zero spectral dispersion / E. S. Moiseev, A. Tashchilina, S. A. Moiseev, B. C. Sanders // New Journal of Physics. – 2021. – Vol. 23. – № 6. – P. 063071. DOI: 10.1088/1367-2630/ac0754.
16. Lago-Rivera, D. Telecom-heralded entanglement between multimode solid-state quantum memories / D. Lago-Rivera, S. Grandi, J. V. Rakonjac et al. // Nature. – 2021. – Vol. 594. – № 7861. – P. 37–40. DOI: 10.1038/s41586-021-03481-8.
17. Askarani, M. F. Long-Lived Solid-State Optical Memory for High-Rate Quantum Repeaters / M. F. Askarani, A. Das, J. H. Davidson et al. // Physical Review Letters. – 2021. – Vol. 127. – № 22. – P. 220502. DOI: 10.1103/PhysRevLett.127.220502.
18. Wang, P.-C. Proposal and proof-of-principle demonstration of fast-switching broadband frequency shifting for a frequency-multiplexed quantum repeater / P.-C. Wang, O. Pietx-Casas, M. Falamarzi Askarani, G. C. do Amaral // Journal of the Optical Society of America B. – 2021. – Vol. 38. – № 4. – P. 1140. DOI: 10.1364/JOSAB.412517.
19. Bustard, P. J. Toward a Quantum Memory in a Fiber Cavity Controlled by Intracavity Frequency Translation / P. J. Bustard, K. Bonsma-Fisher, C. Hnatovsky et al. // Physical Review Letters. – 2022. – Vol. 128. – № 12. – P. 120501. DOI: 10.1103/PhysRevLett.128.120501.
20. Businger, M. Non-classical correlations over 1250 modes between telecom photons and 979-nm photons stored in $^{171}Yb^{3+}:Y_2SiO_5$ / M. Businger, L. Nicolas, T. S. Mejia et al. // Nature Communications. – 2022. – Vol. 13. – № 1. – P. 6438. DOI: 10.1038/s41467-022-33929-y.
21. Senkalla, K. Germanium Vacancy in Diamond Quantum Memory Exceeding 20 ms / K. Senkalla, G. Genov, M. H. Metsch et al. // Physical Review Letters. – 2024. – Vol. 132. – № 2. – P. 026901. DOI: 10.1103/PhysRevLett.132.026901.
22. Moiseev, S. A. Optical Quantum Memory on Macroscopic Coherence / S. A. Moiseev, K. I. Gerasimov, M. M. Minnegaliev, E. S. Moiseev // Physical Review Letters. – 2025. – Vol. 134. – № 7. – P. 070803. DOI: 10.1103/PhysRevLett.134.070803.
23. Моисеев, С. А. Оптическая квантовая память на атомных ансамблях: физические принципы, эксперименты и возможности применения в квантовом повторителе / С. А. Моисеев, М. М. Миннегалиев, К. И. Герасимов и др. // Успехи физических наук. – 2025. – Т. 195. – № 5. – С. 455–477. DOI: 10.3367/UFNr.2024.06.039694.
24. Brand, S. Efficient Computation of the Waiting Time and Fidelity in Quantum Repeater Chains / S. Brand, T. Coopmans, D. Elkouss // IEEE Journal on Selected Areas in Communications. – 2020. – Vol. 38. – № 3. – P. 619-639. DOI: 10.1109/JSAC.2020.2969037.
25. Asadi, F. K. Protocols for long-distance quantum communication with single 167 Er ions / F. K. Asadi, S. C. Wein, C. Simon // Quantum Science and Technology. – 2020. – Vol. 5. – № 4. – P. 045015. DOI: 10.1088/2058-9565/abae7c.
26. Yan, P.-S. A survey on advances of quantum repeater / P.-S. Yan, L. Zhou, W. Zhong, Y.-B. Sheng // EPL (Europhysics Letters). – 2021. – Vol. 136. – № 1. – P. 14001. DOI: 10.1209/0295-5075/ac37d0.
27. Hu, T. Quantum Network Routing Based on Surface Code Error Correction / T. Hu, J. Wu, Q. Li // 2024 IEEE 44th International Conference on Distributed Computing Systems (ICDCS). – IEEE, 2024. – P. 1236–1247.
28. Schmidt, F. Error-corrected quantum repeaters with Gottesman-Kitaev-Preskill qudits / F. Schmidt, D. Miller, P. van Loock // Physical Review A. – 2024. – Vol. 109. – № 4. – P. 042427. DOI: 10.1103/PhysRevA.109.042427.
29. Azuma, K. Networking quantum networks with minimum cost aggregation / K. Azuma // npj Quantum Information. – 2025. – Vol. 11. – № 1. – P. 51. DOI: 10.1038/s41534-025-01000-5.
30. Djordjevic, I. B. Hybrid CV-DV Quantum Communications and Quantum Networks / I.B. Djordjevic // IEEE Access. – 2022. – Vol. 10. – P. 23284–23292. DOI: 10.1109/ACCESS.2022.3154468.
31. Сукачёв, Д. Д. Протяжённые квантовые сети / Д. Д. Сукачёв // Успехи физических наук. – 2021. – Т. 191. – № 10. – С. 1077–1094. DOI: 10.3367/UFNe.2020.11.038888.
32. Azuma, K. Tools for quantum network design / K. Azuma, S. Bäuml, T. Coopmans et al. // AVS Quantum Science. – 2021. – Vol. 3. – № 1. – P. 14101. DOI: 10.1116/5.0024062.
33. Li, P.-Z. Memoryless Quantum Repeaters Based on Cavity-QED and Coherent States / P.-Z. Li, P. van Loock // Advanced Quantum Technologies. – 2023. – Vol. 6. – № 8. – P. 2200151. DOI: 10.1002/qute.202200151.
34. Benchasattabuse, N. Engineering Challenges in All-Photonic Quantum Repeaters / N. Benchasattabuse, M. Hajdušek, R. Van Meter // IEEE Network. – 2025. – Vol. 39. – № 1. – P. 132–139. DOI: 10.1109/MNET.2024.3411802.
35. Wei, S. Towards Real-World Quantum Networks: A Review / S. Wei, B. Jing, X. Zhang et al. // Laser & Photonics Reviews. – 2022. – Vol. 16. – № 3. – P. 2100219. DOI: 10.1002/lpor.202100219.
36. Li, Y. A Survey of Quantum Internet Protocols From a Layered Perspective / Y. Li, H. Zhang, C. Zhang et al. // IEEE Communications Surveys & Tutorials. – 2024. – Vol. PP. – P. 1-1. DOI: 10.1109/COMST.2024.3361662.

37. Singh, A. Quantum Internet—Applications, Functionalities, Enabling Technologies, Challenges, and Research Directions / A. Singh, K. Dev, H. Siljak et al. // IEEE Communications Surveys & Tutorials. – 2021. – Vol. 23. – № 4. – P. 2218-2247. DOI: 10.1109/COMST.2021.3109944.
38. Марков, А. С. Важная веха в безопасности открытого программного обеспечения / А. С. Марков // Вопросы кибербезопасности. – 2023. – № 1(53). – С. 2–12. DOI: 10.21681/2311-3456-2023-1-2-12.

RESEARCH OF APPROACHES TO THE IMPLEMENTATION OF A QUANTUM REPEATER

Goncharov R.[21], Kiselev A. D.[22], Egorov V.[23]

Keywords: *quantum network, entanglement, elementary link, classification.*

Purpose of the study: *systematization and critical analysis of existing approaches to error suppression in quantum repeaters, as well as to assess their advantages and limitations for the implementation of scalable quantum networks and the quantum internet.*

Methods of research: *the paper provides a detailed analysis of modern literature, including a comparison of various quantum repeater schemes, as well as an assessment of resource costs and performance.*

Result(s): *quantum repeaters can be divided into three generations, each of which demonstrates optimal efficiency under certain conditions. The first generation, implementing probabilistic error suppression by generating heralded entanglement and two-way entanglement distillation, is easy to implement and provides the basic functionality of a quantum network, but requires significant time due to the need to exchange classical signals and long-term storage of quantum states. The second generation combines probabilistic entanglement generation with deterministic operation error suppression using quantum error correction, which reduces the requirements for long-term quantum memory, although the exchange of classical signals between neighboring nodes remains mandatory. The third generation relies entirely on deterministic error suppression using one-way classical communication, which allows for a significant reduction in time delays and high entanglement generation rates, despite the need for a denser arrangement of repeaters and high-quality local gates. In addition, the review covers new areas such as memoryless repeaters and all-photonic repeaters. The conducted comparative analysis of resource costs demonstrates that optimization of repeater operating parameters is a key factor for the implementation of scalable quantum networks, which is of direct importance for quantum key distribution, quantum metrology, and distributed quantum computing.*

Scientific novelty: *the scientific novelty lies in the integration of disparate approaches to the implementation of quantum repeaters into a single holistic representation, which allows for an objective assessment of their efficiency by key parameters. The review highlights the potential for new classes of repeaters, such as memoryless repeaters and all-photonic circuits, to become important elements in the development of a quantum internet.*

References

1. Pirandola, S. Advances in quantum cryptography / S. Pirandola, U. L. Andersen, L. Banchi et al. // Advances in Optics and Photonics. – 2020. – Vol. 12. – № 4. – P. 1012. DOI: 10.1364/AOP.361502.
2. Aver'janov, V. S. O pervichnyh tehnicheskih ustrojstvah i trebovanijah k kljucham bezopasnosti kvantovyh sistem / V. S. Aver'janov, I. N. Karcan // Voprosy kiberbezopasnosti. – 2023. – № 2(54). – S. 65–72. DOI: 10.21681/2311-3456-2023-2-65-72.
3. Petrenko, S. A. Model' kvantovyh ugroz bezopasnosti informacii dlja nacional'nyh blokchejn-jekosistem i platform / S. A. Petrenko, A. A. Baljabin // Voprosy kiberbezopasnosti. – 2025. – № 1(65). – S. 7–17. DOI: 10.21681/2311-3456-2025-1-7-17.
4. Azuma, K. Quantum repeaters: From quantum networks to the quantum internet / K. Azuma, S. E. Economou, D. Elkouss et al. // Reviews of Modern Physics. – 2023. – Vol. 95. – № 4. – P. 045006. DOI: 10.1103/RevModPhys.95.045006.
5. Wallnöfer, J. Faithfully Simulating Near-Term Quantum Repeaters / J. Wallnöfer, F. Hahn, F. Wiesner et al. // PRX Quantum. – 2024. – Vol. 5. – № 1. – P. 010351. DOI: 10.1103/PRXQuantum.5.010351.
6. Chakraborty, T. Towards a spectrally multiplexed quantum repeater / T. Chakraborty, A. Das, H. van Brug et al. // npj Quantum Information. – 2025. – Vol. 11. – № 1. – P. 3. DOI: 10.1038/s41534-024-00946-2.
7. Avis, G. Analysis of multipartite entanglement distribution using a central quantum-network node / G. Avis, F. Rozpędek, S. Wehner // Phys. Rev. A. – 2023. – Vol. 107. – № 1. – P. 12609. DOI: 10.1103/PhysRevA.107.012609.
8. Krutyanskiy, V. Telecom-Wavelength Quantum Repeater Node Based on a Trapped-Ion Processor / V. Krutyanskiy, M. Canteri, M. Meraner et al. // Physical Review Letters. – 2023. – Vol. 130. – № 21. – P. 213601. DOI: 10.1103/PhysRevLett.130.213601.
9. Kucera, S. Demonstration of quantum network protocols over a 14-km urban fiber link / S. Kucera, C. Haen, E. Arenskötter et al. // npj Quantum Information. – 2024. – Vol. 10. – № 1. – P. 88. DOI: 10.1038/s41534-024-00886-x.

21 Roman Goncharov, junior research fellow, Laboratory for quantum communications, ITMO University, Saint Petersburg. E-mail: rkgoncharov@itmo.ru
22 Aleksei D. Kiselev, Dr.Sc., Associate Professor, Research and Educational Center for Photonics and Optical IT, ITMO University, Saint Petersburg. E-mail: alexei.d.kiselev@gmail.com
23 Vladimir Egorov, Ph.D., Associate Professor, Research and Educational Center for Photonics and Optical IT, ITMO University, Saint Petersburg. E-mail: viegorov@itmo.ru

10. Liu, J.-L. Creation of memory–memory entanglement in a metropolitan quantum network / J.-L. Liu, X.-Y. Luo, Y. Yu et al. // Nature. – 2024. – Vol. 629. – № 8012. – P. 579–585. DOI: 10.1038/s41586-024-07308-0.
11. Knaut, C. M. Entanglement of nanophotonic quantum memory nodes in a telecom network / C. M. Knaut, A. Suleymanzade, Y. C. Wei et al. // Nature. – 2024. – Vol. 629. – № 8012. – P. 573–578. DOI: 10.1038/s41586-024-07252-z.
12. Goncharov, R. Performance of Quantum Repeaters Using Multimode Schrödinger Cat States / R. Goncharov, A. D. Kiselev, V. Egorov // Bulletin of the Russian Academy of Sciences: Physics. – 2024. – Vol. 88. – № 6. – P. 901–908. DOI: 10.1134/S1062873824706809.
13. Goncharov, R. Quantum repeaters and teleportation via entangled phase-modulated multimode coherent states / R. Goncharov, A. D. Kiselev, E. S. Moiseev et al. // Physical Review Applied. – 2023. – Vol. 20. – № 4. – P. 044030. DOI: 10.1103/PhysRevApplied.20.044030.
14. Davidson, J. H. Improved light-matter interaction for storage of quantum states of light in a thulium-doped crystal cavity / J. H. Davidson, P. Lefebvre, J. Zhang et al. // Physical Review A. – 2020. – Vol. 101. – № 4. – P. 042333. DOI: 10.1103/PhysRevA.101.042333.
15. Moiseev, E. S. Broadband quantum memory in a cavity via zero spectral dispersion / E. S. Moiseev, A. Tashchilina, S. A. Moiseev, B. C. Sanders // New Journal of Physics. – 2021. – Vol. 23. – № 6. – P. 063071. DOI: 10.1088/1367-2630/ac0754.
16. Lago-Rivera, D. Telecom-heralded entanglement between multimode solid-state quantum memories / D. Lago-Rivera, S. Grandi, J. V. Rakonjac et al. // Nature. – 2021. – Vol. 594. – № 7861. – P. 37–40. DOI: 10.1038/s41586-021-03481-8.
17. Askarani, M. F. Long-Lived Solid-State Optical Memory for High-Rate Quantum Repeaters / M. F. Askarani, A. Das, J. H. Davidson et al. // Physical Review Letters. – 2021. – Vol. 127. – № 22. – P. 220502. DOI: 10.1103/PhysRevLett.127.220502.
18. Wang, P.-C. Proposal and proof-of-principle demonstration of fast-switching broadband frequency shifting for a frequency-multiplexed quantum repeater / P.-C. Wang, O. Pietx-Casas, M. Falamarzi Askarani, G. C. do Amaral // Journal of the Optical Society of America B. – 2021. – Vol. 38. – № 4. – P. 1140. DOI: 10.1364/JOSAB.412517.
19. Bustard, P. J. Toward a Quantum Memory in a Fiber Cavity Controlled by Intracavity Frequency Translation / P. J. Bustard, K. Bonsma-Fisher, C. Hnatovsky et al. // Physical Review Letters. – 2022. – Vol. 128. – № 12. – P. 120501. DOI: 10.1103/PhysRevLett.128.120501.
20. Businger, M. Non-classical correlations over 1250 modes between telecom photons and 979-nm photons stored in $^{171}Yb^{3+}$:Y_2SiO_5 / M. Businger, L. Nicolas, T. S. Mejia et al. // Nature Communications. – 2022. – Vol. 13. – № 1. – P. 6438. DOI: 10.1038/s41467-022-33929-y.
21. Senkalla, K. Germanium Vacancy in Diamond Quantum Memory Exceeding 20 ms / K. Senkalla, G. Genov, M. H. Metsch et al. // Physical Review Letters. – 2024. – Vol. 132. – № 2. – P. 026901. DOI: 10.1103/PhysRevLett.132.026901.
22. Moiseev, S. A. Optical Quantum Memory on Macroscopic Coherence / S. A. Moiseev, K. I. Gerasimov, M. M. Minnegaliev, E. S. Moiseev // Physical Review Letters. – 2025. – Vol. 134. – № 7. – P. 070803. DOI: 10.1103/PhysRevLett.134.070803.
23. Moiseev, S. A. Opticheskaja kvantovaja pamjat' na atomnyh ansambljah: fizicheskie principy, jeksperimenty i vozmozhnosti primenenija v kvantovom povtoritele / S. A. Moiseev, M. M. Minnegaliev, K. I. Gerasimov i dr. // Uspehi fizicheskih nauk. – 2025. – T. 195. – № 5. – S. 455–477. DOI: 10.3367/UFNr.2024.06.039694.
24. Brand, S. Efficient Computation of the Waiting Time and Fidelity in Quantum Repeater Chains / S. Brand, T. Coopmans, D. Elkouss // IEEE Journal on Selected Areas in Communications. – 2020. – Vol. 38. – № 3. – P. 619–639. DOI: 10.1109/JSAC.2020.2969037.
25. Asadi, F. K. Protocols for long-distance quantum communication with single 167 Er ions / F. K. Asadi, S. C. Wein, C. Simon // Quantum Science and Technology. – 2020. – Vol. 5. – № 4. – P. 045015. DOI: 10.1088/2058-9565/abae7c.
26. Yan, P.-S. A survey on advances of quantum repeater / P.-S. Yan, L. Zhou, W. Zhong, Y.-B. Sheng // EPL (Europhysics Letters). – 2021. – Vol. 136. – № 1. – P. 14001. DOI: 10.1209/0295-5075/ac37d0.
27. Hu, T. Quantum Network Routing Based on Surface Code Error Correction / T. Hu, J. Wu, Q. Li // 2024 IEEE 44th International Conference on Distributed Computing Systems (ICDCS). – IEEE, 2024. – P. 1236–1247.
28. Schmidt, F. Error-corrected quantum repeaters with Gottesman-Kitaev-Preskill qudits / F. Schmidt, D. Miller, P. van Loock // Physical Review A. – 2024. – Vol. 109. – № 4. – P. 042427. DOI: 10.1103/PhysRevA.109.042427.
29. Azuma, K. Networking quantum networks with minimum cost aggregation / K. Azuma // npj Quantum Information. – 2025. – Vol. 11. – № 1. – P. 51. DOI: 10.1038/s41534-025-01000-5.
30. Djordjevic, I. B. Hybrid CV-DV Quantum Communications and Quantum Networks / I. B. Djordjevic // IEEE Access. – 2022. – Vol. 10. – P. 23284–23292. DOI: 10.1109/ACCESS.2022.3154468.
31. Sukachjov, D. D. Protjazhjonnye kvantovye seti / D. D. Sukachjov // Uspehi fizicheskih nauk. – 2021. – T. 191. – № 10. – S. 1077–1094. DOI: 10.3367/UFNe.2020.11.038888.
32. Azuma, K. Tools for quantum network design / K. Azuma, S. Bäuml, T. Coopmans et al. // AVS Quantum Science. – 2021. – Vol. 3. – № 1. – P. 14101. DOI: 10.1116/5.0024062.
33. Li, P.-Z. Memoryless Quantum Repeaters Based on Cavity-QED and Coherent States / P.-Z. Li, P. van Loock // Advanced Quantum Technologies. – 2023. – Vol. 6. – № 8. – P. 2200151. DOI: 10.1002/qute.202200151.
34. Benchasattabuse, N. Engineering Challenges in All-Photonic Quantum Repeaters / N. Benchasattabuse, M. Hajdušek, R. Van Meter // IEEE Network. – 2025. – Vol. 39. – № 1. – P. 132–139. DOI: 10.1109/MNET.2024.3411802.
35. Wei, S. Towards Real-World Quantum Networks: A Review / S. Wei, B. Jing, X. Zhang et al. // Laser & Photonics Reviews. – 2022. – Vol. 16. – № 3. – P. 2100219. DOI: 10.1002/lpor.202100219.
36. Li, Y. A Survey of Quantum Internet Protocols From a Layered Perspective / Y. Li, H. Zhang, C. Zhang et al. // IEEE Communications Surveys & Tutorials. – 2024. – Vol. PP. – P. 1-1. DOI: 10.1109/COMST.2024.3361662.
37. Singh, A. Quantum Internet—Applications, Functionalities, Enabling Technologies, Challenges, and Research Directions / A. Singh, K. Dev, H. Siljak et al. // IEEE Communications Surveys & Tutorials. – 2021. – Vol. 23. – № 4. – P. 2218–2247. DOI: 10.1109/COMST.2021.3109944.
38. Markov, A. S. Vazhnaja veha v bezopasnosti otkrytogo programmnogo obespechenija / A. S. Markov // Voprosy kiberbezopasnosti. – 2023. – № 1(53). – S. 2–12. DOI: 10.21681/2311-3456-2023-1-2-12.

КРИТЕРИИ И ПОКАЗАТЕЛИ КОНСТРУКТИВНОЙ ЗАЩИТЫ РАСПРЕДЕЛЕННОГО РЕЕСТРА

Сундеев П. В.[1]

DOI: 10.21681/2311-3456-2025-5-103-108

Цель исследования: *исследовать проблему формальных критериев и показателей для оценки конструктивной защиты информационных систем с учетом особенностей архитектуры распределенного реестра в условиях квантовой угрозы, предложить подход к формированию критериев конструктивной защиты, позволяющих автоматизировать анализ и синтез безопасной архитектуры.*

Методы исследования: *объектно-ориентированный анализ сложных систем, системный анализ, теория модульно-кластерных сетей, теория графов, теория матриц, математическая логика.*

Результат исследования: *предложен подход к формированию системы критериев и показателей конструктивной защиты архитектуры информационной системы на основе кластерной модели защиты с полным перекрытием, представленной в виде кластерного мультиграфа модульно-кластерной сети, и анализа ее топологии. Оценка безопасности архитектуры по критериям конструктивной защиты заключается в формальном анализе наличия или отсутствия определенных типов вершин и дуг кластерного мультиграфа. Показатель непрерывности конструктивной защиты для оценки безопасности траекторий информационного процесса основан на сравнении весов помеченных вершин, обозначающих модули защиты, с установленным значением. Показана применимость критериев конструктивной защиты к оценке технологии распределённого реестра.*

Научная новизна: *разработан новый подход к формированию формальных критериев и показателей конструктивной защиты информационных систем на основе кластерной модели защиты с полным перекрытием в терминах теории модульно-кластерных сетей.*

Ключевые слова: *модульно-кластерная сеть, квантовая угроза, защита информации.*

Введение

При разработке защищенных информационных систем необходимо оценивать безопасность архитектуры, программного кода, учитывать политику доступа [1–4]. Критериями безопасности архитектуры, как правило, являются требования к защите информации и их реализации, установленные стандартами[2] и нормативными актами, которые не формальны и не позволяют доказывать безопасность архитектуры. Формальные показатели эффективности средств защиты информации, например, защиты от побочных электромагнитных излучений, стойкости криптографии, безопасности программного кода и т.д. применимы к оценке элементов, но не позволяют оценить безопасность архитектуры системы. В случае критических систем оценка должна быть формальной, чтобы минимизировать риск нарушения безопасности информации с неприемлемым ущербом. Технология распределенного реестра (DLT) используется в национальных блокчейн-экосистемах, критических приложениях в финансовой, медицинской и других сферах деятельности. Из-за нарушения безопасности информации может быть нанесен неприемлемый ущерб. Поэтому обеспечение конструктивной защиты архитектуры является непременным условием безопасности систем распределенного реестра (DLTS). Особенностью конструктивной безопасности «классической» архитектуры DLT является конструктивная защита, которая обеспечивается децентрализованной топологией информационного взаимодействия недоверенных субъектов и криптографией[3]. Масштабное использование криптографии с неидеальной стойкостью создает высокий риск компрометации из-за «квантовой угрозы», связанной с появлением эффективных методов атаки на ключевые системы, и квантовых компьютеров с высокой скоростью вычислений [5–9].

Система критериев конструктивной защиты позволяет оценивать безопасность архитектуры по результатам анализа топологии мультиграфа, который является информационной моделью системы с учетом ее защиты. В зависимости от постановки задачи при оценке безопасности может проверяться наличие или отсутствие определенных типов вершин, дуг и их веса. Применение формальных

1 Сундеев Павел Викторович, доктор технических наук, главный инженер-исследователь, Научный центр информационных технологий и искусственного интеллекта, Научно-технологический университет «Сириус», Адрес: Россия, 354340, Краснодарский край, Федеральная территория «Сириус», пгт Сириус, проспект Олимпийский, д.1., ORCID: 0009-0005-8442-8921, E-mail: sundeev.pv@talantiuspeh.ru

2 ГОСТ Р 72118-2025 «Защита информации. Системы с конструктивной информационной безопасностью. Методология разработки».

3 Recommendation ITU-T X.1410 (03/2023), Distributed ledger technology (DLT) security. Security architecture of data sharing management based on the distributed ledger technology.

критериев позволяет автоматизировать процесс оценки конструктивной защиты информационной системы и улучшить ее достоверность.

Концепция оценки конструктивной защиты

Термин «конструктивная защита» информационной системы может трактоваться в узком смысле, когда обработка информации технологически невозможна без использования функций защиты, и в широком смысле, когда к ней относят дополнительные функции (средства) защиты, которые компенсируют отсутствующие функции защиты и уязвимости архитектуры. Примерами конструктивной защиты могут быть технология квантового распределения ключей, биометрическая аутентификация, DLT.

Критерии конструктивной защиты информационной системы разработаны на основе положений теории модульно-кластерных сетей (МК-сеть) с использованием кластерной модели защиты информации с полным перекрытием [10]. В терминах МК-сетей конструктивная защита информации рассматривается как архитектурные ограничения топологии информационного взаимодействия, а также функции защиты, которые реализуют управление доступом и компенсируют недостатки архитектуры с целью ограничения несанкционированного взаимодействия. Формальным критерием или показателем конструктивной защиты в терминах теории МК-сетей может быть оценка наличия, отсутствия и свойств элементов топологии кластерного мультиграфа (вершин, дуг и их весов), которые влияют на доступ субъектов к объектам.

Для формальной оценки безопасности информационная система и ее окружение представляются в виде объектно-ориентированной модели в терминах теории МК-сетей, в которых используется декомпозиция информационных взаимодействий на физические (F), синтаксические (L) и семантические (S) отношения между модулями [10]. Если объектно-ориентированную информационную модель Q представить в виде графа, то элементы множеств M_N и R_M^{FLS} будут составлять множество вершин кластерного мультиграфа $G^{FLS}(M_N,R_M^{FLS})$, где N – число вершин, и R_M^{FLS} – множество дуг, которые обозначают информационные связи между вершинами, обозначающими физические и логические модули, имеющие функциональное значение в процессах обработки и защиты информации. При этом множество дуг R_M^{FLS} состоит из подмножеств внутренних дуг инцидентных вершинам одного кластера R^K и подмножества внешних дуг инцидентных вершинам из разных кластеров $R^{\bar{K}}$, таких что $R^K \cap R^{\bar{K}} = \emptyset$. Таким образом, модульно-кластерная сеть представляет собой FLS-мультиграф $G^{FLS}(M_N,R_M^{FLS})$, вершины которого обозначают модули $M_N = \{m_1,m,\ldots,m_n\}$, где N их число, а дуги $R_M^{FLS} = \{R^F \cup R^L \cup R^S\}$, где $R^F = \{r_x^F\}$, $R^L = \{r_y^L\}$, $R^S = \{r_z^S\}$, определяются наличием у модулей FLS-интерфейсов, через которые возможна реализация информационных FLS-отношений.

Декомпозиция элементов множества дуг $R_M \longrightarrow R_M^{FLS}$ расширяет понятие «смежности» вершин графа.

Утверждение 1. Любые две вершины $\{m_i,m_j\}$ мультиграфа $G^{FLS}(M_N,R_M^{FLS})$ являются смежными, если и только если в остовных FLS-подграфах между этими вершинами существует хотя бы одно подмножество кратных дуг вида $\{r_{i,j}^F,r_{i,j}^L,r_{i,j}^S\}$, которое называется полной FLS-дугой.

Утверждение 2. Если любые две вершины $\{m_i,m_j\}$ графа $G^{FLS}(M_N,R_M^{FLS})$, обозначающие модули МК-сети, смежные, то возможно информационное взаимодействие между этими модулями.

Утверждение 3. Путь P_{ij}^{FLS} между произвольной парой вершин $\{m_i,m_j\}$ в FLS-мультиграфе $G^{FLS}(M_N,R_M^{FLS})$ существует, если и только если существуют пути P_{ij}^F, P_{ij}^L, P_{ij}^S между этими вершинами в остовных FLS-подграфах.

Критериями распределения вершин по кластерам являются одинаковые для них внешние правила доступа, установленные в эталонной кластерной матрице, которые в свою очередь определяются принадлежностью модуля к системе W, окружению V или средствам защиты D, а также правилами разграничения доступа и системными критериями конструктивной защиты.

Политика доступа разбивает систему на множество кластеров $K = \{K^W,K^V,K^D\}$. Для вершин одного кластера установлены общие правила доступа, то есть взаимодействие не ограничено. Для вершин разных кластеров установлены разные правила доступа, то есть взаимодействие ограничено или запрещено. Возможна вложенность кластеров с иерархическими правилами доступа. Если правила выполняются для произвольной пары вершин $\{m_i,m_j\}$ и отсутствует путь P_{ij}^{FLS} между вершинами разных кластеров мультиграфа, между которыми политикой доступа запрещено взаимодействие, то архитектура безопасна. При этом политика доступа должна разрешать взаимодействие с модулями подмножества M_d^D, обозначающими модули защиты из кластеров подмножества K^D, через которые осуществляется управление доступом и в соответствии с политикой может быть создана дуга в остовном F, L или S подграфе для образования полной FLS-дуги.

Таким образом, кластерная структура системы Q определяется на множествах модулей $M_N=\{M_w^W,M_v^V,M_d^D\}$ информационной системы W и ее окружения V, множествах FLS-отношений $R_M^{FLS} = \{R^F,R^L,R^S\}$ и кластеров $K_Q = \{K^W,K^V,K^D\}$,

$$G_Q^{FLS}(M_N,R_M^{FLS}) \equiv$$
$$\equiv \begin{cases} M_N = \{M_w^W,M_v^V,M_d^D\} | \{M_w^W \cap M_v^V \cap M_d^D\} = \emptyset; \\ R_M^{FLS} = \{R_W^{FLS};R_V^{FLS};R_D^{FLS};R_{W^*}^{FLS};R_{V^*}^{FLS};R_{D^*}^{FLS}\}; \\ K_Q = \{K^W,K^V,K^D\} | \{K^W \cap K^V \cap K^D\} = \emptyset; \end{cases} \quad (1)$$

где $\{R_W^{FLS};R_V^{FLS};R_D^{FLS};R_{W^*}^{FLS};R_{V^*}^{FLS};R_{D^*}^{FLS}\}$ – подмножества полных внутренних $\{R_W^{FLS};R_V^{FLS};R_D^{FLS}\}$ и внешних $\{R_{W^*}^{FLS};R_{V^*}^{FLS};R_{D^*}^{FLS}\}$ FLS-дуг кластеров $\{K^W,K^V,K^D\}$.

Системные критерии конструктивной защиты определяют наличие или отсутствие вершин или дуг внутри и между кластерными подмножествами K_Q мультиграфа $G^{FLS}(M_N,R_M^{FLS})$, что указывает на наличие или отсутствие пути P_{ij}^{FLS} между любыми произвольными вершинами $\{m_i,m_j\}$ из кластерных подмножеств K^W и K^V или разных кластеров подмножества K^W. Для разных топологий и уровней защиты может использоваться один или несколько критериев конструктивной защиты. При выборе критериев должны учитываться топология информационного взаимодействия, уровни защиты и доверия, которые необходимо обеспечить для конкретной системы, а также используемые технологии обработки и защиты информации.

Формальные критерии и показатели конструктивной защиты

Рассмотрены критерии и показатели конструктивной защиты для наиболее распространённых случаев.

Критерий 1. Если все внешние дуги $R_{W^*}^{FLS}$ кластера K^W инцидентны вершинам из кластера K^D, то конструктивно архитектура системы безопасна.

В этом случае кластеры K^W и K^V не имеют смежных вершин. Критерий актуален для информационных систем с единой политикой для всех собственных субъектов и объектов доступа. Соответственно, оценка конструктивной защиты архитектуры проводится поиском полных FLS-дуг между вершинами кластеров K^W и K^V. Критерий актуален для случая периметровой защиты при отсутствии разделения прав доступа между субъектами внутри системы.

Более сильный критерий 2 конструктивной защиты учитывает внутренние угрозы, связанные с разными политиками доступа для разных кластеров системы W.

Критерий 2. Если все внутренние R_W^{FLS} и внешние $R_{W^*}^{FLS}$ дуги кластеров $\{K_1^W,K_2^W,...,K_N^W\} \subseteq K^W$ инцидентны вершинам из кластеров K^D, то конструктивно архитектура системы безопасна.

Критерий актуален для случая групповых политик, когда в информационной системе установлены одинаковые правила доступа в выделенных группах субъектов и объектов доступа.

Для взаимодействия недоверенных субъектов актуален критерий конструктивной защиты, предусматривающий возможность взаимодействия только через средства защиты. В случае однотипной политики доступа для всех субъектов, например, в «классической» архитектуре DLT, все вершины становятся кластерами, что приводит к трактовке критерия 1.

Критерий 3. Если каждый кластер $\{K_1^W,K_2^W,...,K_N^W\} \subseteq K^W$ имеет только одну вершину и внешние дуги $R_{W^*}^{FLS}$ кластеров K^W инцидентны вершинам из кластеров K^D, то конструктивно архитектура системы безопасна.

Оценка конструктивной защиты архитектуры по критерию 3 заключается в проверке отсутствия полных FLS-дуг между вершинами кластеров $K_1^W,K_2^W,...,K_N^W$ и с вершинами из K^V, наличия только одной вершины в каждомкластере K^W и наличия у него FLS-дуги с вершиной из кластера K^D.

Следующий критерий безопасности актуален для критических информационных систем с необходимостью обеспечения относительно высокого уровня доверия. В этом случае в каждый кластер $K_1^W,K_2^W,...,K_N^W$ включаются кластеры K^D или вершины из подмножества M_d^D, которые обозначают функции (средства) защиты информации. Критерий применим, когда каждый субъект имеет собственные «доверенные» средства защиты информации и может управлять доступом к своим объектам в кластере без участия посредников.

Критерий 4. Если каждый кластер $K_1^W,K_2^W,...,K_N^W$ имеет вершины из подмножества M_d^D и внешние дуги $R_{W^*}^{FLS}$ каждого кластера инцидентны только этим вершинам кластера, то конструктивно архитектура системы безопасна.

На практике чаще встречаются гибридные схемы с встроенными в K^W и дополнительными средствами защиты из подмножества M_d^D, для оценки которых необходимо использовать несколько критериев.

Для всех критериев средства защиты из подмножества M_d^D должны иметь установленный уровень доверия, выбор которого зависит от оценки рисков. Уровень доверия может учитываться в весах вершин.

При использовании критериев конструктивной защиты анализ безопасности топологии архитектуры сводится к поиску вершин, которые должны отсутствовать или присутствовать в МК-модели системы, и в оценке смежности определенных типов вершин мультиграфа. В систему критериев могут быть включены требования к наличию, смежности и месту размещения в мультиграфе вершин, обозначающих модули с определенными функциями защиты. Сравнение значений весов определенных типов вершин мультиграфа с установленным знамением, является показателем уровня защиты.

Высокий уровень конструктивной защиты в узкой трактовке термина дает использование функций

защиты технологически не отделимых от процедур обработки информации. Например, такими технологиями являются квантовое распределение ключей и биометрическая аутентификация. Формально в мультиграфе такие свойства могут задаваться объединением функций защиты в один кластер средств защиты, например, K^{FL}, а также весами дуг, на одном или более уровней FLS-отношений, рассчитанных по внешним методикам.

Показателем эффективности элементов защиты являются веса вершин или дуг мультиграфа. Если значение весов вершин из подмножества M_d^D принадлежащих пути P_{ij}^{FLS} между произвольной парой вершин $\{m_i, m_j\}$ в FLS-мультиграфе $G^{FLS}(M_N, R_M^{FLS})$ не ниже установленного значения, то конструктивно архитектура системы безопасна.

Значение веса, например, для средств криптографической защиты DLT может быть задано значением показателя стойкости криптографии с учетом «квантовой угрозы» [6–9]. При заданных значениях весов вершин оценка проводится сравнением веса каждой вершины с установленным значением с учетом требуемого уровня защиты.

Для учета надежности функций защиты в оценке безопасности архитектуры вершинам мультиграфа из подмножества M_d^D присваиваются нормированные весовые коэффициенты, характеризующие надежность защиты на основе внешних данных

$$D = |d_1^{(k)}, d_2^{(k)}, \ldots, d_n^{(k)}|. \quad (2)$$

Каждой d-ой вершине приписывается вес, где k – функция веса. Непрерывность уровня защиты оценивается сравнением значений весовых коэффициентов помеченных вершин графа для каждого пути относительно нормированного заданного значения, которое может быть установлено классом защиты и уровнем доверия к средству защиты или расчетным значением, например, стойкости криптографии к атакам с использованием квантового суперкомпьютера с высокой скоростью вычислений. Например, компания IBM анонсировала технологический прорыв в обработке ошибок и создание промышленного квантового компьютера на 200 логических кубитов к 2029 году [12, 13], что приближает горизонт реализации «квантовой» угрозы для широко используемых информационных систем, в которых применяется асимметричная или не стойкая симметричная криптография. В системах, для которых криптография является основой конструктивной защиты, например, для национальных распределенных реестров, разработка надежных постквантовых криптографических алгоритмов становится приоритетной задачей [14].

Кластерная модель защиты позволяет продемонстрировать использование системных критериев для оценки конструктивной защиты архитектуры на FLS-уровнях информационных отношений. На рис. 1 пример а) демонстрирует неконтролируемое взаимодействие $R_{v.w}^{FLS}$ модулей v_z и w_y из кластеров K_3^V и K_2^W, что по критерию 1 оценивается как угроза безопасности.

В примере б) показана блокировка связи между модулями v_z и w_y кластеров K_3^V и K_2^W средством защиты семантического уровня $d_s^{K_S^D}$ из подмножества

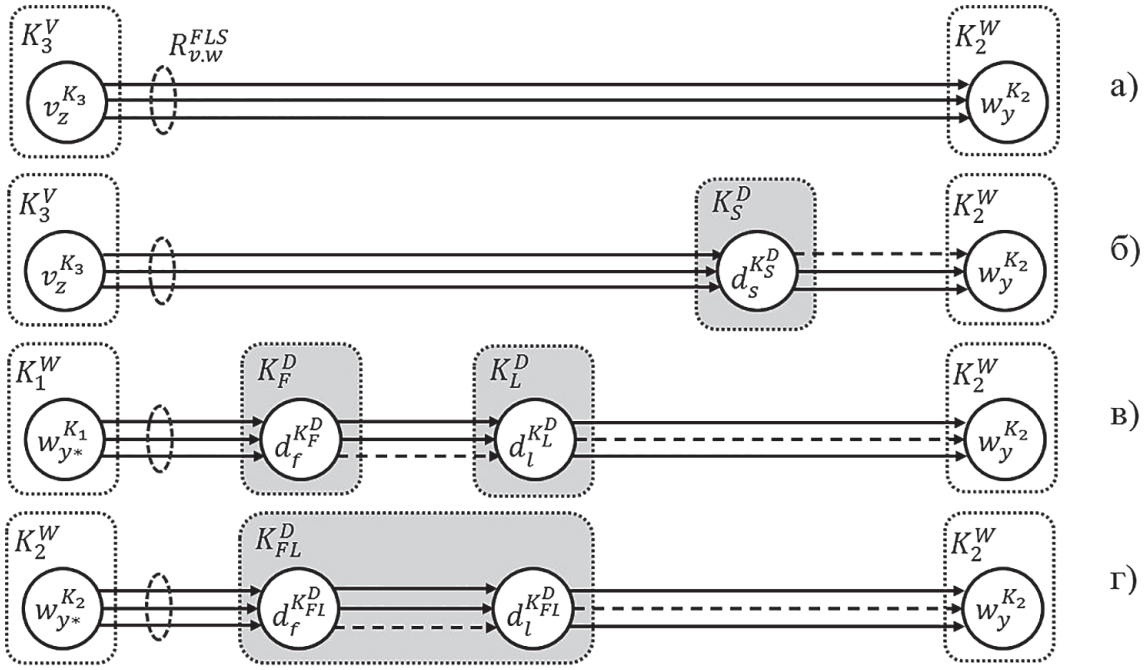


Рис. 1. Критерии конструктивной защиты архитектуры

M_d^D, например, средством разграничения доступа. В примере в) показано взаимодействие модулей $w_{y^*}^{K_1}$ и $w_y^{K_2}$ из кластеров K_1^W и K_2^W контролируемое средствами защиты физического $d_f^{K_F^D}$ и синтаксического $d_l^{K_L^D}$ уровней, например, средством физической защиты периметра и средством шифрования. Пример г) демонстрирует защиту взаимодействия модулей w_y и w_{y^*} из одного кластера K_2^W через средство, которое реализует конструктивную защиту в узкой трактовке термина на физическом $d_1^{K_1^F}$ и синтаксическом уровнях $d_1^{K_1^L}$, например, средство квантового распределения ключей.

Выводы

Предложен подход к формированию формальных критериев и показателей конструктивной защиты архитектуры информационных систем, которые сформированы на основе кластерной модели защиты в терминах теории МК-сетей. Критерии конструктивной защиты позволяют проводить формальный анализ FLS-мультиграфа, который является моделью информационной архитектуры, с учетом топологии информационного взаимодействия и политики доступа, оценивать непрерывность уровня защиты для траекторий информационного процесса сравнением нормированных весов средств защиты информации с установленным уровнем защиты.

Оценка конструктивной защиты архитектуры информационной системы обеспечивается формальным анализом наличия или отсутствия вершин и дуг мультиграфа. Критерии конструктивной защиты инвариантны к широкому классу задач оценки безопасности архитектуры и вариантов топологии информационного взаимодействия, могут быть использованы при разработке стандартов конструктивной безопасности информационных систем, создании систем с «иммунной защитой». Результаты исследований позволяют разработать систему критериев конструктивной защиты для решения проблемы оценки безопасности, автоматизации анализа и синтеза DLTS с доказательством уровня безопасности с применением аппарата математической логики.

Результаты получены при финансовой поддержке проекта «Технологии противодействия ранее неизвестным квантовым киберугрозам», реализуемого в рамках государственной программы федеральной территории «Сириус» «Научно-технологическое развитие федеральной территории «Сириус» (Соглашение № 23-03 от 27.09.2024 г.)

Литература

1. Antipov I. S., Arustamyan S. S., Ganichev A. A. Markov A. S. et al. Intelligent Fuzzing Method for Aviation Information Systems as Part of the Secure Software Development Cycle. Russian engineering research. 45, 685–690 (2025). DOI: 10.3103/S1068798X25700728.
2. Ищукова Е. А. О влиянии криптографической стойкости функций хеширования на устойчивость современных блокчейн-экосистем и платформ // Вопросы кибербезопасности. 2025. № 3(67). С. 63-71. DOI: 10.21681/2311-3456-2025-3-63-71.
3. Markov A. S., Varenitca V. V., Arustamyan S. S. Topical issues in the implementation of secure software development processes. В сборнике: Proceedings of the International Conference on Information Processes and Systems Development and Quality Assurance. IPSQDA-2023. 2023. С. 48–53.
4. Наседкин П. Н. Оценка состояния комплексной системы защиты информации на основе онтологий / П. Н. Наседкин, Л. В. Аршинский // Информационные и математические технологии в науке и управлении. 2023. № 1(29). С. 158–177. – DOI:10.38028/ESI.2023.29.1.014.
5. Балябин А. А., Петренко С. А. Модель блокчейн-платформы с кибериммунитетом в условиях квантовых атак // Вопросы кибербезопасности. 2025. № 3(67). С. 72–82. DOI: 10.21681/2311-3456-2025-3-72-82.
6. Petrenko A. S., Petrenko S. A. Basic Algorithms Quantum Cryptanalysis // Вопросы кибербезопасности. 2023. No. 1(53). P. 100–115. DOI: 10.21681/2311-3456-2023-1-100-115.
7. Petrenko A. S. Applied Quantum Cryptanalysis (scientific monograph). River Publishers. (2023). 256 p. ISBN 9788770227933. DOI: 10.1201/9781003392873.
8. Mark Webber, Vincent Elfving, Sebastian Weidt, Winfried K. Hensinger. The impact of hardware specifications on reaching quantum advantage in the fault tolerant regime. AVS Quantum Sci. 4, 013801 (2022). DOI: 10.1116/5.0073075.
9. Battarbee C., Kahrobaei D., Perret L., Shahandashti S. F. SPDH-Sign: Towards Efficient, Post-quantum Group-Based Signatures. In: Johansson, T., Smith-Tone, D. (eds) Post-Quantum Cryptography. PQCrypto 2023. (2023). Lecture Notes in Computer Science. V. 14154. P. 113–138. Springer, Cham. DOI: 10.1007/978-3-031-40003-2_5.
10. Li L., Lu X., Wang K. Hash-based signature revisited. (2022). Cybersecurity. V. 5. Article no. 13. DOI:10.1186/s42400-022-00117-w.
11. Сундеев П. В. Функциональная стабильность распределенного реестра в условиях появления новой квантовой угрозы // Вопросы кибербезопасности. 2025. № 3(67). С. 83–89. DOI: 10.21681/2311-3456-2025-3-83-89.
12. Tristan Muller, Thomas Alexander, Michael E. Beverland, Markus Buhler, Blake R. Johnson, Thilo Maurer, Drew Vandeth. Improved belief propagation is sufficient for real-time decoding of quantum memory. IBM Quantum. (Dated: June 11, 2025). DOI: 10.48550/arXiv.2506.01779.
13. Theodore J. Yoder1, Eddie Schoute1, Patrick Rall1, Emily Pritchett1, Jay M. Gambetta1, Andrew W. Cross1, Malcolm Carroll1, and Michael E. Beverland. Tour de gross: A modular quantum computer based on bivariate bicycle codes. IBM Quantum. (Dated: June 03, 2025). DOI: 10.48550/arXiv.2506.03094.
14. Skiba V. Yu., Petrenko S. A., Gnidko K. O., Petrenko A. S. Concept of ensuring the resilience of operation of national digital platforms and blockchain ecosystems under the new quantum threat to security. Computing, Telecommunication and Control, 2025, Vol. 18, No. 2, pp. 56–73.

CRITERIA AND INDICATORS FOR CONSTRUCTIVE PROTECTION OF THE DISTRIBUTED REGISTRY

Sundeev P. V.[4]

Keywords: *modular cluster network, quantum threat, information protection.*

The purpose of the study: *to investigate the problem of formal criteria and indicators for evaluating the constructive protection of information systems, taking into account the features of the architecture of a distributed registry in the context of a quantum threat, to propose an approach to the formation of criteria for constructive protection that automate the analysis and synthesis of secure architecture.*

Research methods: *object-oriented analysis of complex systems, system analysis, theory of modular cluster networks, graph theory, matrix theory, mathematical logic.*

Research result: *an approach to the formation of a system of criteria and indicators for constructive protection of an information system architecture based on a cluster protection model with complete overlap, presented as a cluster multigraph of a modular cluster network, and an analysis of its topology is proposed. The assessment of architecture security according to the criteria of constructive protection consists in a formal analysis of the presence or absence of certain types of vertices and arcs of a cluster multigraph. The continuity indicator of constructive protection for assessing the safety of information process trajectories is based on comparing the weights of the marked vertices indicating the protection modules with the set value. The applicability of constructive protection criteria to the evaluation of distributed registry technology is shown.*

Scientific novelty: *a new approach to the formation of formal criteria and indicators for constructive protection of information systems based on a cluster protection model with complete overlap in terms of the theory of modular cluster networks has been developed.*

The results were obtained *with the financial support of the project «Technologies for countering previously unknown quantum cyber threats», implemented within the framework of the state program of the «Sirius» Federal Territory «Scientific and technological development of the «Sirius» Federal Territory (Agreement No. 23-03 dated September 27, 2024).*

References

1. Antipov, I. S., Arustamyan, S. S., Ganichev, A. A., Markov, A. S. et al. Intelligent Fuzzing Method for Aviation Information Systems as Part of the Secure Software Development Cycle. (2025). Russian engineering research. No 45. P. 685–690. DOI: 10.3103/S1068798X25700728.
2. Ishchukova, E. A. On the influence of cryptographic stability of hashing functions on the stability of modern blockchain ecosystems and platforms. (2025). Voprosy Kiberbezopasnosti [Cybersecurity issue]. No 3 (67). P. 63-71. DOI: 10.21681/2311-3456-2025-3-63-71 (Russian Text).
3. Markov, A. S., Varenitca, V. V., Arustamyan, S. S. Topical issues in the implementation of secure software development processes. (2023). In the collection: Proceedings of the International Conference on Information Processes and Systems Development and Quality Assurance. IPSQDA-2023. P. 48–53.
4. Nasedkin, P. N. Assessment of the state of the integrated information security system based on ontologies. (2023). Informacionnye i matematicheskie tekhnologii v nauke i upravlenii. [Information and mathematical technologies in science and management]. No 1(29). P. 158–177. DOI:10.38028/ESI.2023.29.1.014 (Russian Text).
5. Balyabin, A. A., Petrenko, S. A. Model of a blockchain platform with cyber-immunity under quantum attacks. (2025). Voprosy Kiberbezopasnosti [Cybersecurity issue]. No 3(67). P. 72–82. DOI: 10.21681/2311-3456-2025-3-72-82 (Russian Text).
6. Petrenko, A. S., Petrenko, S. A. Basic Algorithms Quantum Cryptanalysis. (2023). Voprosy Kiberbezopasnosti [Cybersecurity issue]. No. 1(53), pp. 100–115. DOI: 10.21681/2311-3456-2023-1-100-115 (Russian Text).
7. Petrenko, A. S. Applied Quantum Cryptanalysis (scientific monograph). (2023). River Publishers. 256 p. ISBN 9788770227933. DOI: 10.1201/9781003392873.
8. Webber, M., Elfving, V., Weidt, S., Hensinger, W. The impact of hardware specifications on reaching quantum advantage in the fault tolerant regime. (2022). AVS Quantum Sci. 4, 013801. DOI: 10.1116/5.0073075.
9. Battarbee, C., Kahrobaei, D., Perret, L., Shahandashti, S. F. SPDH-Sign: Towards Efficient, Post-quantum Group-Based Signatures. In: Johansson, T., Smith-Tone, D. (eds). (2023). Post-Quantum Cryptography. PQCrypto 2023. Lecture Notes in Computer Science. V. 14154. P. 113–138. Springer, Cham. DOI: 10.1007/978-3-031-40003-2_5.
10. Li, L., Lu, X., Wang, K. Hash-based signature revisited. (2022). Cybersecurity. V. 5. Article No. 13. DOI:10.1186/s42400-022-00117-w.
11. Sundeev, P. V. Functional stability of a distributed registry in the context of the emergence of a new quantum threat. (2025). Voprosy Kiberbezopasnosti [Cybersecurity issue]. No 3 (67). P. 83–89. DOI: 10.21681/2311-3456-2025-3-83-89 (Russian Text).
12. Muller, T., Alexander, T., Beverland, M., Buhler, M., Johnson, B., Maurer, T., Vandeth, D. Improved belief propagation is sufficient for real-time decoding of quantum memory. (2025). IBM Quantum. DOI: 10.48550/arXiv.2506.01779.
13. Yoder, T., Schoute, E., Rall, P., Pritchett, E., Gambetta, J., Cross, A., Carroll, M., Beverland, M. (2025). Tour de gross: A modular quantum computer based on bivariate bicycle codes. IBM Quantum. DOI: 10.48550/arXiv.2506.03094.
14. Skiba, V. Y., Petrenko, S. A., Gnidko, K. O., Petrenko, A. S. Concept of ensuring the resilience of operation of national digital platforms and blockchain ecosystems under the new quantum threat to security. (2025). Computing, Telecommunication and Control. Vol. 18, No. 2, pp. 56–73.

4 Pavel V. Sundeev, Dr. Sc. (Technical), Chief researcher, Scientific Center for Information Technologies and Artificial Intelligence, Sirius University of Science and Technology, Address: Olimpiyskiy ave. b.1, Sirius, Sirius Federal Territory, Krasnodar region, Russia, 354340, ORCID: 0009-0005-8442-8921. E-mail: Sundeev.pv@talantiuspeh.ru

ПОДХОД К АНАЛИЗУ И ОЦЕНКЕ ЗАЩИЩЕННОСТИ СИСТЕМ УПРАВЛЕНИЯ БОЛЬШИМИ ДАННЫМИ

Полтавцева М. А.[1], Зегжда Д. П.[2]

DOI: 10.21681/2311-3456-2025-5-109-118

Цель исследования: *разработка подхода к анализу и оценке защищенности систем управления большими данными, с учетом технологических особенностей данного класса решений, отличающих их от традиционных облачных систем обработки данных.*

Метод(ы) исследования: *в работе проводится анализ особенностей целевых систем, а также предлагаемых исследователями методов сбора данных и оценки защищенности. Выделяются их недостатки в контексте современных требований. Предлагается использовать подход к сбору данных и моделированию целевой системы с использованием теоретико-множественной агрегатной модели данных, а также интеграция модифицированной оценки NIST контроля доступа в системах больших данных и авторской оценки на основе учета грануляции данных и доверия к узлам-обработчикам.*

Результат(ы) исследования: *в результате работы были сформулированы технологические особенности целевых систем с точки зрения оценки защищенности, такие как распределенность, гетерогенность (мультимодельность), сложный жизненный цикл данных. Анализ научных работ показал, с одной стороны, интерес исследователей к задаче оценки защищенности систем управления большими данными, а с другой – отсутствие оценок, предложенных для целевого класса систем. Авторами сформированы требования к оценке защищенности к системам управления большими данными как к специфическому компоненту современных информационных систем. Также предложен новый метод оценки защищенности, впервые учитывающий специфические свойства систем управления большими данными. Предлагаемый метод, дополнительно к ранее предложенным оценкам, учитывает недостатки контроля доступа вызванные различной грануляцией данных в компонентах целевой системы, а также большое число доверенных пользователей, и, как следствие, необходимость обработки конфиденциальных данных либо на доверенных узлах, либо в скрытом (обфусцированном или зашифрованном) виде. Предложенная оценка является нормированной, может быть детализирована до оценки каждого конкретного инструмента обработки данных, легко расширена или встроена в более высокоуровневые оценки. Показана достоверность и возможность практического применения предложенной оценки путем разработки программного прототипа на основе ранее известных и апробированных программных решений.*

Научная новизна: *новизна заключается в авторском методе оценки защищенности систем управления большими данными, отличающимся впервые предложенным учетом недостатков контроля доступа вызванных различной грануляцией данных и учетом доверия к отдельным узлам обработчикам данных.*

Ключевые слова: *большие данные, информационная безопасность, оценка защищенности, грануляция данных, контроль доступа, доверие.*

Введение

Сегодня в результате развития цифровых технологий практически во всех отраслях широко используется сбор, обработка, хранение и использование больших данных. Технологически, формируется новый класс информационных систем, в составе которых вместо классического сервера систем управления базами данных (СУБД), для хранения и обработки информации используется сочетание разнородных компонентов: набор СУБД различного типа, инструменты потоковой обработки информации, иное программное обеспечение различного типа.

Такие системы сталкиваются с большим количеством угроз и проблем безопасности, связанных с особенностями их построения [1]. В свою очередь недостатками систем защиты таких решений, из-за их технологических особенностей, являются децентрализованные и несогласованные механизмы контроля доступа отдельных инструментов хранения и обработки данных, высокое доверие между инструментами, сложность согласованного администрирования.

Фундаментальными причинами проблемы безопасности в рассматриваемом классе систем является, в первую очередь различная грануляция данных в инструментах их обработки [2]. Эта особенность, в свою очередь, порождает различия в используемых моделях контроля доступа и их реализациях в разных компонентах системы, что, в свою очередь, зачастую является причиной несанкционированного доступа к данным.

1 Полтавцева Мария Анатольевна, доктор технических наук, доцент, Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский политехнический университет Петра Великого». Санкт-Петербург, Россия. E-mail: poltavtseva@ibks.spbstu.ru

2 Зегжда Дмитрий Петрович, член корреспондент РАН, доктор технических наук., профессор, Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский политехнический университет Петра Великого». Санкт-Петербург, Россия. E-mail: zegzhda_dp@spbstu.ru

Целью данной работы является создание подхода к анализу и последующей оценке защищенности систем управления большими данными с учетом ключевых особенностей, отличающих данный класс решений от других типов информационных систем.

Степень разработанности проблемы

Проблемой безопасности в системах обработки и хранения больших данных занимаются многие исследователи. Нужно отметить, что на данный момент большая часть работ посвящена контролю доступа в таких системах. Также исследователи отмечают задачи сбора данных и оценки защищенности, однако эти области являются в значительной степени менее исследованными [3].

Наибольшая часть современных научных работ посвящена мониторингу и анализу управления большими данными в облачных системах. В частности, можно отметить исследования посвященные безопасности облачных систем и процессов разработки [4]. Однако распределенные гетерогенные системы больших данных исследователи в этой области не рассматривают, ограничиваясь рамками одного центра обработки данных и одной экосистемы инструментов обработки. Ряд работ посвящены использованию blockchain и технологиям распределенных реестров. Это работы как по управлению большими данными при помощи данной технологии [5], так и по мониторингу отдельных технологических сред, в частности – интернета вещей [6], промышленных решений [7]. Недостатком этих работ также является их ориентированность на гомогенную среду, в крайнем случае – экосистему и решения одного производителя. Предложенное в [8] решение пока не имеет известных аналогов для гетерогенных систем.

С точки зрения оценки защищенности сегодня работ также крайне мало, в том числе потому, что такая оценка должна базироваться на данных, полученных в результате анализа системы. А задача сбора данных для такого анализа, типового решения, как показано выше, еще не имеет. Как правило, исследователи оценивают безопасность на уровне информационных систем в целом. Для больших данных предлагается оценка безопасности на основе теории принятия решений [9], различные подходы по оценке рисков (например, [10]). Проблема безопасности и оценки защищенности гетерогенных архитектур больших данных является фокусом исследования в работе [11], однако конкретного решения ее авторы не предлагают. Необходимость оценки защищенности через анализ контроля доступа авторы рассматривают и обосновывают в работе [12], однако сама оценка не приводится. А оценка из работы [13] также не учитывает иных свойств данного класса систем, кроме необходимости согласования контроля доступа.

С точки зрения оценки защищенности гетерогенные системы обработки и хранения больших данных являются частью информационных систем, и оценки защищенности информационных систем релевантны в их отношении [14]. В свою очередь, исследования в области оценки защищенности информационных систем можно отнести к следующим категориям. Во-первых, это работы, которые рассматривают отдельные аспекты защищенности систем и технологий (например, [15]). Во-вторых, работы, посвященные конкретным инфраструктурам [16], включая оценку рисков безопасности и облачные инфраструктуры [17]. Здесь стоит отдельно отметить работу [18]. В ней авторы анализируют гетерогенные архитектуры, однако специфические проблемы гетерогенных систем обработки и хранения данных ими также не рассматриваются.

Таким образом, задача оценки защищенности гетерогенных систем управления большими данными, включая как сбор и анализ информации, так и вычисление самой метрики защищенности, является высоко актуальной.

Особенности гетерогенных систем управления большими данными как объекта защиты

Гетерогенные системы обработки и хранения больших данных обладают рядом ключевых особенностей [19]. Эти особенности являются их нативными технологическими чертами систем управления большими данными и отличают их как от классических систем управления базами данных, так и от более общих классов решений (например, информационных систем). Ключевыми технологическими чертами нового класса, гетерогенных систем управления большими данными, являются:

1. Распределенная в широком смысле среда обработки.
2. Использование нескольких разнородных (основанных на разных моделях данных) инструментов обработки данных.
3. Единый жизненный цикл разнородных данных.

В рамках рассматриваемых систем данные обрабатываются, как правило, в рамках нескольких связанных СУБД функционирующих на основе различных моделей данных. При этом, в силу разнородности и сложности задач, каждая такая система является самостоятельным компонентом со своими пользователями и жизненным циклом данных [1]. Как правило, все или часть таких СУБД развернуты отдельно друг от друга на географически распределенных серверах или центрах обработки данных. В результате в организации используется распределенная в широком смысле среда обработки данных, вся информация в которой связана в единый жизненный цикл. Результаты обработки или сырые данные

Рис. 1. Гетерогенные системы обработки и хранения больших данных

с одного этапа или из одной системы являются входными данными для другой и служат целевой информацией для ее пользователей. В результате одни и те же данные обрабатываются на различных узлах под управлением различных СУБД в распределенной корпоративной среде (рисунок 1).

С точки зрения решения задач безопасности это приводит к необходимости учета следующих моментов при оценке защищенности:

1. Каждый отдельный инструмент хранения и обработки данных нуждается в собственной оценке и анализе защищенности [20].
2. Политики безопасности в рамках отдельных инструментов должны быть согласованы, отсутствие такого согласования или его несоответствия, вызванные технологической невозможностью должной грануляции данных, должны учитываться как потенциальные уязвимости [21].
3. Доверие в отношении узлов – обработчиков данных должно оцениваться и обработка открытых данных на узлах с низким уровнем доверия (большим числом доверенных привилегированных пользователей или иными факторами риска) должна быть минимизирована [22].
4. Необходима возможность получения интегрированной оценки защищенности для информационной системы в целом.

В этих условиях основной задачей становится сбор данных и оценка защищенности, которая, с одной стороны, учитывает безопасность отдельного инструмента, с другой – технологические особенности целевого класса систем (систем управления большими данными), и с третьей - может быть интегрирована в оценку защищенности информационной системы организации в целом. В оценке защищенности таких систем требуется учесть несколько факторов.

Во-первых, это недостатки системы контроля доступа. Как показано в работах [2, 12], при использовании инструментов с различной грануляцией данных неизбежно возникают проблемы согласования доступа. Также характерная особенность таких систем — это повышение привилегий отдельных пользователей из-за недостаточной грануляции слабо структурированных данных для обеспечения бизнес-процессов.

Во-вторых, это динамичность компонентов обработки и хранения данных. Указанные организации как правило имеют большой объем legacy, в том числе в области организации бизнес-процессов, что вызывает дублирование данных. Организационные изменения и работа с персоналом также приводят к возникновению новых этапов и шагов в обработке информации. Отслеживание и учет таких этапов это также важная задача.

В любом случае, оценка защищенности (в том числе для систем обработки и хранения больших данных) базируется на результатах сбора и анализа данных в процессе мониторинга и определении на их основе численных метрик.

Оценка защищенности гетерогенных систем больших данных

Сбор и анализ данных для оценки защищенности

Сбор и анализ информации в гетерогенных системах обработки и хранения больших данных для решения задачи оценки защищенности должен обеспечивать информированность о движении потоков данных (по крайней мере, между инструментами обработки) и сведениями о состоянии и согласованности контроля доступа в инструментах обработки и хранения информации.

Для решения этой задачи была использована система мониторинга и моделирования процессов обработки данных в системах управления большими данными на основе распределенного реестра [12]. Моделирование процесса обработки данных, на основе реальных потоков данных в целевой системе, позволяет установить изменения в процессах обработки и хранения данных для оценки сведений о движении фрагментов данных. Над полученными данными авторским коллективом ранее был разработан метод анализа политик безопасности, с целью поиска ошибок и уязвимостей. Анализ политики безопасности решает несколько задач:

1. Проверка соответствия частных политик безопасности на инструментах обработки и хранения данных политике верхнего уровня.
2. Обнаружение возможных каналов логического вывода со стороны получателей данных.
3. Итерационный поиск оптимальной политики безопасности, соответствующей принципу минимального доверия, с учетом технологических ограничений.

В результате может быть установлено соответствие реализации политики безопасности и ее оптимальной конфигурации, а также выявлены неустранимые недостатки, которые должны быть компенсированы иными мерами. Общая схема сбора данных и их применения на последующих шагах по расчету оценки защищенности приведена на рисунке 2.

Рис. 2. Схема сбора данных и оценки защищенности в гетерогенных системах обработки и хранения больших данных

На первых этапах выполняется сбор данных на узлах-обработчиках, на уровне отдельных инструментов, и сохранение в цепочке распределенного реестра. Собираемые данные включают информацию об узлах системы управления большими данными, типах фрагментов данных и выполняемых операциях над ними. Технологически сбор данных может быть основан на традиционных технологиях, однако использование распределенного реестра позволяет гарантировать целостность информации об узлах и выполняемых операциях, поступающих для дальнейшей оценки защищенности.

Затем осуществляется анализ полученных цепочек распределенного реестра уже в блоке управления безопасностью, включая извлечение данных о последовательностях операций над каждым фрагментом больших данных. На основе извлеченных данных блоком моделирования выполняется построение модели системы сбора, хранения и обработки больших данных в виде графа обработки информации в целевой гетерогенной системе. Вершинами такого графа являются операции над данными, а ребрами – передаваемые фрагменты данных.

Для согласования логических представлениий инструментов обработки данных (моделей данных) при составлении модели системы обработки информации в целом, на основе аппарата теории множеств была предложена агрегатная модель данных [23].

Пусть $d_i \in D$ – фрагмент данных, передаваемый между инструментами и/или обработки информации. Тогда каждый i-й фрагмент, точнее – тип фрагмента, представляется кортежем ключ – значение $d_i = <Key, Val>$, где ключом является уникальный идентификатор типа фрагмента данных. Значение фрагмента данных – содержащаяся в нем информация, включая, в некоторых случаях, пустое множество: $d_i.Value = \{Val, \{d_{ij} | d_i \in \{\emptyset, d\}\}$. Каждый элемент данных может быть атомарным, и тогда ключ является произвольным UID внутри системы. Либо же элемент данных может представлять собой агрегацию других элементов (записи в кортеже, кортежи в таблице, документы в коллекции и т.д.). Идентификатор (ключ) части составного фрагмента данных представляет собой конкатенацию собственного UID и UID обещающего фрагмента: $d_i.Key = \overline{Key} = (k_n, \ldots, k_1)$.

В рамках модели выделяются основные операции над фрагментами данных, такие как:

Создание фрагмента:

$$Create: (\{d\}, Key) \longrightarrow d_i = \{Key, Value\}.$$

Удаление фрагмента: $Delete: (d_i) \longrightarrow \emptyset$.

Включение одного фрагмента в другой:

$$Incl{:}(d_i,\{d\}) \longrightarrow d_i`.$$

Исключение части составного фрагмента:

$$Exception\ Extr{:}(d_i`,\{Key_j\}) \longrightarrow (d_i,d_j).$$

Изменение фрагмента данных, в общем случае структурное (например, сортировка) или семантическое (например, вычисление над ним некоторой функции или иное преобразование): $Transform{:}(d_i) \longrightarrow d_j$.

С точки зрения безопасности и оценки защищенности все приведенные операции разделяются на те, которые могут выполняться без доступа к семантике данных на текущей технологической платформе и те, которые требуют такого доступа.

Без доступа к семантике, как правило, выполняются операции удаления, включения и, часто, исключения фрагментов данных. Это обусловлено тем, что при удалении чтение содержимого фрагмента не требуется, при включении фрагментов также не нужен доступ к значению данных, а только к синтетическому ключу-идентификатору, а при исключении фрагментов большинство систем также позволяют извлечь часть не оперируя к целому. Здесь безусловно есть некоторые нюансы, связанные, в частности, с механизмами ускорения доступа в системах управления базами данных. Например, для выполнения запроса требуется сравнение критерия запроса со значениями индекса, и таким образом, происходит обращение к семантике. Однако на сегодня известны технологии сквозного шифрования для устранения этой проблемы и для NoSQL решений на базе более простых моделей задача может решаться проще. Поэтому операция исключения в каждом отдельном случае может как требовать доступа к открытой семантике данных, так и быть реализована без нее.

Операции создания и трансформации фрагментов данных очевидным образом требуют доступа к семантике сведений и должны выполняться на доверенных узлах или в защищенном режиме, т.е. над зашифрованными или обфусцированными данными.

Метод оценки защищенности

Определение оценки защищенности (нижний блок на рисунке 2) можно разделить на два этапа, которые могут выполняться параллельно. Первый этап заключается в анализе реализованных политик безопасности в гетерогенной системе сбора и обработки больших данных. На этом этапе для гетерогенных систем обработки и хранения больших данных целесообразно применение двух типов анализа. Во-первых, анализ и оценка защищенности на основе подхода NIST, детально описанный в [13, 24]. Во-вторых, детальный анализ политик безопасности гетерогенных мульти модельных систем, с учетом технологических ограничений и разницы в грануляции данных [12].

Анализ согласно рекомендациям NIST проводится на основе цепочки доверия, в которой выделяются источник данных, ведущая и ведомая системы [24]. Цепочка доверия формируется от источника данных к ведущей системе через соглашение о безопасности, а от ведущей системы к ведомой через список доверенных систем, политику контроля доступа и словарь атрибутов. Обратная связь от ведомой системы к ведущей формируется через локальную политику ведомой системы и словарь атрибутов, тем самым формируя невозможность неавторизованного доступа без участия обеих систем. Ситуация, когда реальная реализация контроля доступа шире, чем должна быть на основе модели, называется нарушением. При нарушении оценка безопасности рассматриваемой системы уменьшается, тем больше, чем выше ее класс секретности. Порядок расчета такой оценки Ecn приведен на рисунке 3, где n и n’– число правд доступа реализованных с нарушениями текущее и итоговое.

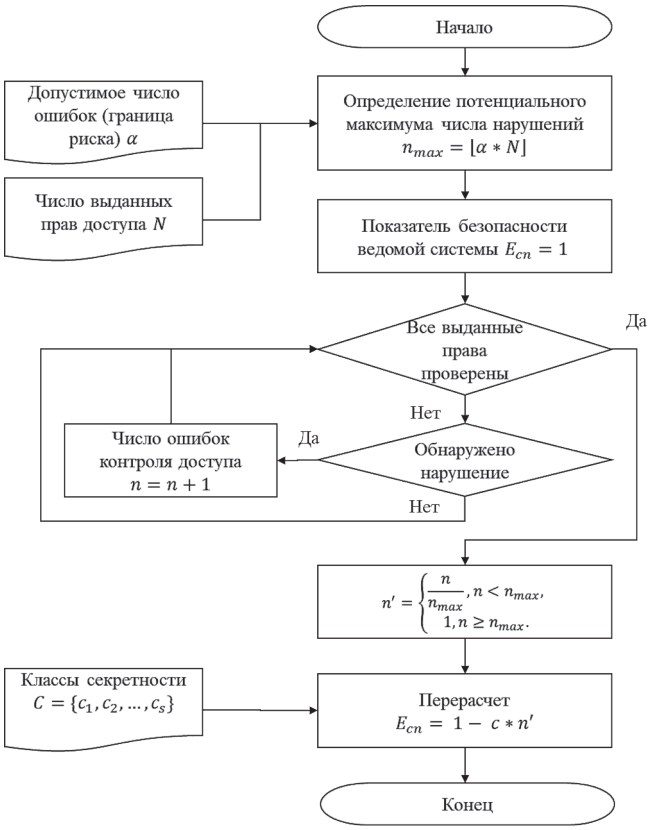


Рис. 3. Алгоритм расчета оценки контроля доступа

При детальном анализе политики безопасности на основе полученной в результате сбора данных модели потоков данных конкретной системы последовательно проводится две оценки. Анализ соответствия политике безопасности верхнего уровня заключается в оценке соответствия общей политике безопасности

политик безопасности реализованных на узлах обработки данных. В результате него формируется оценка $Ec_{,i}$. В наиболее простом случае эта оценка может быть бинарной для узла $Ec_{,i} \in \{0,1\}$, где 0 – не выполняется, 1 – выполняется. Наиболее рациональной представляется следующая оценка:

$$Ec_i = \frac{n_i}{N} \quad (1)$$

где n_i – число отклонений от правил политики безопасности на узле i, а N – общее число правил политики. Полученное значение $Ec_{,i} \in [0,1]$. При этом не учитывается объем отклонений, корректная оценка которого сегодня представляется сложной задачей. Оценка Ec_i является более детальной относительно оценки по NIST (Ec_n).

На втором этапе проводится анализ узлов, выполняющих операции над данными, с учетом степени их доверенности. Цель этой части оценки – выявление операций над открытыми данными на узлах обработки и хранения больших данных, не обладающих должным уровнем доверия. При ее проведении фактически выполняется правило «*все операции с открытыми данными должны выполняться только на доверенных узлах системы*».

Для проведения такого анализа используется отображение графа обработки данных на узлы обработки в модели системы. Оценка самого доверия к узлам при этом может проводится стандартными общеизвестными способами, уровень доверия владелец системы может установить самостоятельно на основ критичности или степени конфиденциальности обрабатываемых данных. Входными данными для дальнейшего расчета является статус узлов: условно «доверенный» или «не доверенный» с точки зрения возможности обработки открытых данных. При анализе узлов системы, выполняющих операции над данными, с учетом степени их доверенности, режим обработки данных считается безопасным в том случае, если обработка данных осуществляется в зашифрованном виде на уровне пользователя либо же данные обфусцированы, то есть, недоступны для злоумышленника, имеющего доступ к узлу – обработчику в том числе, на уровне администратора узла.

Оценка защищенности с учетом доверия в отношении операций по обработке данных рассчитывается как нормированная оценка безопасности операций для каждого i-го узла системы:

$$Eo_i = \frac{\sum_{j=1}^{M} s_j}{M}, \quad (2)$$

где M – число различных типов операций, s_j – признак безопасности j-й операции. Для обработки в безопасном режиме s_i = 1, для обработки в небезопасном режиме s_j = 0. Таким образом, оценка $Eo_i \in [0,1]$ показывает соотношение безопасных и небезопасных операций. На безопасном узле Eo_i = 1. Если Eo_i = 0 – это значит все операции выполняются в небезопасном режиме.

Для системы в целом может быть вычислена суммарная нормированная оценка на основе формул (1) и (2) вида:

$$\mathrm{Es} = \frac{\sum_{i=1}^{N}(Ec_i + Eo_i)}{2N}. \quad (3)$$

Также одним из требований к оценке защищенности была необходимость ее интеграции с другими оценками, характеризующими информационную систему в целом. Так как обе разработанные оценки Ec и Eo являются нормированными, прочие характеристики защищенности при необходимости вводятся дополнительно владельцем системы также как нормированная метрика $E_{ex} \in [0,1]$.

Итоговая оценка защищенности может рассчитывается согласно формуле:

$$E = \frac{E_{ex} + Es}{3}, \quad (4)$$

где E_{ex} – внешняя оценка, E_s – оценка защищенности на основе анализа контроля доступа и безопасности операций, отражающая специфические особенности и уязвимости системы. Полученная оценка является нормализованной, при необходимости детализируется и может входить в более высокоуровневые оценки, включая, при необходимости, показатель состояния технической защиты информации[3]. В формулы (3) и (4) также могут быть добавлены веса для акцентирования различных аспектов функционирования системы управления большими данными.

Стоит отметить, что предложенная оценка защищенности распределенных гетерогенных систем управления большими данными включает не только оценки, полученные известными ранее способами, но и составляющую, определяемую соответствием процессов сбора, хранения и обработки больших данных политике безопасности с учетом различий в структуризации данных на различных узлах целевой системы. Предложенный метод расчета оценки защищенности позволяет как детализировать ее до отдельных узлов, определяя наиболее уязвимые компоненты системы, так и получить общее усредненное значение для сравнения различных конфигураций системы, ее эволюционных версий и иных задач.

Реализация предложенных решений

Для реализации приведенного метода в рамках технологии обеспечения защищенности систем обработки и хранения больших данных была разработана

3 Методика оценки показателя состояния технической защиты информации и обеспечения безопасности значимых объектов критической информационной инфраструктуры Российской Федерации. Методический документ от 2 мая 2024 г. Утвержден ФСТЭК 2 мая 2024 г.

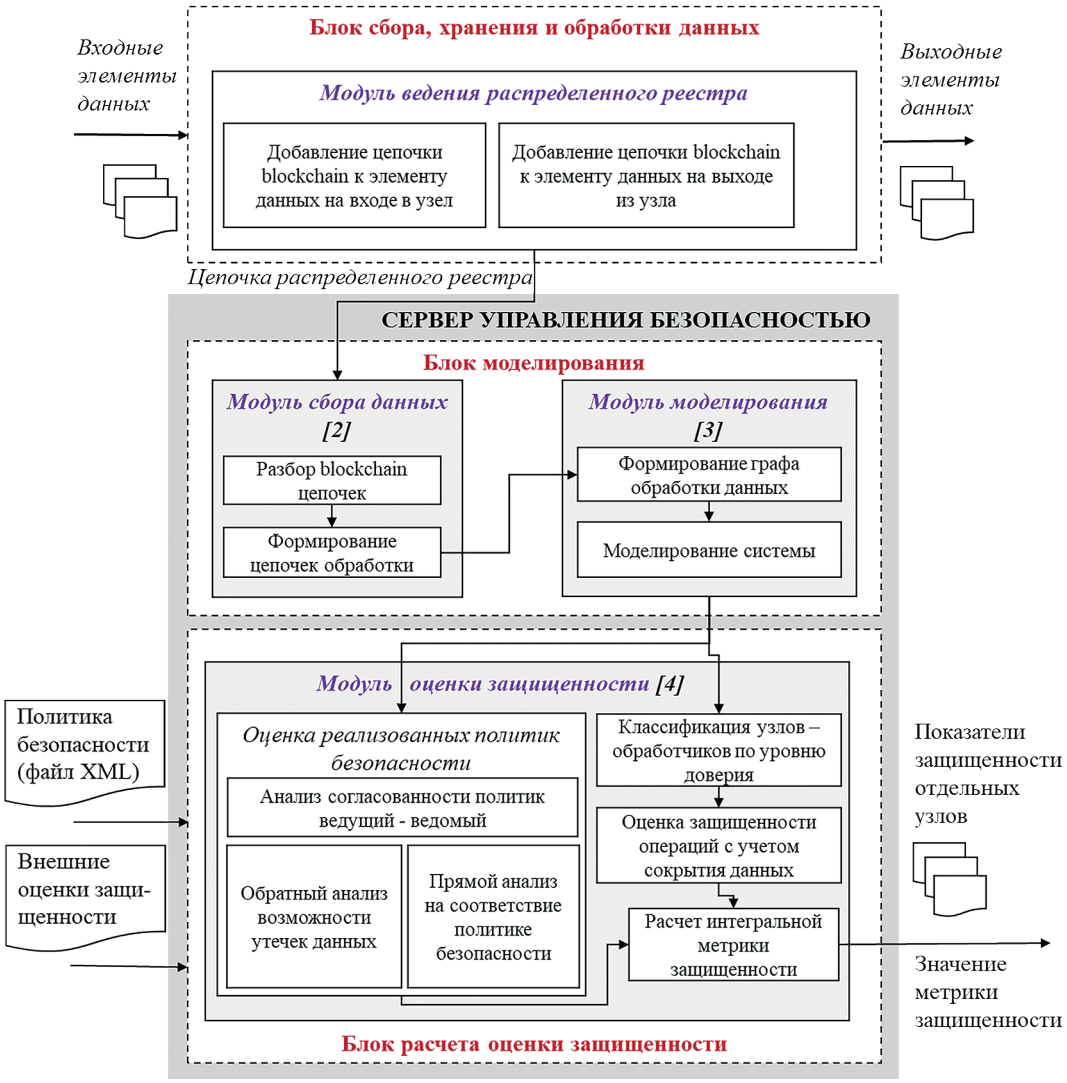


Рис. 4. Система оценки защищенности систем управления большими данными

архитектура системы безопасности и программный прототип для оценки защищенности. Программный прототип реализован на языке python с применением технологии blockchain на этапе сбора и моделирования данных[4].

На узлах – обработчиках данных, на которых функционируют отдельные инструменты гетерогенной обработки информации (PostgreSQL, CassandraDB, MongoDB) и реализуется сложный жизненный цикл обработки данных реализуются программные агенты, собирающие сведения для построения модели системы на основе потоков данных. Центральный сервер управления безопасностью осуществляет сбор и моделирование данных, анализ политики безопасности и расчет оценок (рисунок 4).

Управление указанными фреймворками осуществляется через консоль, на входе принимается json – файлы конфигурации, включая политику безопасности верхнего уровня и внешние оценки безопасности. Влияние на производительность целевой системы со стороны фреймворка сбора данных показывает увеличение нагрузки на систему в пределах 4–10 % [8], вычисление оценки защищенности и работа с политикой безопасности выполняется сервером управления безопасностью не влияет на производительность целевой системы.

4 Свидетельство о государственной регистрации программы для ЭВМ № 2024688201 Российская Федерация. Программа гранулированного аудита в гетерогенных распределенных системах обработки и хранения больших данных: № 2024686659: заявл. 08.11.2024: опубл. 26.11.2024 / М. А. Полтавцева, М. О. Калинин; заявитель Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский политехнический университет Петра Великого». – EDN NEJLCQ
Свидетельство о государственной регистрации программы для ЭВМ № 2024687565 Российская Федерация. Программа автоматического анализа политик безопасности и контроля доступа в системах обработки и хранения больших данных: № 2024686731: заявл. 08.11.2024: опубл. 20.11.2024 / М. А. Полтавцева, М. О. Калинин; заявитель федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский политехнический университет Петра Великого». – EDN CZTMNY

Заключение

При решении поставленной задачи были сформулированы особенности систем управления большими данными, которые отличают их от других программных решений: гетерогенность, географическая распределенность и единый жизненный цикл данных в рамках экосистемы в целом. Эти особенности являются отличительными чертами целевых систем, прямо влияющими на их безопасность и должны быть приняты во внимание при формировании оценки защищенности.

В основе сбора и подготовки данных при оценке защищенности, для отражения особенностей гетерогенных мультимодельных систем управления большими данными, предлагается использовать ранее разработанную систему на основе технологий распределенного реестра. Построенная в результате сбора информации о движении фрагментов данных в целевой системе модель позволяет оценить выполнение политики безопасности не только в формате ведущей - ведомой систем хранения, но и учесть различную грануляцию данных на протяжении их жизненного цикла и связанные с этим возможные нарушения политики безопасности системы.

При расчете оценки защищенности в части, специфической для целевых систем управления большими данными, предложенный новый метод учитывает несколько факторов. Это согласование политик безопасности отдельных инструментов на основе оценки грануляции и подхода NIST и результат анализа узлов, выполняющих операции над данными с точки зрения доверия. В итоге формируется нормированная интегральная оценка, включающая все приведенные аспекты. За счет нормализации предложенная оценка может быть легко интегрирована с иными техниками и оценками защищенности, как применяемыми к целевой системе, так и на более высоком уровне.

Сформированная архитектура программного прототипа и модуль оценки защищенности опираются на использование ранее разработанного фреймворка, апробированы на практике и позволяют рассчитать оценку без дополнительной нагрузки на производительность, за исключением аспекта сбора данных при мониторинге и построении модели целевой системы.

Предложенный метод и система оценки защищенности систем управления большими данными позволяют достичь повышения безопасности целевых систем, в том числе, за счет естественной детализации компонентов оценки до отдельного инструмента обработки данных. Повышение гарантированности выполнения требований безопасности достигается за счет использования распределенного реестра при сборе данных об узлах и операциях над данными в системе сбора, хранения и обработки больших данных, не позволяющего осуществить подмену собранных данных. Автоматизация работы по сбору данных о системе и о процессах обработки данных, а также автоматизация анализа политик безопасности и расчета оценки защищенности позволяют сократить временные и ресурсные затраты на оценку защищенности.

Как дальнейшие направлениями развития данной области можно обозначить повышение качества мониторинга и сбора данных об информационных потоках в системе управления большими данными и повышение степени интеграции предложенных оценок с категориями конфиденциальности данных, принятыми в конкретной целевой организации.

Исследование выполнено за счет гранта Российского научного фонда № 23-11-20003.
https://rscf.ru/project/23-11-20003/, грант Санкт-Петербургского научного фонда
(Соглашение №23-11-20003 о предоставлении регионального гранта).

Литература

1. Минзов А. С., Невский А. Ю., Баронов О. Р. Безопасность персональных данных: новый взгляд на старую проблему // Вопросы кибербезопасности. 2022. №. 4 (50). С. 2–12. DOI:10.21681/2311-3456-2022-4-2-12
2. Colombo P., Ferrari E. Access control technologies for Big Data management systems: literature review and future trends // Cybersecurity. 2019. Т. 2. №. 1. С. 1–13.
3. Rafiq F. et al. Privacy Prevention of Big Data Applications: A Systematic Literature Review // SAGE Open. T.12(2). DOI: 10.1177/21582440221096445
4. Markov A. S., Varenitca V. V., Arustamyan S. S. Topical Issues in the Implementation of Secure Software Development Processes // Proceedings of the International Conference on Information Processes and Systems Development and Quality Assurance IPSQDA-2023 (March 22–24, 2023, St. Petersburg Russia). IEEE. 2023. С. 48–54.
5. Alhazmi H. E., Eassa F. E., Sandokji S. M. Towards Big Data Security Framework by Leveraging Fragmentation and Blockchain Technology // IEEE Access. 2022. Т. 10. С. 10768–10782. DOI: 10.1109/ACCESS.2022.3144632.
6. Wang T. et al. Edge-based auditing method for data security in resource-constrained Internet of Things // Journal of Systems Architecture. 2021. Т. 114. С.1–10. DOI: 10.1016/j.sysarc.2020.101971.

7. Stodt, J. at al. Security Audit of a Blockchain-Based Industrial Application Platform // Algorithms. 2021. Т. 14(4), 121 с. DOI:10.3390/a14040121
8. Kalinin M., Poltavtseva M., Zegzhda D. Ensuring the Big Data Traceability in Heterogeneous Data Systems // 2023 International Russian Automation Conference (RusAutoCon). Sochi, Russian Federation. 2023. С. 775-780. DOI: 10.1109/RusAutoCon58002.2023.10272905.
9. Attaallah A. et al. Analyzing the Big Data Security Through a Unified Decision-Making Approach // Intelligent Automation & Soft Computing, 2022. Т. 32(2). С. 1071–1088. DOI: 10.32604/iasc.2022.022569.
10. Yang M. Information security risk management model for big data // Advances in Multimedia 2022. Т.1 С. 1-10 DOI: 10.1155/2022/3383251.
11. Theodorakopoulos L., Theodoropoulou A., Stamatiou Y. A State-of-the-Art Review in Big Data Management Engineering: Real-Life Case Studies, Challenges, and Future Research Directions // Eng. 2024. Т. 5(3). С. 1266–1297. DOI: 10.3390/eng5030068.
12. Kalinin M., Poltavtseva M. Big Data Security Evaluation by Bidirectional Analysis of Access Control Policy // 2024 International Russian Smart Industry Conference (SmartIndustryCon). Sochi, Russian Federation. 2024. С. 98–103. DOI: 10.1109/SmartIndustryCon61328.2024.10515459.
13. Poltavtseva, M. A., Zaitseva, V. V., Ivanov, D. V. Assessing the Security of Big Data Systems // Aut. Control Comp. Sci. 2024. Т. 58. С. 1352–1364. DOI: 10.3103/S0146411624701025.
14. Dhillon G., Smith K., Dissanayaka I. Information systems security research agenda: Exploring the gap between research and practice // The Journal of Strategic Information Systems. 2021. Т. 30. № 4. DOI: 10.1016/j.jsis.2021.101693.
15. Костогрызов А. И. Методические положения по вероятностном прогнозированию качества функционирования информационных систем ч. 1–3 / А. И. Костогрызов, А. А. Нистратов, П. Е. Голосов // Вопросы кибербезопасности. 2025. № 2(66). С. 2–19. DOI: 10.21681/2311-3456-2025-2-2-19.
16. Оценка уязвимостей автоматизированных систем с применением теории вероятностей, распределения Стьюдента и нормальных случайных величин / И. В. Атласов, А. О. Ефимов, Е. А. Рогозин, А. С. Черкасова // Вопросы кибербезопасности. 2025. № 2(66). С. 124–131. DOI: 10.21681/2311-3456-2025-2-124-131.
17. Ali, T., Al-Khalidi, M., Al-Zaidi, R. Information Security Risk Assessment Methods in Cloud Computing: Comprehensive Review // Journal of Computer Information Systems. 2024. С. 1–28. DOI: 10.1080/08874417.2024.2329985.
18. Крюков Р. О., Федорченко Е. В., Котенко И. В., Новикова Е. С., Зима В. М. Оценивание защищенности гетерогенных инфраструктур на основе графов атак с использованием баз данных NVD и MITRE ATT & CK / Р. О Крюков, Е. В Федорченко, И. В. Котенко, Е. С Новикова, В. М. Зима / Информационно-управляющие системы. 2024. Т.2., С. 39-50. DOI: 10.31799/1684-8853-2024-2-39-50.
19. Wang J. et al. Big data service architecture: a survey // Journal of Internet Technology. 2020. Т. 21. №. 2. С. 393–405.
20. Omotunde H., Ahmed M. A comprehensive review of security measures in database systems: Assessing authentication, access control, and beyond // Mesopotamian Journal of CyberSecurity. 2023. Т. 2023. С. 115–133.
21. El Ahdab L. et al. Unified Models and Framework for Querying Distributed Data Across Polystores //International Conference on Research Challenges in Information Science. Cham: Springer Nature Switzerland. 2024. С. 3–18.
22. Wang S. et al. Data privacy and cybersecurity challenges in the digital transformation of the banking sector // Computers & security. 2024. Т. 147.
23. Poltavtseva M., Aleksandrova E., Izotova O. Data modeling for consistent access control in heterogeneous big data systems // 2024 Ivannikov Memorial Workshop (IVMEM). Velikiy Novgorod, Russian Federation. 2024. С. 42–48. DOI: 10.1109/IVMEM63006.2024.10659707
24. Hu V. C. et al. An access control scheme for big data processing // 10th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing. IEEE. 2014. С. 1–7.

AN APPROACH TO ANALYZING AND EVALUATING BIG DATA MANAGEMENT SYSTEMS SECURITY

Poltavtseva M. A.,[5], Zegzhda D. P.[6]

Keywords: *big data, information security, security assessment, data granulation, access control, trustworthiness.*

Purpose of the study: *development of an approach to analyzing and evaluating the security of big data management systems, taking into account the technological features of this class of solutions that distinguish them from traditional cloud data processing systems.*

Methods of research: *the paper analyzes the features of target systems, as well as the methods of data collection and security assessment proposed by another researchers. Their disadvantages are highlighted in the context of modern requirements. It is proposed to use an approach to data collection and modeling of the target system using the aggregate data model based on set theory, as well as the integration of a modified NIST assessment of access control in big data systems and an author's assessment, based on data granulation and trust in processing nodes.*

Result(s): *as the result of the work, the technological features of the target systems were formulated in terms of security assessment. They are distribution, heterogeneity (multimodality), and a complex data lifecycle. The analysis of scientific papers showed, on the one hand, the interest of researchers in the task of assessing the security of big data management*

5 Maria A. Poltavtseva, Dr.Sc. of Technical Sciences, Associate Professor, Peter the Great St. Petersburg Polytechnic University. St. Petersburg, Russia. E-mail: poltavtseva@ibks.spbstu.ru

6 Dmitry P. Zegzha, Corresponding Member of the Russian Academy of Sciences, Dr.Sc. of Technical Sciences, Professor, Federal State Autonomous Educational Institution of Higher Education «Peter the Great St. Petersburg Polytechnic University». St. Petersburg, Russia. E-mail: zegzhda_dp@spbstu.ru

systems, and on the other hand, the lack of estimates proposed for the target class of systems. The authors have formulated security assessment requirements for big data management systems as a specific component of modern information systems. A new security assessment method is also proposed, which for the first time takes into account the specific properties of big data management systems. The proposed method, in addition to the previously proposed estimates, takes into account the disadvantages of access control caused by various data granulation in the target system components. As well, as a large number of trusted users. And, as a result, the need to process confidential data either on trusted nodes or in a hidden (obfuscated or encrypted) form. The proposed estimate is normalized, can be detailed to the evaluation of each specific data processing tool, easily expanded or integrated into higher-level estimates. The reliability and possibility of practical application of the proposed assessment is shown by developing a software prototype based on previously known and tested software solutions.

***Scientific novelty:** the novelty lies in the author's method of assessing the security of big data management systems, which differs for the first time by taking into account the disadvantages of access control caused by different granulation of data and taking into account the trust in individual data processing nodes.*

References

1. Minzov A. S., Nevskij A. Ju., Baronov O. R. Bezopasnost' personal'nyh dannyh: novyj vzgljad na staruju problemu // Voprosy kiberbezopasnosti. 2022. №. 4(50). S. 2–12. DOI:10.21681/2311-3456-2022-4-2-12.
2. Colombo P., Ferrari E. Access control technologies for Big Data management systems: literature review and future trends // Cybersecurity. 2019. T. 2. №. 1. S. 1–13.
3. Rafiq F. et al. Privacy Prevention of Big Data Applications: A Systematic Literature Review // SAGE Open. T.12(2). DOI: 10.1177/21582440221096445
4. Markov A. S., Varenitca V. V., Arustamyan S. S. Topical Issues in the Implementation of Secure Software Development Processes // Proceedings of the International Conference on Information Processes and Systems Development and Quality Assurance IPSQDA-2023 (March 22–24, 2023, St. Petersburg Russia). IEEE. 2023. C. 48–54.
5. Alhazmi H. E., Eassa F. E., Sandokji S. M. Towards Big Data Security Framework by Leveraging Fragmentation and Blockchain Technology // IEEE Access. 2022. T. 10. S. 10768–10782. DOI: 10.1109/ACCESS.2022.3144632.
6. Wang T. et al. Edge-based auditing method for data security in resource-constrained Internet of Things // Journal of Systems Architecture. 2021. T. 114. C. 1–10. DOI: 10.1016/j.sysarc.2020.101971.
7. Stodt, J. at al. Security Audit of a Blockchain-Based Industrial Application Platform // Algorithms. 2021. T. 14(4), 121 c. DOI:10.3390/a14040121.
8. Kalinin M., Poltavtseva M., Zegzhda D. Ensuring the Big Data Traceability in Heterogeneous Data Systems // 2023 International Russian Automation Conference (RusAutoCon). Sochi, Russian Federation. 2023. C. 775–780. DOI: 10.1109/RusAutoCon58002.2023.10272905.
9. Attaallah A. et al. Analyzing the Big Data Security Through a Unified Decision-Making Approach // Intelligent Automation & Soft Computing, 2022. T. 32(2). C. 1071–1088. DOI: 10.32604/iasc.2022.022569.
10. Yang M. Information security risk management model for big data // Advances in Multimedia 2022. T.1 C. 1–10 DOI: 10.1155/2022/3383251
11. Theodorakopoulos L., Theodoropoulou A., Stamatiou Y. A State-of-the-Art Review in Big Data Management Engineering: Real-Life Case Studies, Challenges, and Future Research Directions // Eng. 2024. T. 5(3). C. 1266–1297. DOI: 10.3390/eng5030068.
12. Kalinin M., Poltavtseva M. Big Data Security Evaluation by Bidirectional Analysis of Access Control Policy // 2024 International Russian Smart Industry Conference (SmartIndustryCon). Sochi, Russian Federation. 2024. C. 98–103. DOI: 10.1109/SmartIndustryCon61328.2024.10515459.
13. Poltavtseva, M. A., Zaitseva, V. V., Ivanov, D. V. Assessing the Security of Big Data Systems // Aut. Control Comp. Sci. 2024. T. 58. C. 1352–1364. DOI: 10.3103/S0146411624701025.
14. Dhillon G., Smith K., Dissanayaka I. Information systems security research agenda: Exploring the gap between research and practice // The Journal of Strategic Information Systems. 2021. T. 30. № 4. DOI: 10.1016/j.jsis.2021.101693.
15. Kostogryzov, A. I. Metodicheskie polozhenija po verojatnostnom prognozirovaniju kachestva funkcionirovanija informacionnyh sistem ch. 1–3 / A. I. Kostogryzov, A. A. Nistratov, P. E. Golosov // Voprosy kiberbezopasnosti. 2025. № 2(66). S. 2–19. DOI: 10.21681/2311-3456-2025-2-2-19.
16. Ocenka ujazvimostej avtomatizirovannyh sistem s primeneniem teorii verojatnostej, raspredelenija St'judenta i normal'nyh sluchajnyh velichin / I. V. Atlasov, A. O. Efimov, E. A. Rogozin, A. S. Cherkasova // Voprosy kiberbezopasnosti. 2025. № 2(66). S. 124–131. DOI: 10.21681/2311-3456-2025-2-124-131.
17. Ali, T., Al-Khalidi, M., Al-Zaidi, R. Information Security Risk Assessment Methods in Cloud Computing: Comprehensive Review // Journal of Computer Information Systems. 2024. C. 1–28. DOI: 10.1080/08874417.2024.2329985.
18. Krjukov, R. O., Fedorchenko, E. V., Kotenko, I. V., Novikova, E. S., Zima, V. M. Ocenivanie zashhishhennosti geterogennyh infrastruktur na osnove grafov atak s ispol'zovaniem baz dannyh NVD i MITRE ATT & CK. / R. O Krjukov, E. V Fedorchenko, I. V. Kotenko, E. S Novikova, V. M. Zima / Informacionno-upravljajushhie sistemy. 2024. T. 2., C. 39–50. DOI: 10.31799/1684-8853-2024-2-39-50.
19. Wang J. et al. Big data service architecture: a survey //Journal of Internet Technology. 2020. T. 21. №. 2. S. 393–405.
20. Omotunde H., Ahmed M. A comprehensive review of security measures in database systems: Assessing authentication, access control, and beyond //Mesopotamian Journal of CyberSecurity. 2023. T. 2023. S. 115–133.
21. El Ahdab L. et al. Unified Models and Framework for Querying Distributed Data Across Polystores //International Conference on Research Challenges in Information Science. Cham: Springer Nature Switzerland. 2024. S. 3–18.
22. Wang S. et al. Data privacy and cybersecurity challenges in the digital transformation of the banking sector //Computers & security. 2024. T. 147.
23. Poltavtseva M., Aleksandrova E., Izotova O. Data modeling for consistent access control in heterogeneous big data systems // 2024 Ivannikov Memorial Workshop (IVMEM). Velikiy Novgorod, Russian Federation. 2024. C. 42–48. DOI: 10.1109/IVMEM63006.2024.10659707.
24. Hu V. C. et al. An access control scheme for big data processing //10th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing. IEEE. 2014. C. 1–7.

МНОГОУРОВНЕВАЯ АРХИТЕКТУРА СИСТЕМЫ МОНИТОРИНГА И РЕАГИРОВАНИЯ НА ВОЗДЕЙСТВИЯ В ЭРГАТИЧЕСКИХ СИСТЕМАХ

Мещеряков Р. В.[1], Селиверстов Д. Е.[2], Русаков К. Д.[3]

DOI: 10.21681/2311-3456-2025-5-119-127

Цель исследования: *проектирование многоуровневой системы мониторинга и реагирования на воздействия, обеспечивающей устойчивость сложных эргатических систем за счет использования резервного контура управления и адаптивного распределения ресурсов.*

Методы исследования: *системный анализ, моделирование, синтез архитектуры, распределение ресурсов.*

Результаты исследования: *разработана многоуровневая архитектура системы мониторинга и реагирования на воздействия, направленная на повышение устойчивости сложных эргатических систем. Предложено структурное решение, включающее основной и резервный контуры управления, что обеспечивает непрерывность мониторинга и координацию реагирования даже при частичной деградации или отказе коммуникационной инфраструктуры. Спроектирован механизм адаптивного перераспределения ресурсов между компонентами архитектуры, обеспечивающий эффективное функционирование системы в условиях переменной нагрузки и ограниченных вычислительных и сетевых возможностей. Теоретическая значимость работы заключается в развитии научных представлений о построении многоуровневых архитектур защиты эргатических систем, обеспечивающих их функционирование в условиях сложных воздействий. Практическая значимость определяется возможностью применения спроектированных архитектурных решений при создании и модернизации распределенных автоматизированных комплексов различного назначения с целью повышения их устойчивости и эффективности мониторинга.*

Научная новизна: *впервые предложена многоуровневая архитектура системы мониторинга и реагирования на воздействия, в которой реализовано разделение на основной и резервный контуры управления, обеспечивающее устойчивость функционирования при деградации связи и отказах. Впервые разработан механизм адаптивного перераспределения ресурсов между уровнями системы, позволяющий поддерживать эффективность мониторинга и реагирования в условиях переменной нагрузки и ограниченных вычислительных и сетевых возможностей.*

Ключевые слова: *отказоустойчивость, киберустойчивость, федеративное взаимодействие, резервный контур управления, адаптивное распределение ресурсов, эргатические системы, автоматизированные комплексы, кибербезопасность.*

Введение

Современные эргатические системы управления представляют собой сложные интеграции технических средств, программного обеспечения и операторов, обеспечивающие выполнение критичных процессов в промышленности, транспорте и других социально-значимых сферах [1]. Их устойчивое функционирование напрямую зависит от способности противостоять разнообразным информационно-техническим воздействиям (ИТВ), включая преднамеренные кибератаки, случайные сбои оборудования, ошибки операторов и деградацию каналов связи. С ростом масштабов цифровизации и взаимосвязанности компонентов таких систем значительно возрастает риск каскадных отказов, когда локальные нарушения могут привести к разрушению всей функциональной цепочки [2]. Особенно критично это для систем, обеспечивающих безопасность государства и общества, где последствия отказов могут быть катастрофическими.

Традиционные централизованные системы мониторинга и реагирования, в том числе классические системы обнаружения и предотвращения вторжений (IDS/IPS), остаются уязвимыми к ряду фундаментальных проблем. Основной недостаток – наличие единой точки отказа, при выходе из строя которой вся система защиты становится неработоспособной. Централизованные решения плохо масштабируются и неэффективно обрабатывают большие объёмы данных при динамически изменяющейся нагрузке [3]. В современных исследованиях активно развиваются многоуровневые и распределённые архитектуры, которые позволяют повысить отказоустойчивость и гибкость реагирования. Примером являются гибридные системы, объединяющие периферийные и облачные

1 Мещеряков Роман Валерьевич, доктор технических наук, профессор РАН, ФГБУН «Институт проблем управления им. В. А. Трапезникова Российской академии наук», г. Москва, Россия. ORCID: 0000-0002-1129-8434. E-mail: mrv@ipu.ru

2 Селиверстов Дмитрий Евгеньевич, кандидат технических наук, ФГБУН «Институт проблем управления им. В. А. Трапезникова Российской академии наук», г. Москва, Россия. ORCID: 0009-0004-8412-7873. E-mail: Seliverstov_dmitriyy@rambler.ru

3 Русаков Константин Дмитриевич, «Институт проблем управления им. В. А. Трапезникова Российской академии наук», г. Москва, Россия. ORCID: 0009-0004-8412-7873. E-mail: rusakov@ipu.ru

вычислительные узлы, что позволяет распределять вычислительную нагрузку и обеспечивать баланс между быстротой реагирования и глубиной аналитики [4]. Важным направлением развития является также федеративный принцип взаимодействия (распределённое взаимодействие автономных уровней без передачи исходных данных), реализованный, например, в многоагентных системах на основе методов глубокого обучения, позволяющих обмениваться знаниями между узлами без передачи чувствительных данных [5]. Однако анализ отечественных и зарубежных решений показал, что существующие разработки недостаточно учитывают необходимость резервирования управления и адаптивного перераспределения ресурсов, что особенно важно в условиях деградации коммуникационной инфраструктуры и ограниченных вычислительных возможностей [6].

В аналитическом исследовании был проведён анализ патентов и научных публикаций, посвящённых системам мониторинга и реагирования на ИТВ. Результаты анализа показали, что несмотря на значительный прогресс, сохраняется проблема отсутствия комплексных архитектур, которые могли бы функционировать в условиях неопределённости и частичных отказов. Это определяет актуальность задач, связанных с проектированием многоуровневых систем нового поколения, в которых сочетаются принципы федеративного взаимодействия, резервирования и адаптивного управления ресурсами.

Цель настоящей работы заключается в проектировании многоуровневой архитектуры системы мониторинга и реагирования на ИТВ с резервным контуром управления и механизмом адаптивного распределения ресурсов, направленной на обеспечение устойчивости сложных эргатических систем в условиях динамических внешних и внутренних воздействий. В статье представлено структурное решение проектируемой архитектуры и механизм адаптивности, а также обсуждается теоретическая и практическая значимость предложенного подход.

Анализ архитектурных решений

В эргатических системах устойчивость функционирования напрямую зависит от архитектуры средств мониторинга и реагирования на информационно-технические воздействия (ИТВ). Архитектурное решение определяет, насколько эффективно система способна противостоять отказам отдельных элементов, деградации каналов связи и динамически изменяющимся нагрузкам [7]. Современные подходы к построению систем мониторинга и реагирования эволюционировали от централизованных решений к многоуровневым и распределённым системам. Однако, даже самые современные разработки имеют свои ограничения, что обуславливает необходимость поиска новых архитектурных принципов. В данном разделе проводится анализ существующих решений с выделением их преимуществ и недостатков для определения требований к проектируемой системе.

Централизованные архитектуры исторически стали первыми решениями для систем защиты информации и мониторинга. Они характеризуются наличием единого центра, который принимает все решения по обработке данных и реагированию. Такие системы просты в администрировании и понятны с точки зрения логики управления [8], однако их главный недостаток – наличие единой точки отказа. Сбой или атака на центральный узел приводит к полной потере контроля, что делает такие решения непригодными для высоконагруженных и критичных систем. Централизованные архитектуры плохо масштабируются при росте числа элементов и генерируемых событий [9].

Следующим этапом развития стали многоуровневые системы, где функции мониторинга и реагирования распределяются между локальным, промежуточным и глобальным уровнями. Пример такой системы представлен в [10], где локальные узлы выполняют первичный анализ, промежуточные – корреляцию данных и принятие решений в своих сегментах, а глобальный уровень обеспечивает стратегическую координацию. Многоуровневая структура позволяет повысить отказоустойчивость и снизить нагрузку на центральные компоненты, но остаётся зависимость от верхнего уровня: при его отказе координация сегментов нарушается.

Современные исследования активно развивают распределённые архитектуры, где все узлы равноправны и обмениваются информацией напрямую без выделенного центра управления. Такой подход позволяет минимизировать риски полного отказа системы и повысить её гибкость [11], однако он предъявляет повышенные требования к алгоритмам синхронизации и маршрутизации данных. В работе [12] показано, что распределённые решения обеспечивают высокую устойчивость к динамическим изменениям структуры сети, но сложность их внедрения и администрирования пока остаётся серьёзным ограничением.

Федеративные архитектуры сочетают принципы многоуровневых и распределённых систем. Узлы обрабатывают данные локально и передают на верхние уровни только агрегированные результаты. Такой подход позволяет снизить нагрузку на каналы связи, сохранить конфиденциальность данных и повысить масштабируемость [13]. Тем не менее, федеративные решения требуют сложных алгоритмов адаптивного управления ресурсами, чтобы обеспечить баланс между локальной автономностью и глобальной координацией [14].

Анализ существующих подходов показал, что ни одна из рассмотренных архитектур не обеспечивает комплексного решения по ключевым направлениям – устойчивости, адаптивности и резервированию управления. Для построения многоуровневой системы мониторинга и реагирования на информационно-технические воздействия необходимо выделить совокупность требований, которые станут основой для проектируемой архитектуры и механизма её функционирования. На основе проведённого системного анализа и современных научных исследований [15–20] сформулирован ряд ключевых требований. Рассмотрим их подробно. Первым требованием является отказоустойчивость и отсутствие единой точки отказа. Здесь каждый уровень и узел системы должны сохранять базовые функции мониторинга и реагирования при отказе центральных компонентов или деградации каналов связи. Второе требование касается непрерывности мониторинга за счёт резервного контура управления. Оно подразумевает существование отдельного резервного контура, способного автоматически перехватывать управление при сбое основного уровня и обеспечивать координацию действий в условиях критической ситуации. Третье требование относится к адаптивному распределению ресурсов, то есть, система должна динамически перераспределять вычислительные и сетевые ресурсы в зависимости от текущей нагрузки, доступности каналов и приоритетности задач, минимизируя задержки обработки событий. Четвертое требование предъявляется к масштабируемости и модульности архитектуры, которая должна поддерживать гибкое добавление новых узлов и сегментов без необходимости глобальной реконфигурации системы.

Таким образом можно сделать вывод, что для эффективного противодействия информационно-техническим воздействиям архитектура должна быть спроектирована таким образом, чтобы обеспечивать устойчивость, адаптивность и резервирование управления. Эти три фактора являются взаимосвязанными и определяют целостность решения. Таким образом, архитектура должна объединять преимущества многоуровневых и распределённых решений, дополняясь механизмами адаптивности и резервирования, что позволяет создать основу для высокоэффективных систем мониторинга и реагирования.

Постановка задачи исследования

Исходя из результатов проведенного анализа, задача исследования формулируется следующим образом: необходимо разработать многоуровневую архитектуру системы мониторинга и реагирования на информационно-технические воздействия эргатических систем, в которой реализованы требования устойчивости, адаптивности и резервирования управления. Архитектура должна включать основной и резервный контуры управления, а также механизм адаптивного перераспределения вычислительных и сетевых ресурсов, обеспечивающий эффективное функционирование системы в условиях динамически изменяющихся нагрузок, частичных отказов и деградации коммуникационной инфраструктуры. Формально задача исследования примет следующий вид. Дано: множество исходных данных $D = \langle V,E,\Lambda,\Omega\rangle$, V – множество узлов (компонентов системы мониторинга и реагирования), E – множество каналов связи между узлами, Λ – характеристики потоков данных и динамически изменяющихся нагрузок, Ω – множество сценариев частичных отказов и деградации коммуникационной инфраструктуры. Требуется: разработать многоуровневую архитектуру $A = \langle L,C,M\rangle$ где L – структура уровней архитектуры (локальный, сегментный, глобальный), C – основной и резервный контуры управления, M – механизм адаптивного перераспределения ресурсов между уровнями и узлами. При этом, архитектура A определяется на основе исходных данных $A = F(D)$, и обеспечивает выполнение таких требований как: устойчивость ($\forall\omega \in \Omega$: $\Phi_{work}(A,\omega) \geq \Phi_{min}$, где Φ_{work} – функциональность архитектуры при сценарии ω, а Φ_{min} – минимально допустимый уровень функционирования системы); адаптивность (M:$(\Lambda) \longrightarrow \Lambda'$, где M обеспечивает динамическое перераспределение ресурсов и потоков данных при изменениях нагрузки и состояния системы; резервирование управления ($\forall v \in V$: $\exists\ p(v,C_{main}) \vee p(v,C_{res})$, где $p(v,C)$ – наличие связного пути от узла v к основному C_{main} или резервному C_{res} контуру управления.

Разработка архитектуры

В предыдущих разделах статьи были определены исходные данные $D = \langle V,E,\Lambda,\Omega\rangle$ и сформулированы ключевые требования к разрабатываемой архитектуре: устойчивость, адаптивность и резервирование управления. Эти требования отражают специфику функционирования эргатических систем и являются основой для дальнейших проектных решений. На данном этапе основная цель – разработка многоуровневой архитектуры $A = \langle L,C,M\rangle$, которая позволит системе мониторинга и реагирования эффективно функционировать в условиях: динамически изменяющихся нагрузок Λ, частичных отказов и деградации инфраструктуры Ω, повышенных требований к непрерывности и надежности процессов управления.

Проектирование архитектуры предполагает последовательное выполнение нескольких шагов:

- определение уровневой структуры системы L, обеспечивающей устойчивость и разделение функций между компонентами;

- разработка основного и резервного контуров управления C, которые обеспечивают непрерывность мониторинга и реагирования при штатном режиме и в условиях отказов;
- создание механизма адаптивного перераспределения ресурсов M, необходимого для эффективного функционирования системы при изменении нагрузки и состояния сети;
- интеграция всех элементов в единую архитектуру и проверка её соответствия поставленным требованиям.

Формирование уровневой структуры архитектуры является ключевым этапом проектирования, поскольку именно структура уровней определяет, каким образом система будет обеспечивать устойчивость, адаптивность и резервирование управления. Основой проектируемого решения является концепция многоуровневой организации, позволяющая гибко распределять функции между компонентами системы, минимизировать риски возникновения единой точки отказа и обеспечивать непрерывность работы в условиях деградации или частичных отказов инфраструктуры. Процесс проектирования уровней начинается с анализа особенностей информационно-технических воздействий и требований, вытекающих из постановки задачи. В частности, необходимо, чтобы каждый уровень выполнял собственный набор функций, а взаимодействие между уровнями обеспечивало согласованность процессов мониторинга и реагирования. Для достижения этих целей в рамках исследования предложено выделить три функциональных уровня: локальный, сегментный (промежуточный) и глобальный, которые формируют множество $L = \{L_1, L_2, L_3\}$.

Одноуровневые архитектуры, несмотря на простоту реализации, обладают существенными ограничениями. Централизация всех функций на одном уровне неизбежно создаёт единую точку отказа: сбой ключевого узла или канала связи приводит к полной потере работоспособности системы. Кроме того, такие решения не масштабируются при увеличении числа узлов и объёмов данных, что особенно критично для современных эргатических систем.

В противоположность этому, полностью распределённые решения обеспечивают высокий уровень отказоустойчивости за счёт равноправного взаимодействия всех узлов, однако требуют сложных алгоритмов синхронизации и маршрутизации данных. Это делает их уязвимыми к деградации связей и затрудняет централизованную координацию действий в условиях кризисных ситуаций.

Многоуровневый подход занимает промежуточное положение между этими крайностями. Он позволяет сохранить локальную автономность нижних уровней, при этом обеспечивая согласованность работы всей системы за счёт верхнего уровня, ответственного за стратегическое управление. Такая организация делает возможным как горизонтальное взаимодействие между сегментами, так и вертикальное управление потоками данных и ресурсов, что напрямую соответствует требованиям устойчивости, адаптивности и резервирования, определённым в постановке задачи.

В разработанной архитектуре выделяются три уровня, каждый из которых выполняет собственные функции и взаимодействует с другими уровнями по строго определённым правилам: локальный, сегментный и глобальный – см. рис. 1.

Рис. 1. Многоуровневая архитектура системы мониторинга и реагирования на информационно-технические воздействия

Локальный уровень (L_1) расположен на границе взаимодействия с физической или виртуальной средой. Его задача – первичная обработка данных, формируемых источниками информации. На этом уровне выполняются: фильтрация и нормализация поступающих событий; обнаружение простейших аномалий и известных сигнатур угроз; предварительное ранжирование событий по степени критичности; передача агрегированных данных на вышестоящие уровни для последующего анализа; выполнение локальных реакций, не требующих сложной координации (например, временная блокировка узла или ограничение сетевого трафика). Наличие автономных функций на локальном уровне позволяет системе реагировать на инциденты даже при частичной потере связи с верхними уровнями.

Сегментный уровень (L_2) выполняет роль промежуточного слоя, который объединяет несколько локальных узлов в рамках одного технологического сегмента. Основные задачи сегментного уровня включают: корреляцию событий, поступающих с локальных узлов; консолидацию контекстной информации о состоянии сегмента; принятие решений в границах сегмента при недоступности глобального уровня; управление распределением ресурсов между узлами данного сегмента; временное исполнение функций резервного контура управления в случае отказа глобального уровня. Благодаря этому сегментный уровень выступает в роли буфера, обеспечивая баланс между автономностью локальных узлов и стратегическим управлением всей системой.

Глобальный уровень предназначен для координации работы всей системы в целом. Его задачами являются: межсегментная корреляция данных и построение целостной картины состояния системы; формирование и распространение политик реагирования на информационно-технические воздействия; стратегическое распределение ресурсов между сегментами; активация или деактивация основного контура управления и передача функций резервному контуру в случае необходимости; анализ эффективности функционирования нижних уровней и адаптация их параметров. Таким образом, глобальный уровень обеспечивает высокий уровень согласованности действий всех компонентов и является ядром стратегического управления.

Взаимодействие между уровнями носит двунаправленный характер. Восходящие потоки данных ($L_1 \longrightarrow L_2 \longrightarrow L_3$) включают события, метрики и агрегированные результаты анализа, необходимые для формирования глобальной картины состояния системы. Нисходящие потоки управления ($L_3 \longrightarrow L_2 \longrightarrow L_1$) содержат политики реагирования, приоритеты и распределение ресурсов. Особая роль отводится резервному контуру управления: при отказе глобального уровня функции координации временно передаются сегментному уровню. Если же деградация затрагивает часть сегментных узлов, соседние сегменты берут на себя их управление, обеспечивая непрерывность работы всей системы.

Сформированная многоуровневая структура $L = \{L_1, L_2, L_3\}$ обеспечивает: устранение единой точки отказа за счёт распределения функций по уровням; адаптивность через возможность перераспределения ролей и ресурсов в ответ на изменения нагрузки и состояния сети; реализацию резервирования управления за счёт автоматического перераспределения функций между уровнями при отказах отдельных компонентов. Таким образом, предложенная уровневая структура создаёт фундамент для построения основной и резервной логики управления системой.

Проектирование контуров управления

Функционирование многоуровневой архитектуры невозможно без чётко организованных контуров управления, которые обеспечивают согласованность действий всех компонентов системы. В рамках разработанной структуры $L = \{L_1, L_2, L_3\}$ выделяются два взаимосвязанных контура управления – основной (C_{main}) и резервный (C_{res}). Их совместная работа направлена на обеспечение непрерывности процессов мониторинга и реагирования даже при частичных отказах инфраструктуры или деградации каналов связи. Основной контур обеспечивает работу системы в штатных условиях, тогда как резервный вступает в действие только при обнаружении сбоев или потере связи с ключевыми компонентами верхнего уровня. Такое разделение позволит не только снизить нагрузку на систему при нормальной работе, но и гарантировать устойчивость управления в кризисных ситуациях.

Основной контур управления предназначен для организации потоков данных и принятия решений в штатных условиях функционирования системы. Он базируется на принципах иерархической координации: нижние уровни (L_1, L_2) предоставляют агрегированные данные и локальные решения, а глобальный уровень (L_3) осуществляет анализ и формирование стратегических команд.

В рамках C_{main} реализуются следующие процессы:

- сбор и агрегация данных на локальном и сегментном уровнях (локальные узлы (L_1) выполняют первичную обработку и передают результаты на сегментный уровень (L_2), где осуществляется корреляция и выделение критических событий);
- формирование стратегических решений на глобальном уровне (L_3) (данные анализируются в межсегментном контексте, что позволяет выявлять комплексные угрозы и определять приоритеты реагирования;

- нисходящая передача команд (глобальный уровень передаёт сегментам и локальным узлам инструкции, которые реализуют заданные сценарии реагирования и распределения ресурсов);
- обратная связь – на каждом цикле управления данные о выполнении команд возвращаются вверх, что позволяет корректировать стратегию и предотвращать каскадные сбои.

Таким образом, основной контур обеспечивает скоординированное взаимодействие всех уровней в нормальном режиме работы, поддерживая баланс между скоростью реакции и глубиной анализа.

Несмотря на эффективность основного контура, реальная эксплуатация сложных эргатических систем неизбежно связана с рисками отказов или временной недоступности ключевых компонентов. Для минимизации последствий подобных ситуаций необходим резервный контур управления, который обеспечивает непрерывность функционирования системы при нарушениях штатной структуры.

Резервный контур активируется в следующих случаях:

- потеря связи между сегментным (L_2) и глобальным (L_3) уровнями;
- отказ центрального управляющего узла или деградация критических каналов связи;
- резкий рост нагрузки, при котором основной контур не успевает обрабатывать входящие данные.

В режиме работы резервного контура часть функций глобального уровня временно передаётся сегментным узлам (L_2). Эти узлы берут на себя задачи по координации локальных узлов и принятию тактических решений на уровне сегмента. Взаимодействие между сегментами осуществляется по принципу горизонтальной федерации: соседние сегментные узлы могут обмениваться агрегированными данными и координировать свои действия без участия глобального уровня.

Ключевым элементом C_{res} является механизм синхронизации данных. При восстановлении основного контура резервный обеспечивает передачу накопленной информации о принятых решениях и текущем состоянии системы на глобальный уровень, что позволяет быстро вернуть архитектуру в штатный режим без потери данных и конфликтов между уровнями.

Разработанные основной и резервный контуры управления образуют единый комплекс $\{C_{main}, C_{res}\}$, обеспечивающий высокий уровень отказоустойчивости и непрерывности функционирования системы. Их взаимодействие позволяет: поддерживать нормальную работу системы в штатных условиях; сохранять способность к принятию решений при частичных отказах или деградации инфраструктуры; минимизировать последствия отказов за счёт быстрого автоматического переключения на резервный режим; гарантировать согласованность данных и процессов при возврате в стандартное состояние.

Таким образом, спроектированные контуры управления напрямую реализуют требование резервирования, сформулированное на этапе постановки задачи, и создают основу для дальнейшего синтеза механизма адаптивного распределения ресурсов.

Механизм адаптивного распределения ресурсов и интеграция архитектуры

Разработанная многоуровневая архитектура требует наличия управляющего механизма, который обеспечит её гибкость и способность реагировать на изменения внешних и внутренних условий функционирования. В этой роли выступает механизм адаптивного распределения ресурсов M, предназначенный для динамической настройки процессов мониторинга и реагирования в зависимости от текущей ситуации. Его ключевая задача – перераспределять ресурсы между уровнями и сегментами системы на основе анализа состояния сети и потоков данных Ω.

В основе работы M лежит использование параметров, описывающих текущее состояние системы:

- интенсивность событий $\lambda(t)$ – поток данных, поступающих на узлы в момент времени t, что отражает текущую активность сети и уровень информационно-технических воздействий;
- состояние каналов связи $\delta(t)$ – метрики качества каналов: пропускная способность, задержки, уровень ошибок, что позволяет оценивать доступность и надежность коммуникаций;
- загрузка узлов $\rho(t)$ – уровень использования вычислительных мощностей на каждом уровне архитектур;
- критичность потоков z – приоритеты обработки данных, зависящие от типа события и его значимости для защищаемой системы.

Данные параметры формируют динамическую карту состояния, которая служит входными данными для принятия решений о перераспределении ресурсов.

Механизм M работает циклично и включает три последовательных этапа. На первом этапе осуществляется оценка состояния системы. Здесь сегментный уровень (L_2) собирает телеметрию о нагрузке, доступности каналов и узлов. Данные агрегируются и передаются на глобальный уровень (L_3), где формируется целостная картина состояния системы. На втором этапе реализуется определение приоритетов, на основе критичности потоков z определяется порядок обработки событий. Критические инциденты

и процессы, влияющие на устойчивость системы, получают наивысший приоритет. Третий этап включает в себя перераспределение ресурсов, где часть задач узлов L_1 может быть передана на уровень L_2 для разгрузки периферии. Такой подход позволяет системе поддерживать стабильную работу даже при резких изменениях нагрузки или частичных отказах инфраструктуры.

Механизм M не функционирует изолированно, а тесно связан с уровнями архитектуры L и контурами управления C, формируя единую систему. Уровни L обеспечивают физическую и логическую основу системы, контуры C поддерживают непрерывность и резервирование функций, гарантируя, что даже при отказе отдельных компонентов система сохраняет способность к координации, а механизм M выступает связующим звеном, адаптивно настраивая взаимодействие уровней и обеспечивая эффективное использование доступных ресурсов. Интеграция всех элементов в единую структуру позволяет архитектуре удовлетворять требованиям устойчивости, адаптивности и резервирования, определённым в постановке задачи.

В результате интеграции сформирована архитектура $A = \langle L,C,M \rangle$, в которой уровни, контуры управления и механизм адаптивности функционируют как единое целое. Такая организация позволяет системе не только противостоять текущим информационно-техническим воздействиям, но и активно адаптироваться к изменяющимся условиям эксплуатации, обеспечивая высокий уровень надёжности и непрерывности работы.

Заключение

В ходе проведенного исследования предложена многоуровневая архитектура системы мониторинга и реагирования на информационно-технические воздействия в сложных эргатических системах. Разработанная архитектура включает три взаимосвязанных компонента:

- уровневую структуру L, обеспечивающую распределение функций между локальными, сегментными и глобальными узлами;
- два контура управления C_{main} и C_{res} реализующих резервирование и непрерывность процессов мониторинга и реагирования;
- механизм адаптивного распределения ресурсов M, позволяющий динамически настраивать работу системы в условиях изменяющихся нагрузок и частичных отказов.

Впервые предложено объединение этих элементов в единую архитектуру $A = \langle L,C,M \rangle$ которая удовлетворяет требованиям устойчивости, адаптивности, резервирования управления. Проведённая аналитическая проверка и сценарное моделирование показали, что данное решение позволяет исключить наличие единой точки отказа и поддерживать функционирование системы даже при деградации каналов связи, резком росте нагрузки или отказе глобального уровня.

Теоретическая значимость работы заключается в развитии научных представлений о проектировании многоуровневых архитектур для защиты эргатических систем и формализации принципов их построения на основе системного анализа. Практическая значимость определяется возможностью применения полученных результатов при создании или модернизации распределённых автоматизированных комплексов различного назначения – в промышленности, транспорте, энергетике, связи – для повышения их устойчивости и эффективности процессов мониторинга и реагирования.

Таким образом, предложенная архитектура формирует основу для построения новых поколений интеллектуальных систем обеспечения безопасности, способных адаптироваться к изменяющимся условиям эксплуатации и противостоять современным информационно-техническим воздействиям.

Литература

1. Железнов Э. Г., Комиссаров П. В., Цымай Ю. В. Исследование эргатических систем управления // Современные наукоёмкие технологии. – 2021. – № 4. – С. 45–53.
2. Al-Khaysat H., et al. Risk Assessment for Cyber Resilience of Critical Infrastructures // Applied Sciences. – 2024. – Vol. 14, № 24. – Article 11807. – DOI: 10.3390/app142411807.
3. Diana L., Dini P., Paolini D. Overview on Intrusion Detection Systems for Computers Networking Security // Computers. – 2025. – Vol. 14, no. 3. – P. 87. – DOI: 10.3390/computers14030087.
4. Lezzi M., Corallo A., Lazoi M., Nimis A. Measuring Cyber Resilience in Industrial IoT: A Systematic Literature Review // Management Review Quarterly. – 2025. – Vol. 75, № 4. – С. 1213–1235. – DOI: 10.1007/s11301-025-00495-8.
5. Soltani M., Khajavi K., Jafari Siavoshani M., Jahangir A. H. A multi-agent adaptive deep learning framework for online intrusion detection // Cybersecurity. 2023. Vol. 6, Iss. 2. P. 45–59. DOI 10.1186/s42400-023-00199-0.
6. Калашников А. О., Бугайский К. А., Аникина Е. В., Перескоков И. С., Петров Ан. О., Петров Ал. О., Храмченкова Е. С., Молотов А. А. Применение логико-вероятностного метода в информационной безопасности (Часть 2) // Вопросы кибербезопасности. 2023. № 5(57). С. 113–127. DOI 10.21681/2311-3456-2023-5-113-127.
7. Власов Д. С. Мультикритериальная модель систематизации способов обнаружения инсайдера // Вопросы кибербезопасности. 2024. № 2(60). С. 66–73. DOI 10.21681/2311-3456-2024-2-66-73.

8. Lagraa S., Husak M., Seba H., Vuppala S., State R., & Ouedraogo M. A review on graph-based approaches for network security monitoring and botnet detection // International Journal of Information Security. 2024. Vol. 23. P. 119–140. DOI 10.1007/s10207-023-00742-7.
9. Hu Q., Yu S. -Y., Asghar M. R. Analysing performance issues of open-source intrusion detection systems in high-speed networks // Journal of Information Security and Applications. 2020. Vol. 51. Article 102426. DOI 10.1016/j.jisa.2019.102426.
10. Furrer F. J. Safe and secure system architectures for cyber-physical systems // Informatik Spektrum. – 2023. – Vol. 46. – № 2. – С. 96-103. – DOI: 10.1007/s00287-023-01533-z.
11. Sharma S., Sahay S. K. Evolution and impact of distributed intrusion detection systems in network security and management // Computer Networks. 2022. Vol. 206. Article 108784. DOI 10.1016/j.comnet.2021.108784.
12. Sharma A., Rani S., Boulila W. Blockchain-based zero trust networks with federated transfer learning for IoT security in industry 5.0 // PLOS ONE. 2025. Vol. 20, Iss. 6. Article e0323241. DOI 10.1371/journal.pone.0323241.
13. Lim W.Y.B., Xiong Z., Niyato D., et al. Federated Learning in Mobile Edge Networks: A Comprehensive Survey // IEEE Communications Surveys & Tutorials. 2020. Vol. 22. Iss. 3. PP. 2031–2063. DOI: 10.1109/COMST.2020.2986024.
14. Xu R., Hang L., Jin W., Kim D. Distributed Secure Edge Computing Architecture Based on Blockchain for Real-Time Data Integrity in IoT Environments // Actuators. 2021. Vol. 10, Iss. 8. Article 197. DOI 10.3390/act10080197.
15. Ji R., Padha D., Singh Y. Survey and analysis of intrusion detection frameworks for cyber-physical systems: A comprehensive study // Recent Innovations in Computing. Lecture Notes in Electrical Engineering, vol. 1194. 2024. P. 307–317. DOI 10.1007/978-981-97-2839-8_21.
16. Singh S., Ahmed J., Raghuvanshi K. K., Agarwal P. Adaptive Resource Management Framework for Secure and Resilient IoT Communication Using Federated Learning and Quantum Encryption // Journal of Information Systems Engineering and Management. 2025. Vol. 10, No. 21s. DOI 10.52783/jisem.v10i21s.3405.
17. Rostami M., Goli-Bidgoli S. An overview of QoS-aware load balancing techniques in SDN-based IoT networks // Journal of Cloud Computing. 2024. Vol. 13. Article 89. DOI 10.1186/s13677-024-00651-7.
18. Belenguer A., Navaridas J., Pascual J. A. A review of federated learning in intrusion detection systems for IoT // arXiv. 2022. DOI 10.48550/arXiv.2024.12443.
19. Язов Ю. К., Авсентьев А. О. Пути построения многоагентной системы защиты информации от утечки по техническим каналам // Вопросы кибербезопасности. 2022. № 5(51). С. 2–13. DOI 10.21681/2311-3456-2022-5-2-13.
20. Zareian Jahromi M., Yaghoubi E., Yaghoubi E., Yusupov Z., Maghami M. R. An Innovative Real-Time Recursive Framework for Techno-Economical Self-Healing in Large Power Microgrids Against Cyber-Physical Attacks Using Large Change Sensitivity Analysis // Energies. 2025. Vol. 18, Iss. 1. Article 190. DOI 10.3390/en18010190.

MULTI-LEVEL ARCHITECTURE OF A MONITORING AND RESPONSE SYSTEM TO IMPACTS WITH BACKUP CONTROL AND ADAPTIVE RESOURCE ALLOCATION IN ERGATIC SYSTEMS

Meshcheryakov R. V.[4], Seliverstov D. E.[5], Rusakov K. D.[6]

Keywords: *fault tolerance, cyber resilience, federated interaction, backup control loop, adaptive resource allocation, ergatic systems, automated complexes, cybersecurity.*

Purpose of the article: *the design of a multi-level monitoring and response system for impacts, ensuring the resilience of complex ergatic systems through the use of a backup control loop and adaptive resource allocation.*

Research methods: *system analysis, modeling, architecture synthesis, resource allocation.*

Research results: *a multi-level architecture of a monitoring and response system for impacts has been developed, aimed at improving the resilience of complex ergatic systems. A structural solution is proposed, including primary and backup control loops, which ensures continuous monitoring and coordinated response even in cases of partial degradation or failure of the communication infrastructure. A mechanism for adaptive resource reallocation between architecture components has been designed, ensuring efficient system operation under variable loads and limited computing and network capabilities. The theoretical significance of the work lies in advancing scientific knowledge on the design of multi-level architectures for the protection of ergatic systems, ensuring their functionality under complex impacts. The practical significance is determined by the possibility of applying the designed architectural solutions in the creation and modernization of distributed automated complexes of various purposes to increase their resilience and the efficiency of monitoring processes.*

4 Roman V. Meshcheryakov, D.Sc., Professor of the Russian Academy of Sciences, V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia. ORCID: 0000-0002-1129-8434. E-mail: mrv@ipu.ru
5 Dmitry E. Seliverstov, Ph.D., V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia. ru ORCID: 0009-0004-8412-7873. E-mail: Seliverstov_dmitriyy@rambler
6 Konstantin D. Rusakov, V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia. ORCID: 0009-0004-8412-7873. E-mail: rusakov@ipu.ru

Scientific novelty: *for the first time, a multi-level architecture of a monitoring and response system for impacts is proposed, implementing a separation into primary and backup control loops to ensure operational resilience under communication degradation and failures. For the first time, a mechanism for adaptive resource reallocation between system levels has been developed, allowing the maintenance of monitoring and response efficiency under variable loads and limited computing and network capabilities.*

References

1. Zheleznov, E. G., Komissarov, P. V., & Tsymai, Y. V. (2021). Issledovanie ergaticheskikh sistem upravleniya [Research of ergatic control systems]. Sovremennye Naukoemkie Tekhnologii [Modern High Technologies], (4), 45–53.
2. Al-Khaysat, H., et al. (2024). Risk Assessment for Cyber Resilience of Critical Infrastructures. Applied Sciences, 14(24), Article 11807. https://doi.org/10.3390/app142411807.
3. Diana, L., Dini, P., & Paolini, D. (2025). Overview on Intrusion Detection Systems for Computers Networking Security. Computers, 14(3), Article 87. https://doi.org/10.3390/computers14030087.
4. Lezzi, M., Corallo, A., Lazoi, M., & Nimis, A. (2025). Measuring Cyber Resilience in Industrial IoT: A Systematic Literature Review. Management Review Quarterly, 75(4), 1213–1235. https://doi.org/10.1007/s11301-025-00495-8.
5. Soltani, M., Khajavi, K., Jafari Siavoshani, M., & Jahangir, A. H. (2023). A multi-agent adaptive deep learning framework for online intrusion detection. Cybersecurity, 6(2), 45–59. https://doi.org/10.1186/s42400-023-00199-0.
6. Kalashnikov, A. O., Bugayskiy, K. A., Anikina, E. V., Pereskokov, I. S., Petrov, An. O., Petrov, Al. O., Khramchenkova, E. S., & Molotov, A. A. (2023). Primenenie logiko-veroyatnostnogo metoda v informatsionnoy bezopasnosti (Chast' 2) [Application of the logic-probabilistic method in information security (Part 2)]. Voprosy Kiberbezopasnosti [Cybersecurity Issues], 5(57), 113–127. https://doi.org/10.21681/2311-3456-2023-5-113-127.
7. Vlasov, D. S. (2024). Mul'tikriterial'naya model' sistematizatsii sposobov obnaruzheniya insaydera [multi-criteria model of systematization of insider detection methods]. Voprosy Kiberbezopasnosti [Cybersecurity Issues], 2(60), 66–73. https://doi.org/10.21681/2311-3456-2024-2-66-73.
8. Lagraa, S., Husak, M., Seba, H., Vuppala, S., State, R., & Ouedraogo, M. (2024). A review on graph-based approaches for network security monitoring and botnet detection. International Journal of Information Security, 23, 119–140. https://doi.org/10.1007/s10207-023-00742-7.
9. Hu, Q., Yu, S.-Y., & Asghar, M. R. (2020). Analysing performance issues of open-source intrusion detection systems in high-speed networks. Journal of Information Security and Applications, 51, 102426. https://doi.org/10.1016/j.jisa.2019.102426.
10. Zhu, Q., Rieger, C., & Basar, T. (2011). A hierarchical security architecture for cyber-physical systems. In Proceedings of the 4th International Symposium on Resilient Control Systems (ISRCS 2011) (pp. 15–20). https://doi.org/10.1109/ISRCS.2011.6016081.
11. Sharma, S., & Sahay, S. K. (2022). Evolution and impact of distributed intrusion detection systems in network security and management. Computer Networks, 206, 108784. https://doi.org/10.1016/j.comnet.2021.108784.
12. Sharma, A., Rani, S., & Boulila, W. (2025). Blockchain-based zero trust networks with federated transfer learning for IoT security in industry 5.0. PLOS ONE, 20(6), e0323241. https://doi.org/10.1371/journal.pone.0323241.
13. Lim, W. Y. B., Xiong, Z., Niyato, D., Miao, C., Yang, Q., & Poor, H. V. (2020). Federated learning in mobile edge networks: A comprehensive survey. IEEE Communications Surveys & Tutorials, 22(3), 2031–2063. https://doi.org/10.1109/COMST.2020.2986024.
14. Xu, R., Hang, L., Jin, W., & Kim, D. (2021). Distributed secure edge computing architecture based on blockchain for real-time data integrity in IoT environments. Actuators, 10(8), Article 197. https://doi.org/10.3390/act10080197.
15. Ji, R., Padha, D., & Singh, Y. (2024). Survey and analysis of intrusion detection frameworks for cyber-physical systems: A comprehensive study. In Recent Innovations in Computing (Vol. 1194, pp. 307–317). Springer. https://doi.org/10.1007/978-981-97-2839-8_21.
16. Singh, S., Ahmed, J., Raghuvanshi, K. K., & Agarwal, P. (2025). Adaptive resource management framework for secure and resilient IoT communication using federated learning and quantum encryption. Journal of Information Systems Engineering and Management, 10(21s). https://doi.org/10.52783/jisem.v10i21s.3405.
17. Rostami, M., & Goli-Bidgoli, S. (2024). An overview of QoS-aware load balancing techniques in SDN-based IoT networks. Journal of Cloud Computing, 13, Article 89. https://doi.org/10.1186/s13677-024-00651-7.
18. Belenguer, A., Navaridas, J., & Pascual, J. A. (2022). A review of federated learning in intrusion detection systems for IoT. arXiv. https://doi.org/10.48550/arXiv.2024.12443.
19. Yazov, Y. K., & Avsentyev, A. O. (2022). Puti postroeniya mnogoagentnoi sistemy zashchity informatsii ot utechki po tekhnicheskim kanalam [Ways to build a multi-agent information security system against leakage through technical channels]. Voprosy Kiberbezopasnosti [Cybersecurity Issues], (5)(51), 2–13. https://doi.org/10.21681/2311-3456-2022-5-2-13.
20. Lin D., He Y., Zhang Q. Real-time optimization of network response under cyber-physical attacks // IEEE Transactions on Industrial Informatics. 2025. Vol. 21. Iss. 2. PP. 1501–1513. DOI: 10.1109/TII.2024.3391750.

ОБ ИСПОЛЬЗОВАНИИ ТЕОРИИ ГРАФОВ ПРИ КЛАССИФИКАЦИИ ИНФОРМАЦИИ

Гордеев Э. Н.[1], *Леонтьев В. К.*[2]

DOI: 10.21681/2311-3456-2025-5-128-138

Скончался выдающийся математик, соавтор данной статьи, доктор физико-математических наук, профессор Владимир Константинович Леонтьев.

Владимир Константинович ушел из жизни 22 июля 2025 года в возрасте 83 лет.

Всю жизнь он посвятил развитию дискретного анализа и теории информации, его исследования в области кодирования и дискретной оптимизации получили мировое признание. Профессор Леонтьев разработал новые границы в задачах о покрытии, решил проблему совершенных кодов и создал оптимальные коды, обнаруживающие ошибки. В теории задач дискретной математики им построена теория устойчивости решений и табулирования в дискретных оптимальных задачах на основе понятия «радиус устойчивости».

Как педагог воспитал множество талантливых математиков, преподавал в МФТИ, МГУ, МГТУ. Среди учеников Владимира Константиновича – 16 кандидатов и 2 доктора наук.

Светлая и благодарная память о Владимире Константиновиче Леонтьеве навсегда останется в истории отечественной науки.

Цель данной работы: *проанализировать возможности применения теории графов для кодирования и классификации изображений, что особенно актуально в связи с использованием методов искусственно интеллекта для классификации изображений.*

Метод исследования: *комбинаторика и теория графов, а также эвристические алгоритмы.*

Полученные результаты: *в работе обсуждается возможность применения классических результатов теории графов, касающихся проблем восстановления и распознавания графов и их характеристик, в области распознавания изображений. При этом анализируются различные аспекты задачи описания (представления) графов с помощью их инвариантов.*

Вводятся и рассматриваются новые классы инвариантов для графов, которые, в частности, могут использоваться для анализа и классификации изображений. Кроме того, доказанные в статье утверждения касаются таких аспектов проблемы как формирование сложных типов инвариантов на основе базисных и нахождение функциональных зависимостей одних инвариантов от других.

Научная новизна: *построены и обоснованы новые составные инварианты графов, которые можно эффективно использовать при распознавании изображений, представленных на основе графов.*

Ключевые слова: *распознавание, признаковые таблицы, эвристики, восстановление, инвариант графа, хроматическое число, число независимости, число внешней устойчивости, число внутренней устойчивости.*

Введение. Признаковое распознавание и графы

В работе рассматриваются вопросы, использующие язык теории графов, которые могут иметь разные прикладные аспекты, в том числе, в области распознавания изображений. В качестве модельной иллюстрации прикладная задача, к которой относятся теоретические рассмотрения, представленные в этой работе, могут быть описана следующим образом.

Дано некоторое множество изображений (представлений) одного и того же объекта $I_1,\ldots,I_m$, полученных с помощью средств $T_1,\ldots,T_m$. Мы хотим определить, какие из выбранных средств предпочтительнее.

Можно посмотреть и по-другому. Дано множество изображений одного объекта, в сочетании с изображениями других объектов. Требуется исключить из данного множества эти последние, быть может, исключив и часть изображений рассматриваемого объекта.

Например, имеются фотоснимки объекта той или иной степени четкости. Их надо как-то «предобработать» для хранения и передачи, сообразуясь с теми правилами, согласно которым это хранение и эта передача осуществляются. Так как в любом случае на вход вычислительного алгоритма подается слово

1 Гордеев Эдуард Николаевич, доктор физико-математических наук, профессор кафедры ИУ-8 «Информационная безопасность» МГТУ им. Н. Э. Баумана. Москва, Россия. E-mail: werhorngord@gmail.com

2 Леонтьев Владимир Константинович, доктор физико-математических наук, профессор. Москва, Россия. E-mail: vkleontiev@yandex.ru

в конечном алфавите, то под «предобработкой» понимается представление объекта в виде такого слова (создание по объекту его кода). Но при этом важно выбрать такое методы предобработки, которые бы учитывали специфику изображения.

В этом случае предобработка может заключаться в определенной классификации частей изображения и некоторых преобразований над ними с целью получения таблицы признаковых значений $W = (w_{ij})$ размеров mxn, где $P_1,\ldots,P_n$ –множество признаков, содержащей признаковые описания изображений $I_1,\ldots,I_m$.

При этом, применяя разные способы предобработки, мы можем подучить несколько различных таких таблиц, каждая из которых несет определенную информацию о распознаваемом объекте. Можно считать, что каждая из отдельных таблиц является некоторым «фрагментом» основной матрицы изображения, несущей наибольшую информацию об исходном объекте, которую, естественно, и легче классифицировать.

Таким образом, в этом случае задача распознавания изображения может быть сформулирована как задача составления признаковой таблицы изображений изображения (первый этап) с последующей классификацией изображений (второй этап).

Признаковые таблицы на основе графов

Можно привести многочисленные примеры использования графов для составления признаковых таблиц при распознавании изображений.

Пример 1.

Сама предметная область, к которой относится изображение естественным образом представима в виде графа. Например, сеть автодорог, географические и спутниковые карты местности с реками, дорогами, населенными пунктами и пр. В поисковых системах или системах типа «антиплагиат», где требуется просматривать тексты со схемами и чертежами, также могут возникать анализируемые изображения, основная информация о которых может быть представлена путем выделения группы объектов и указанием на их попарные связи.

В этом случае строки признаковой таблицы соответствуют изображения, представленным графами, а признаками могут служить характеристики графов, имеющие числовые значения: числа вершин, ребер, компонент связности, значения хроматического числа, числа внешней устойчивости и пр.

Пример 2.

Но графы естественным образом сопоставляются и произвольному изображению. Для простоты рассмотрим черно-белый случай.

На этапе «предобработки» с помощью графов можно построить эвристические процедуры, учитывающие особенности соседства черных и белых областей, их кривизну, размеры и т.д.

Рассмотрим плоское изображение I. Его граница $D(I)$ ограничивает область $W(I)$ на плоскости с фиксированной ориентацией. Для простоты берется черно-белый случай. Через $G(I) = G(A,X)$ обозначим граф, сопоставляемый в каждом примере изображению (черно-белой области) I, со множеством вершин A и множеством ребер X. Так как затем графы будут использоваться в эвристических алгоритмах анализа изображений, то отметим следующие факторы.

1. Мы не требуем, чтобы по графу изображение восстанавливалось однозначно.
2. Граф должен по изображению строиться однозначно и отражать определенные свойства изображения.
3. Предполагается наличие масштабируемости. Например, если одно и то же изображение представлено двумя фотографиями разных размеров с одинаковой ориентацией и без искажений по направлениям осей, то алгоритм сопоставления должен дать одинаковые результаты.

Ниже приведены примеры эвристических алгоритмов для сопоставления графов изображениям, которые отвечают вышеприведенным требованиям.

На самом деле, каждый из этих походов содержит вариабельность как по введенным параметрам, так и по некоторым правилам, используемым при построении графа по изображению (соседству вершин, «цвету», площадям подобластей и пр.).

Сеть правильных многоугольников

Фиксируем параметры: k (натуральное число), d и q $(0 < q < 1)$. Первый задает число вершин правильного многоугольника, а второй – его диаметр. Обозначим такой многоугольник через $M(k,d)$. Вписываем $W(I)$ в прямоугольник $P(I)$ с той же самой ориентацией. Строим покрытие $P(I)$ сетью многоугольников $M(k,d)$. При определенных дополнениях к эвристике его можно построить однозначно. Из построенной сети выбрасываем те многоугольники (граничные), площадь пересечения которых с $W(I)$ меньше половины площади $M(k,d)$. Получим сеть $S(I,k,d)$.

Каждой ячейке сопоставим вершину графа $G(A,X)$. Цветом ячейки назовем черный цвет, если часть площади ячейки, им занятая, превосходит q. В противном случае ее цвет белый.

Варьируя параметры и правила построения ребер (соединение соседних ячеек одно цвета, разных цветов и т.п.) получаются графы, в определенной степени характеризующие изображение.

Полученные графы степени не более k являются плоскими. По укладке на плоскости можно построить изображение (с точностью до разницы негатив-позитив).

DOI: 10.21681/2311-3456-2025-5-128-138

По нему, конечно, нельзя восстановить изображение. Но при увеличении k полученное по графу изображение будет все больше приближаться к оригиналу.

Соотношение цветов

Черно белую область $W(I)$ можно представить как объединение черной $W_b(I)$ и белой $W_w(I)$ частей. Каждую из них можно представить как объединение некоторого множества замкнутых областей.

$$W_w(I) = \cup_{i=1}^{t} W_w^i, W_b(I) = \cup_{i=1}^{r} W_b^i.$$

Внутренности таких областей целиком окрашены в один цвет, а увеличение этой границы на любое $\varepsilon > 0$ приводит к появлению внутри «подобласти» другого цвета.

Так как полученные области не обязательно выпуклы, то с помощью одного определенного эвристического алгоритма с использованием известных методов вычислительной геометрии разбиваем их на выпуклые части:

$$W_w^i = \cup_{j=1}^{t_i} W_w^{ij}, W_b(I) = \cup_{j=1}^{r_i} W_b^{ij}.$$

Конечно, это разбиение зависит от выбранного алгоритма. Мы требуем лишь применения одного и того же алгоритма разбиения для всех анализируемых изображений.

Теперь уже элементам выпуклого разбиения ставим в соответствие вершины графа, а правила построения ребер варьируем аналогично предыдущему случаю. Если «огрубить» эвристику разбиения изображения на выпуклые области, допустив наличия в них определенной части другого цвета (в черной области это белый), то можно уменьшить количество вершин в графе, что важно, так как анализируемые ниже алгоритмы практически применимы для графов с несколькими десятками вершин (как правило, до сотни).

Но с помощью графов можно учитывать не только «соседство», но и «размеры».

Каждой области разбиения

$$W_w^i = \cup_{j=1}^{t_i} W_w^{ij}, W_b(I) = \cup_{j=1}^{r_i} W_b^{ij},$$

сопоставляем число $S(i,j,c)$ – площадь области, где c – цвет области. Вводим параметр $0 \le q \le 1$.

Строим вспомогательный граф соседства областей $\Gamma(W)$, вершинам которого соответствуют области, а ребрами соединяются соседние из них.

Если соотношение площадей двух областей (меньшей к большей) не меньше q, то такие области назовем сопоставимыми.

Теперь строим наш основной граф. Его вершинам, как и в предыдущем случае, соответствуют построенные выпуклые области. Две области x и y соединяются ребром в следующем случае:

1. Они сопоставимы.
2. Любой кратчайший путь в графе соседства, соединяющий эти вершины проходит через области меньшей площади, чем площади x и y, причем эти области не сопоставимы ни с x, ни с y.

Полученный граф дает информацию о числах, размерах и взаимному расположению черных и белых областей.

Можно ввести и еще один параметр S^* – минимальную площадь рассматриваемой области и использовать его как своеобразный фильтр, выключающий из рассмотрения небольшие области.

Вопросы применения теоретико-графовых конструкций для визуализации и распознавания объектов изучались во многих работах с разных точек зрения. В качестве примеров можно привести статьи [1]–[5].

Признаковое распознавание и графы

Если с первым примером все понятно: там значения признаков – числа, то во втором случае мы имеем более сложную ситуацию.

В отличие от первого примера, здесь каждому признаку изображения соответствует **алгоритм** построения соответствующего графа. А значением признака – **построенный граф**. Поэтому встает задача установления эквивалентности значения того или иного признака у разных изображений.

Здесь признаки – тоже объекты и должны быть их описания: $I(p_1),\ldots,I(p_m)$.

Но можно предположить, что это более примитивные объекты по сравнению с объектом B, то есть описания $I(p_1),\ldots,I(p_m)$ мы имеем, а описания самого $I(B)$ пока нет. При этом в такой схеме естественным выглядит следующий вопрос. Нам не нужно знать $I(B)$, целью является нахождение других его признаков: $I(p_m+1),\ldots,I(p_n)$.

Если в качестве объекта B этот граф рассматривать по приведенной выше схеме, а в качестве признаков взять набор инвариантов графа: функций $p_1(B),\ldots,p_m(B)$, то в качестве цели поиска может быть некоторый другой его инвариант: число ребер в этом графе.

Эта схема уже отличается от процедуры распознавания объекта по набору значений его признаков. Информацию об объекте мы черпаем не только из типа признака и его значения. Признаки могут быть взаимосвязаны. Если мы знаем, что один из них p_i есть функция другого $p_i = f(p_j)$, то значения этой функции могут дать дополнительную информацию.

Усложнение постановки можно продолжить. Но мы ограничимся только этими двумя: значение признаков – числа; значения признаков – графы, а сами эти графы-признаки на следующем шаге описываются набором своих признаков – чисел.

В любом случае – это поле для применения результатов теории распознавания.

Представленная так задача является классической в теории распознавания. Фундаментальный подход к решению задач распознавания на основе эвристических алгоритмов предложен Ю. И. Журавлевым.

Использование теоретикографовых построений для представления специфики задач распознавая в области специального рода химических проблем дано, например, в работах К. В. Рудакова и И. Ю. Торшина [6]–[8].

Вопрос использования графов для представления изображений изучался во многих работах с разных точек зрения. См., например, работы [9–10].

В нашей работе большое внимание уделено методам на основе использования инвариантов графа.

В частности, прямое использование инвариантов графа при распознавании объектов описывается в упомянутой выше работе [8].

Актуальность и применимость проблематики построения и анализа инвариантов графа обсуждается, в частности, в работах [11-12].

Возникает вопрос: что такое граф и что такое представление графа? На самом деле, мы, по-видимому, никогда не имеем дело с самим графом, а всегда – только с его «реализаций» или «представлением». Очевидно, что один и тот же объект по-разному можно представить или реализовать.

И вот на этом «стыке» между теорией графов и распознаванием изображений возникает тема, обсуждаемая во втором параграфе. Проблема распознавания графов, известная в теории графов уже более полувека, по-видимому, имеет ограниченное прикладное применение к распознаванию изображений.

С распознаванием графов тесна связано понятия инварианта. Оно появляется при попытке ответить на вопрос: два имеющихся различных представления соответствуют одному объекту или разным?

Во третьем параграфе статьи строятся несколько типов инвариантов графов и обсуждается их возможное применение в эвристических алгоритмах вышеупомянутых прикладных проблем.

Проблема распознавания графов и ее прикладное значение для распознавания изображений

Изображение I представлено графом $G(I)$. И мы теперь имеем дело именно с $G(I)$. В теории графов взаимосвязаны два понятия: «распознавание» и «восстановление».

Введем несколько необходимых определений и сделаем ряд замечаний.

Всюду рассматривается простой неориентированный граф $G = (X,U)$, имеющий n вершин и k ребер. Пусть $X = \{A_1,\ldots,A_n\}$, $U=\{r_1,\ldots,r_k\}$. Пусть $t(A_i)$ – степень вершины A_i в графе G.

Граф с n вершинами и k ребрами можно задать списками ребер и вершин. В этом случае длина его кода (входного слова алгоритма, размера структуры данных) лежит между $4n+10k$ и $4n+10k+(n+2k)[lgn]$. Если граф задается списками соседей его вершин, то длина входа лежит между $2n+8m$ и $2n+8m+2k[lgn]$. Порядок же матрицы инцидентности графа равен n^2-n+1. Таким образом, под эффективным алгоритмом понимается процедура, сложность которой полиномиально зависит от n.

Это, безусловно, не значит, что вычислительные процедуры более высокие сложности неприменимы, но практическое использования подобных методов налагает естественные ограничение на величину n. При нынешнем уровне технологий, например, использование точных алгоритмов экспоненциальной сложности ограничивается графами с десятками или, в лучшем случае, сотнями вершин. См., например, [13].

Это ограничение обуславливает применение вместо точных алгоритмов эвристических.

Определение. Функция, определенная на множестве всех n-вершинных графов и принимающая одно и то же значение для изоморфных графов называется *инвариантом графа*.

Таким образом, инвариант $f(G)$ есть функция, которая может быть вычислена на любой реализации графа G. Она обладает тем свойством, что если между парой графов G_1 и G_2 есть отношение изоморфизма, т.е. $G_1 \sim G_2$, то $f(G_1) = f(G_2)$.

Проблема *восстановления* в теории графов связывается обычно с гипотезой Улама, хотя, безусловно, имеет значительно более широкие аспекты.

Определение. Подграф G_i, подученный из G выбрасыванием i-й вершины со всеми инцидентными ей ребрами, называется примарным.

Гипотеза Улама. Пусть G – n-вершинный неориентированный граф. Если заданы классы изоморфизма всех n примарных подграфов графа G, то при $n \geq 3$ класс изоморфизма графа G определяется однозначно.

В настоящее время гипотеза Улама не доказана и не опровергнута, так что она так и остается гипотезой.

Определение. Свойство T графа G называется *восстанавливаемым* если его можно выявить (решив задачу в форме распознавания, см. [13]), рассматривая все примарные подграфы графа G.

Граф G со свойством T называется *распознаваемым*.

Граф G называется *восстанавливаемым*, если восстанавливаем его класс изоморфизма.

DOI: 10.21681/2311-3456-2025-5-128-138

Как уже было сказано выше, особо актуальна проблема восстановления и распознаваемости в теории графов.

В качестве примеров приведем следующие утверждения.

Утверждение 1. Если G – однородный граф, то он восстанавливаем.

Утверждение 2. Деревья являются распознаваемыми графами.

Это утверждение следует из того факта, что свойства связности графа и свойство отсутствия циклов является восстанавливаемыми. Так как дерево вполне характеризуется этими свойствами, то из упомянутого выше утверждения и следует Утверждение 2.

Утверждение 3. Двусвязные графы распознаваемы.

Утверждение 4. Однородные графы распознаваемы.

Одним из общих результатов, относящихся к восстанавливаемым характеристикам, является лемма Келли. Пусть H и K – графы. Обозначим через $\gamma(H,K)$ – число подграфов графа H, изоморфных графу K.

Лемма Келли. Если мощности множеств вершин графов H и K разные, то $\gamma(H,K)$ – восстанавливаемая характеристика графа G.

Из леммы Келли, среди прочих, можно вывести восстанавливаемость следующих графов и их характеристик: несвязные графа восстанавливаемы; дихромат графа является его восстанавливаемой характеристикой; число гамильтоновых циклов в графе является его восстанавливаемой характеристикой; хроматический многочлен графа является его восстанавливаемой характеристикой; характеристический многочлен графа является его восстанавливаемой характеристикой.

Ряд аналогичных результатов получен и в близкой проблеме реберного восстановления графов.

Уже видно, что восстанавливаемость оперирует с характеристиками графов, для нахождения которых в настоящее время нет полиномиальных алгоритмов. См. [13].

Дело в том, что техника доказательства подобных результатов базируется на смысле определения понятия «восстановление» и обычно проходит по следующей схеме.

1. Распознаваемость графа означает наличие у него восстанавливаемого свойства. Доказываемое свойство фиксируется. Берется конкретный граф G. Способ его задания не описывается. Предполагается, что каким бы этот способ не был, с его помощью строятся все n примарные подграфы графа G. Способ задания или описания таким образом построенных графов также не фиксируется.
2. Следующий шаг – перебор по примарным подграфам. Для каждого такого подграфа G_i в таком переборе считается известным («алгоритмически доступным») весь класс его изоморфизма $K(G_i)$. С теоретической точки естественно, что можно перебрать все изоморфные графы из $K(G_i)$, но без неизвестного на данный момент эффективного алгоритма такого перебора данная процедура не может быть представлена эффективным алгоритмом.
3. При этом переборе уже рассматриваются конкретные графы из и для них анализируется выполнимость восстанавливаемого свойства $K(G_i)$. При этом по каждому такому конкретному графу строится один или несколько конкретных графов из $K(G)$, для которых также устанавливается выполнимость восстанавливаемого свойства.
4. В завершение доказывается корректность перебора. (Как правило, эта часть опускается в силу очевидности вопроса.)

Отсюда следует, что включение в практический алгоритм распознавания изображений, основанный на представлении изображений графами, процедура распознавания графов приводит к тому, что сложность алгоритма не может быть меньше $f(n)$, где $f(n)$ – сложность решения задачи изоморфизма для n-вершинных графов.

Переход к примарным графам приводит к тому, что техника доказательства утверждений в области распознавания графов широко использует метод, который может лечь в основу эвристики для распознавания изображений. Это метод декомпозиции. Он, в каком-то смысле, позволяет обойти ограничение на рост параметра n. Речь идет об очевидной простой схеме.

Пусть в нашем распоряжении есть процедура решения задачи изоморфизма сложности $f(n)$, которая применима для $n < k$, и мы можем себе позволить потратить время $sf(k)$ на решение задачи распознавания графа.

Алгоритм 1. Исходный граф G представляется в виде совокупности графов $G^1_{p_i}$, $i = 1,..,s_1$. В свою очередь, каждый из этих графов представляется совокупностью $G^2_{p_i p_j}, j = 1,..,s_{1i}$. И так далее.

1. Число уровней иерархии l определяется на основе параметров s, $f(k)$ и способа представления.
2. Способ преставления (декомпозиции) зависит от специфики изображения или особенностей задания этого изображения в виде графа. Обращаем внимание, что с прикладной точки зрения, напрашиваются следующие подходы: разбиение на примарные подграфы; физическое разбиение изображения на части (см. примеры из предыдущего раздела), что влечет актуальность таких

характеристик графа, как планарность, хроматическое число, наличие и расположение «мостов», и т.д.

3. Затем организуется процедура восстановления графа по частям, начиная с графов уровня l, и заканчивая графом G.

Другой взгляд на описанное здесь алгоритмическое противоречие между понятием распознавание графа и требованием к эффективности практического алгоритма распознавания изображения, связан с использование инвариантов графа.

Инварианты графа и проблема распознавания

Инварианты и возможности их использования

Примерами широко известных инвариантов являются, так называемые, основные числа теории графов: хроматическое число, число внутренней устойчивости (число независимости), число внешней устойчивости, плотность графа, кликоматическое число, толщина и т.д. Хорошо известными инвариантными свойствами графа являются: гамильтоновость, связность, двудольность, планарность и т.д.

По существу, инвариант – это функция, не зависящая от нумерации вершин графа и потому довольно естественно, что в реальных ситуациях функции на графе являются инвариантами.

Ясно, что многие инварианты графа не являются независимыми между собой и связаны разного рода соотношениями. Например, число независимости $\alpha(G)$ и плотность $\varphi(G)$ графа G связаны между собой сведущим равенством: $\alpha(G) = \varphi(\bar{G})$, где $\bar{G}$ – дополнение графа G. Или хорошо известно неравенство: $\alpha(G)\gamma(G) \geq n$, $\gamma(G)$ – хроматическое число n-вершинного графа G.

Таким образом, имеется очень много различных инвариантов графов. В связи с этим возникает естественная проблема порождения или описания инвариантов в терминах базисных элементов или каким-либо другим способом.

Определение. Система инвариантов $\{f_1,\ldots,f_s\}$ называется полной, если выполнены два условия: из того, что два графа изоморфны, следует, что их полные системы инвариантов совпадают, а для пары неизоморфных графов они не совпадают.

Ясно, что если мы все n вершинные графы разобьем на классы изоморфизма $T_1,\ldots,T_N$, то номер класса изоморфизма является инвариантом графа, через который может быть получен (в принципе) любой другой инвариант, т.к., зная номер класса изоморфизма, мы можем взять любой граф из этого класса и на нем «автоматически» вычислить любой инвариант.

Однако этот «полный инвариант» не является практически удовлетворительным в силу сложности его нахождения. С другой стороны, «обычный» инвариант графа должен вычисляться по любому представлению графа и для его нахождения вовсе не надо иметь таблицу классов изоморфизма всех n-вершинных графов.

В связи со всем сказанным выше возникает два вопроса.

1. Что такое *базисный* и что такое *составной (сложный)* инвариант графа?
2. Что такое *естественный* инвариант графа?

Например, число ребер графа можно считать базисным инвариантом. А такие функции, такие как основные числа графа, связность, числа Бетти и т.д. являются естественными инвариантами графа. В то же время такой инвариант, как номер класса изоморфизма вряд ли можно признать естественным.

Вернемся к практической задаче. Во введении мы сформулировали два взгляда на эту задачу. С точки зрения первого примера, изображению сопоставлен единственный граф, а в признаковой таблице его признаками являются числовые значения характеристик (например, инвариантов) этого графа. Во втором случае значением признака является сам граф, но для установления совпадения значений признака у разных изображений приходим к задаче изоморфизма графов, которая, в свою очередь может быть сведена к анализу инвариантов рассматриваемых графов-признаков.

Оба подхода содержат следующую схему.

Имеется несколько изображений одного и того же объекта $I_1,\ldots,I_m$, полученных с помощью средств $T_1,\ldots,T_m$. В нашем распоряжении есть набор инвариантов $f_1,\ldots,f_s$. (Они, в частности, могут рассматриваться как признаки.) Если их значения на графах $G_1,\ldots,G_m$ могут быть вычислены с помощью эффективного алгоритма, то они практически применимы в эвристической процедуре, основанной на признаковой таблице, элементами которой являются они сами $a_{ij} = f_j(G_i)$ или некоторые функции от них $a_{ij} = F(f_j(G_i))$.

Далее используется предположение, что из неравенства инвариантов следует, что графы не изоморфны, а соответствующие им изображения различны. Практическое применение этого эвристического подхода мы здесь не обсуждаем, а лишь подчеркиваем три его очевидных требования:

1. Наличие эффективной процедуры вычисления инварианта.
2. Необходимость учета специфики задачи для выбора инвариантов.
3. Использование как можно более широкого и разнообразного набора инвариантов.

В качестве примера приведем использование подхода к распознаванию деревьев. На сегодняшний день одним из самых интересных результатов

в этой области является теорема Смоленского-Зарецкого. Ее можно найти, например, в [14].

Для любой пары висячих вершин дерева единственным образом определяется расстояние между ними. Пусть $D(T)$ – набор таких расстояний (с их кратностями) для дерева T. Согласно этой теореме, изоморфные деревья T и $T`$ обязательно имеют одинаковые наборы $D(T)$ и $D(T`)$. Но совпадение $D(T)$ и $D(T`)$ еще не означает изоморфизма деревьев. Нужно еще так проиндексировать висячие вершины обоих деревьев, чтобы совпадали расстояния, чтобы совпадали расстояния между парами одинаково проиндексированных вершин.

А число висячих вершин в дереве уже может быть другого порядка, чем n. Тем самым, если за счет учета специфики задачи можно ограничиться деревьями, то, например, при числе висячих вершин $Olog\ n$) выполняется и первое из приведенных выше требований: наличие эффективного алгоритма проверки изоморфизма.

В следующем разделе мы дадим примеры новых типов инвариантов графов, которые могут быть использованы в описанном подходе.

Примеры новых типов инвариантов

Выше мы говорили, что инварианты можно рассматривать в некоторой иерархии: от простых – базисных, к составным – *сложным*.

В трех ниже приведенных примерах описаны новые, на наш взгляд, практически применимые, типы инвариантов. Они зависят от параметра, значения которого ограничивают их практическую применимость.

Один из самых известных способов задания нумерованного графа – перечень соседей его вершин. Будем называть это множество соседей вершины v множеством вершин, порожденных вершиной v. Аналогично, ребра между v и этими вершинами назовем множеством ребер, порожденным v.

Обозначим через Δ_s^m – число совокупностей из s ребер графа, которые порождают m вершин $(s = 1,...,k;\ m = 2,...,n)$. Этот набор чисел, рассматриваемый как множество, уже не зависит от нумерации вершин графа, поэтому является инвариантом (системой инвариантов) графа.

Из этих инвариантов, как из составных частей, ниже мы построим более сложные.

Теперь рассмотрим еще один пример множества базисных инвариантов. В k-реберном графе можно рассмотреть C_k^s совокупностей, состоящих из s ребер. Они порождают подграфы с разным числом вершин и компонент связности.

Если мы теперь обозначим через $\delta_{m,t}^s$ – число совокупностей из s ребер графа G, образующих t – вершинные подграфы с m компонентами связности, то получим еще один инвариант (систему инвариантов) графа.

Введем еще один инвариант: набор «обобщенных» степеней вершин графа.

Обозначим через γ_s^m число совокупностей из s вершин графа G, которые порождают m вершинные подграфы G. Другими словами, подграф, порожденный вершинами $\{A_{i_1},A_{i_2},...,A_{i_s}\}$ входят сами эти вершины плюс те вершины, которые смежные по крайней мере с одной из входящих в множество $\{A_{i_1},A_{i_2},...,A_{i_s}\}$ вершиной.

Все эти примеры носят «локальный» характер:

1. Используются в их определениях только два понятия: что такое вершина графа, что такое ребро графа.
2. Число-инвариант рассчитывается по одному конкретному объекту: подмножеству вершин или ребер графа.

Теперь на их основе будем строить более сложные инварианты. Эти примеры будут касаться хорошо известных инвариантов, но в качестве следствий из полученных конструкций мы сможем получить новые свойства известных инвариантов.

Использование этих инвариантов в алгоритмах распознавания требует ограничения на их количество.

Поэтому при рассмотрении инвариантов типа Δ_s^m берется случай, когда s и m являются константами. В такой ситуации числа Δ_s^m полиноминально вычислимы.

При рассмотрении инвариантов типа $\delta_{m,t}^s$ берется случай, когда s,t и m являются константами. В такой ситуации числа $\delta_{m,t}^s$ полиноминально вычислимы.

При рассмотрении инвариантов типа γ_s^m берется случай, когда s и m являются константами. В такой ситуации числа γ_s^m полиноминально вычислимы.

Как выше было уже сказано, при декомпозиционных и иерархических представлениях естественно использовать хроматические характеристики графа.

Физическое разбиение изображение при декомпозиционном подходе естественно привлекает внимание к минимизации связей между частями. Некоторые классы изображений имеют определенные ограничения на «схожесть» частей, например, эта «схожесть» различна в случае пейзажа и фото геометрической фигуры, и т.п. Здесь актуальны другие характеристики графа, например, связанные с числом его внешней или внутренней устойчивости.

Конечно, вычисление хроматического числа и чисел его внешней или внутренней устойчивости в настоящее время не может быть осуществлено эффективно. Поэтому нижеприведенное рассмотрение дано для пояснения смысла введенных инвариантов и аргументации их практического применения.

Некоторые элементы техники доказательств приведенных ниже утверждений можно найти в [15].

Пример 3.

Сначала рассмотрим такой инвариант как число независимости графа $\alpha(G)$.

Напомним, что множество вершин графа называется *независимым* (*внутренне устойчивым*), если никакие две вершины в этом множестве не являются смежными. А число вершин в наибольшем по мощности независимом множестве графа называется *числом независимости* (*числом внутренней устойчивости*).

Пусть $G = (X,U)$ n-вершинный неориентированный граф, имеющий k-ребер: $r_1,\ldots,r_k$. Обозначим через Δ_s^m – число совокупностей из s ребер графа, которые порождают m вершин ($s = 1,\ldots,k$; $m = 2,\ldots,n$). Ясно, что если R – некоторая совокупность из s ребер, то максимальная мощность множества вершин, которое она порождает равно $2s$, а минимальное – это $s + 1$. В частности, если граф представляют собой «звезду», имеющая k вершин и $k - 1$ ребро, то и при $m \neq s+1$ справедливо соотношение $\Delta_s^m = 0$, а $\Delta_s^{s+1} = C_{k-1}^s$.

Обозначим через $L(n,r)$ число независимых подмножеств графа G мощности r. Пусть – V_i – множество подграфов графа мощности r, содержащих ребро r_i. Тогда множество

$$V = \bigcup_{i=1}^{k} V_i$$

содержит все подграфы графа G, так как каждый из таких подграфом содержит хотя бы одно ребро графа G.

Отсюда следует соотношение: $L(n,r) = C_n^r - |V|$. Но из формулы включения-исключения имеем равенство:

$$|V| = \sum_{i=1}^{k} |V_i| - \sum_{i<j} |V_i \cap V_j| + \ldots$$

$$\ldots + (-1)^s \sum_{i_1<i_2<\ldots<i_s} |V_{i_1} \cap V_{i_2} \cap \ldots \cap V_{i_s}| + \ldots$$

$$\ldots + (-1)^k |V_{i_1} \cap V_{i_2} \cap \ldots \cap V_{i_k}|.$$

Пусть теперь $\delta(i_1,\ldots,i_s)$ – число вершин, «порождаемых» множеством ребер $(r_{i_1},r_{i_2},\ldots,r_{i_s})$, т.е. инцидентных этим ребрам, а, если вершина инцидентна нескольким ребрам, то в таком подсчете она учитывается только один раз.

С учетом этих обозначений имеем соотношения:

$$|V| = C_{n-2}^{r-2}, \; |V_i \cap V_j| = C_{n-\delta(i,j)}^{r-\delta(i,j)}, \; \sum_{i<j} |V_i \cap V_j| =$$

$$= \sum_{i<j} C_{n-\delta(i,j)}^{r-\delta(i,j)} = \sum_{t=3}^{4} \Delta_2^t C_{n-t}^{r-t}, \; \ldots, \; \sum_{i_1<i_2<\ldots<i_s} |V_{i_1} \cap V_{i_2} \cap \ldots \cap V_{i_s}| =$$

$$= \sum_{i<j} C_{n-\delta(i_1,\ldots,i_s)}^{r-\delta(i_1,\ldots,i_s)} = \sum_{t=s+1}^{2s} \Delta_s^t C_{n-t}^{r-t}.$$

Пусть теперь среди вершин графа есть q изолированных вершин. Так как $\delta(1,2,\ldots,k)$ – это число вершин графа G, которые порождаются всеми его ребрами, то $\delta(1,2,\ldots,k) = n - q$. Тогда с учетом вышеприведенных соотношений получаем выражение:

$$L(n,r) = C_{2s}^r - C_k^1 C_{n-2}^{r-2} + \sum_{t=3}^{4} \Delta_2^t C_{n-t}^{r-t} + \ldots$$
$$\ldots + (-1)^s \sum_{t=s+1} \Delta_s^t C_{n-t}^{r-t} + \ldots + (-1)^k C_q^{r+q-n}.$$

Таким образом, мы получили выражение для $L(n,r)$ через систему инвариантов $\{\Delta_s^m\}$. Теперь заметим, что если $L(n,r) = 0$ для некоторого r, то в графе нет независимых множеств мощности r.

Отсюда следует, что $\alpha(G) \leq r$. Таким образом получается, что число независимости графа есть минимальный натуральный корень полинома $L(n,r)$.

Мы доказали следующую теорему.

Теорема 1. Если r_0 – минимальный натуральный корень полинома $L(n,r)$, то $\alpha(G) = r_0$.

Пример 4.

Теперь рассмотрим другой инвариант – хроматическое число графа. Напомним, что h-раскраской графа называется отображение f множества его вершин на множество $\{1,\ldots,h\}$. Раскраска называется правильной, если для $u \neq v$ справедливо соотношение: $f(u) \neq f(v)$. В этом случае граф называется h-раскрашиваемым. Хроматическим числом графа G называется минимальное натуральное число h, при котором граф G является h-раскрашиваемым.

Пусть, как обычно, граф $G = (X,U)$ имеет n вершин и k ребер.

Обозначим через $\Phi(n,r)$ число правильных раскрасок графа G в r-цветов.

Как и в предыдущем случае через V_i обозначим множество раскрасок графа G r цветов, при которых ребро r_i, раскрашено неправильно. Тогда множество

$$V = \bigcup_{i=1}^{k} V_i$$

содержит все неправильные раскраски графа G в r цветов. Ясно, что $\Phi(n,r) = r^n - |V|$. Вновь, используя формулу включения-исключения, получаем

$$|V| = \sum_{i=1}^{k} |V_i| - \sum_{i<j} |V_i \cap V_j| + \ldots$$

$$\ldots + (-1)^s \sum_{i_1<i_2<\ldots<i_s} |V_{i_1} \cap V_{i_2} \cap \ldots \cap V_{i_s}| + \ldots$$

$$\ldots + (-1)^k |V_{i_1} \cap V_{i_2} \cap \ldots \cap V_{i_k}|.$$

Далее имеем $|V_i| = r \cdot r^{n-2} = r^{n-1}$, $|V_i \cap V_j| = r^{n-2}$.

Эти равенства следуют из того факта, что в случае, когда ребра r_i и r_j не являются инцидентными, то каждое из множеств вершин, порождаемых этими ребрами, можно независимо красить в один цвет, а остальные $n - 4$ вершины красятся произвольно. Таким образом общее число раскрасок в этом варианте равно $r^2 \cdot r^{n-4} = r^{n-2}$.

Если же рассматриваемые ребра образуют цепь, то все ее вершины должны быть окрашены одинаково

и число таких раскрасок равно r. Оставшиеся же $n - 3$ вершины красятся произвольно. Поэтому общее число раскрасок в этом случае равно $r \cdot r^{n-3} = r^{n-2}$.

В общем случае величина $|V_{i_1} \cap V_{i_2} \cap \ldots \cap V_{i_s}|$ вычисляется следующим образом. Если совокупность ребер $(r_{i_1}, r_{i_2}, \ldots, r_{i_s})$ образует подграф с t вершинами и m компонентами связности, то число раскрасок, при которых каждое из ребер этой совокупности будет раскрашено неправильно равно следующей величине:

$$r^m \cdot r^{n-t} = r^{n+m-t}.$$

Доказательство этого утверждения опирается на следующую лемму.

Лемма. Если все ребра связного подграфа окрашены неправильно, то все они окрашены в один цвет.

Если мы теперь обозначим через $\delta^s_{m,t}$ – число совокупностей из s ребер графа G, образующих t – вершинные подграфы с m компонентами связности, то получим, используя вышеприведенные формулы, следующее соотношение.

$$\sum_{i_1<i_2<\ldots<i_s} |V_{i_1} \cap V_{i_2} \cap \ldots \cap V_{i_s}| = \sum_{m,t} \delta^s_{m,t} r^{m+n-t}.$$

Отсюда окончательно получаем формулу

$$\Phi(n,r) = r^n - kr^{n-1} + C_k^2 r^{n-2} - \ldots$$
$$\ldots + (-1)^s \sum_{m,t} \delta^s_{m,t} r^{m+n-t} + \ldots + (-1)^k r^{n+d-k}.$$

В ней через d обозначено число компонент связности графа G.

Таким образом функция $\Phi(n,r)$ выражается через систему инвариантов $\{\delta^s_{m,t}\}$.

Заметим далее, что, если $\Phi(n,r) = 0$, то для хроматического числа $\gamma(G)$ графа G имеем оценку: $\gamma(G) > r$.

В общем случае справедливо следующее утверждение.

Теорема 2. Если $\Phi(n,1) = \Phi(n,2) = \ldots = \Phi(n,p-1) = 0$, а $\Phi(n,p) \neq 0$, то $\gamma(G) = p$.

Пример 5.

Дальнейшее рассмотрение относится к числу внешней устойчивости графа. Напомним, что число внешней устойчивости графа – это наименьшая мощность множества его вершин такого, что любая вершина графа смежная хотя бы с одной вершиной этого множества.

Пусть вновь граф $G = (X,U)$ имеет n вершин и k ребер. $X = \{A_1, \ldots, A_n\}$. Обозначим через $F(n,r)$ число внешне-устойчивых множеств этого графа, состоящих из r вершин. Число внешней устойчивости графа G обозначим через $\beta(G)$.

По аналогии с предыдущими случаями обозначим через V_i совокупность подмножеств вершин мощности r графа G, которые не покрывают вершину A_i. Или, другими словами, в V_i не входят подмножества, содержащие смежные с A_i вершины. Пусть $t(A_i)$ – степень вершины A_i в графе G.

Тогда справедливо соотношение:

$$|V_i| = C^{r-t(A_i)-1}_{n-t(A_i)-1}.$$

Далее, пусть число вершин в подграфе, порожденных вершинами A_i и A_j, т.е. число вершин, каждая из которых смежная либо с A_i, либо с A_j. Тогда

$$|V_i \cap V_j| = C^{r-\gamma(i,j)}_{n-\gamma(i,j)}.$$

Аналогично определим функцию $\gamma(A_{i_1}, A_{i_2}, \ldots, A_{i_s})$ как число вершин в подграфе, порожденном вершинами $\{A_{i_1}, A_{i_2}, \ldots, A_{i_s}\}$.

Тогда справедливо равенство:

$$|V_{i_1} \cap V_{i_2} \cap \ldots \cap V_{i_s}| = C^{r-\gamma(A_{i_1},\ldots,A_{i_s})}_{n-\gamma(A_{i_1},\ldots,A_{i_s})}.$$

Далее заметим, что

$$F(n,r) = C^r_n - \left|\bigcup_{i=1}^n V_i\right|.$$

Учитывая изложенное выше, имеем

$$F(n,r) = C^r_n - \sum_{i<j} |V_i \cap V_j| + \ldots$$
$$\ldots + (-1)^s \sum_{i_1<i_2<\ldots<i_s} |V_{i_1} \cap V_{i_2} \cap \ldots \cap V_{i_s}| + \ldots$$
$$\ldots + (-1)^n |V_{i_1} \cap V_{i_2} \cap \ldots \cap V_{i_n}| = \sum_{m=1}^n \gamma^m_1 C^{r-m}_{n-m} -$$
$$- \sum_{m=1}^n \gamma^m_2 C^{r-m}_{n-m} + \ldots + (-1)^s \sum_{m=1}^n \gamma^m_s C^{s-m}_{n-m}.$$

Здесь через γ^m_s обозначено число совокупностей из s вершин графа G, который порождают m вершинные подграфы G. Другими словами, подграф, порожденный вершинами $\{A_{i_1}, A_{i_2}, \ldots, A_{i_s}\}$ входят сами эти вершины плюс те вершины, которые смежные по крайней мере с одной из входящих в множество $\{A_{i_1}, A_{i_2}, \ldots, A_{i_s}\}$ вершиной.

Обозначим через t_{r-} число вершин степени r в графе G. С учетом этого обозначения, заметим, в частности, что $\gamma^m_1 = t_{m-1}$.

Применяя ту же технику, что и в предыдущих случаях, окончательно получаем следующее соотношение.

$$F(n,r) = C^r_n - \sum_{m=1}^n t_{m-1} C^{r-m}_{n-m} +$$
$$+ \sum_{m=1}^n \gamma^m_2 C^{r-m}_{n-m} - \ldots + (-1)^s \sum_m \gamma^m_s C^{s-m}_{n-m}.$$

Если мы определим понятие обобщенной степени для заданной совокупности вершин как множество вершин графа G, смежных хотя бы с одной из входящих в заданную совокупность вершиной, то γ^m_s будет означать число s-совокупностей с обобщенной степенью m.

Теорема 3. Если $F(n,1) = F(n,2) = \ldots = F(n,p-1) = 0$, а $F(n,p) \neq 0$, то $\beta(G) = p$.

Таким образом инвариант графа $\beta(G)$ выражается через систему инвариантов $\{\gamma_s^m\}$ являющихся обобщенными степенями совокупностей вершин.

В свете введенного определения «обобщенной степени», нетрудно заметить, что система инвариантов $\{\Delta_s^m\}$ представляет собой обобщенные степени совокупностей ребер графа G.

Заключительные замечания

Теоремы 1–3 представляют интерес в области теории графов, посвященной построению и анализу различных инвариантов графа.

Алгоритмы, построенные на их основе, в общем случае являются переборными. Поэтому в этом случае они могут быть использованы либо для графов с небольшим числом вершин, либо на основе декомпозиционного эвристического подхода, описанного во втором параграфе.

Однако пример с теоремой Смоленского-Зарецкого говорит о том, что на специальных классах графов трудоемкость алгоритмов может быть иная.

В заключение сделаем одно замечание. При распознавании на основе признаков важно иметь «качественные» (эффективно различающие) признаки. См., например, [1].

Для случая инвариантов, которые могут принимать одинаковые значения и для неизоморфных графов, можно считать, что инвариант I разбивает все множество нумерованных n-вершинных графов на $k(n)$ подмножеств $W_1,\ldots W_{k(n)}$, на элементах каждого из которых он принимает одинаковые значения. В свою очередь, каждое из этих подмножеств W_j разбивается на $l(j)$ частей: W_{ju}, состоящих из изоморфных графов.

В этой связи качество (силу) инварианта можно, например, задать следующим образом:

$$\xi_{I(n)} = \max_{j=1,\ldots,k(n)} \max_{i=1,\ldots,l(j)} \frac{|W_{ij}|}{|W_j|}.$$

То есть, чем «сильнее» инвариант, тем больше доля изоморфных графов среди всех графов, на которых он принимает одинаковое значение.

Так как в классической задаче распознавания большую роль играет «вес» признака, то «сила инварианта» также может быть при этом использована.

Она, в частности связана с вопросом, что несет больше информации о графе G: хроматическое число $\gamma(G)$ или число независимости $\alpha(G)$?

Конечно, если бы на множестве n-вершинных графов G_n было, например, задано равномерное распределение и мы бы знали $p(\gamma(G))$ – вероятность того, что значение инварианта (например, хроматического числа) некоторого графа G равно $\gamma(G)$, тогда энтропия

$$H(n) = \sum_{G \in G_n} p_{\gamma(G)} \log \frac{1}{p_{\gamma(G)}}$$

и может названа «информацией» о графе по этому инварианту.

Литература

1. Каркищенко А. Н., Мнухин В. Б. Метод детекции характерных точек изображения с помощью знакового представления // Известия ЮФУ. Технические науки, 2020. Том 214. № 4, стр. 59–70.
2. Баженов А. В., Филякин А. А. Теорема графов как основа построения систем связи // Universum: технические науки: электрон. научн. журнал, 2022. № 3(96).
3. Акбашева Е. А., Акбашева Г. А., Тлупов И. З. Методы представления текстовых документов на основе графов в задачах обработки естественного языка // Информатика, вычислительная техника и управление. Серия: Естественные и технические науки. 2022. № 11, стр. 67–72.
4. Степкина А. В., Степкина А. С. Алгоритмы распознавания простых графов коллективным агентом // Компьютерные исследования и моделирование, 2021., том 13, № 1. С.33–45. DOI: 10.20537/2076-7633-2021-13-1-33-45.
5. Nagavarapu S. C., Vachhani L., Sinha A. et al. Generalizing Multi-agent Graph Exploration Techniques // International Journal of Control, Automation and Systems. 2020. Vol. 19. P. 491–504. https://doi.org/10.1007/s12555-019-0067-8.
6. Torshin I. Yu., Rudakov K. V. Topological Chemograph Analysis Theory as a Promising Approach to Simulation Modeling of Quantum-Mechanical Properties of Molecules. Part II: Quantum-Chemical Interpretations of Chemograph TheoryPattern // Recognition and Image Analysis. 2022. Vol. 22. P. 205–217.
7. Torshin I. Yu., Rudakov K. V. Topological Chemograph Analysis Theory As a Promising Approach to Simulation Modeling of Quantum-Mechanical Properties of Molecules. Part I: Quantum-Chemical Interpretations of Chemograph TheoryPattern // Recognition and Image Analysis. 2021. Vol. 21. P. 800–810.
8. Torshin I. Yu., Rudakov K. V. Local completeness of the ‘chemographs’ invariants in view of the combinatorial theory of solvability // Pattern Recognition and Image Analysis. 2014. Vol. 24. P. 196–208.
9. Абгалдаева А. А., Пушкин А. Ю. Применение теории графов в сфере информационных технологий // Universum: технические науки: электрон. научн. журн. 2023. № 2(107). URL: https://7universum.com/ru/tech/archive/item/15061.
10. Сапунов С. В., Сенченко А. С. Лингвистическое представление графов с помеченными вершинами // Доповіді Національної академії наук України. 2019. № 11. С. 17–24.
11. Курапов С. В., Давидовский М. В. Вычислительные методы определения инвариантов графа///International Journal of Open Information Technologies ISSN: 2307-8162. 2021. Vol. 9, № 2. С. 1–8.
12. Тутыгин Р. А., Зяблицева Л. В. Эффективность инвариантов графов, соответствующих полугруппам // Сб трудов конференции: Материалы VI Международной научно-практической конференции (школы-семинара) молодых ученых. Тольятти, 2020. С. 114–117.
13. Гэри М., Джонсон Д. Вычислительные машины и труднорешаемые задачи // М.: Мир, 2012.
14. Zykov A. A., Graphs Theory (Nauka, Moscow, 1986) [in Russian].
15. Леонтьев В. К. Комбинаторика и информация. Часть 1. Комбинаторный анализ. М.: МФТИ, 2015. 174 с.

ON THE USE OF GRAPH THEORY IN THE CLASSIFICATION INFORMATION

Gordeev E. N.[3], Leontiev V. K.[4]

Keywords: *recognition, feature tables, heuristics, reconstruction, graph invariant, chromatic number, independence number, number of external stabilities, number of internal stabilities.*

The purpose of this work *is to analyze the possibilities of applying graph theory for image coding and classification, which is especially relevant in connection with the use of artificial intelligence methods for image classification.*

Research method: *combinatorics and graph theory, as well as heuristic algorithms.*

Results: *The paper discusses the possibility of applying the classical results of graph theory concerning the problems of graph recovery and recognition and their characteristics in the field of image recognition. At the same time, various aspects of the problem of describing (representing) graphs using their invariants are analyzed.*

New classes of invariants for graphs are introduced and considered, which, in particular, can be used for image analysis and classification. In addition, the statements proved in the article relate to such aspects of the problem as the formation of complex types of invariants on the basis of basic ones and the finding of functional dependencies of some invariants on others.

Scientific novelty: *new composite invariants of graphs that can be effectively used in the recognition of graph-based images are constructed and substantiated.*

References

1. Karkishhenko A. N., Mnuhin V. B. Metod detekcii harakternyh tochek izobrazhenija s pomoshh'ju znakovogo predstavlenija // Izvestija JuFU. Tehnicheskie nauki, 2020. Tom 214. № 4, str. 59–70.
2. Bazhenov A. V., Filjakin A. A. Teorema grafov kak osnova postroenija sistem svjazi // Universum: tehnicheskie nauki: jelektron. nauchn. zhurnal, 2022. № 3(96).
3. Akbasheva E. A., Akbasheva G. A., Tlupov I. Z. Metody predstavlenija tekstovyh dokumentov na osnove grafov v zadachah obrabotki estestvennogo jazyka // Informatika, vychislitel'naja tehnika i upravlenie. Serija: Estestvennye i tehnicheskie nauki. 2022. № 11, str. 67–72.
4. Stepkina A. V., Stepkina A. S. Algoritmy raspoznavanija prostyh grafov kollektivnym agentom // Komp'juternye issledovanija i modelirovanie, 2021., tom 13, № 1. S.33–45. DOI: 10.20537/2076-7633-2021-13-1-33-45.
5. Nagavarapu S. C., Vachhani L., Sinha A. et al. Generalizing Multi-agent Graph Exploration Techniques // International Journal of Control, Automation and Systems. 2020. Vol. 19. P. 491–504. https://doi.org/10.1007/s12555-019-0067-8.
6. Torshin I. Yu., Rudakov K. V. Topological Chemograph Analysis Theory as a Promising Approach to Simulation Modeling of Quantum-Mechanical Properties of Molecules. Part II: Quantum-Chemical Interpretations of Chemograph TheoryPattern // Recognition and Image Analysis. 2022. Vol. 22. P. 205–217.
7. Torshin I. Yu., Rudakov K. V. Topological Chemograph Analysis Theory as a Promising Approach to Simulation Modeling of Quantum-Mechanical Properties of Molecules. Part I: Quantum-Chemical Interpretations of Chemograph TheoryPattern // Recognition and Image Analysis. 2021. Vol. 21. P. 800–810.
8. Torshin I. Yu., Rudakov K. V. Local completeness of the 'chemographs' invariants in view of the combinatorial theory of solvability // Pattern Recognition and Image Analysis. 2014. Vol. 24. P. 196–208.
9. Abgaldaeva A. A., Pushkin A. Ju. Primenenie teorii grafov v sfere informacionnyh tehnologij // Universum: tehnicheskie nauki: jelektron. nauchn. zhurn. 2023. № 2(107). URL: https://7universum.com/ru/tech/archive/item/15061.
10. Sapunov S. V., Senchenko A. S. Lingvisticheskoe predstavlenie grafov s pomechennymi vershinami // Dopovidi Nacional'noï akademiï nauk Ukraïni. 2019. № 11. S. 17–24.
11. Kurapov S. V., Davidovskij M. V. Vychislitel'nye metody opredelenija invariantov grafa///International Journal of Open Information Technologies ISSN: 2307-8162. 2021. Vol. 9, № 2. S. 1–8.
12. Tutygin R. A., Zjabliceva L. V. Jeffektivnost' invariantov grafov, sootvetstvujushhih polugruppam // Sb trudov konferencii: Materialy VI Mezhdunarodnoj nauchno-prakticheskoj konferencii (shkoly-seminara) molodyh uchenyh. Tol'jatti, 2020. S.114–117.
13. Gjeri M., Dzhonson D. Vychislitel'nye mashiny i trudnoreshaemye zadachi // M.: Mir, 2012.
14. Zykov A. A., Graphs Theory (Nauka, Moscow, 1986) [in Russian].
15. Leont'ev V. K. Kombinatorika i informacija. Chast' 1. Kombinatornyj analiz. M.: MFTI, 2015. 174 s.

3 Eduard N. Gordeev, Dr.Sc. of Physical and Mathematical Sciences, Professor of the Department of IU-8 «Information Security» of the Bauman Moscow State Technical University. Moscow, Russian Federation. E-mail: werhorngord@gmail.com

4 Vladimir K. Leontiev, Dr.Sc. of Physical and Mathematical Sciences, Computer Center of the Institute of Computer Science «Informatics and Control», Moscow, Russia. E-mail: vkleontiev@yandex.ru

АЛГОРИТМЫ ОБРАБОТКИ ЦИФРОВОГО ВОДЯНОГО ЗНАКА ДЛЯ ЗАЩИТЫ АВТОРСКИХ ПРАВ НА ГРАФИЧЕСКИЕ ФАЙЛЫ

Сысоев В. В.[1], Быков А. Ю.[2]

DOI: 10.21681/2311-3456-2025-5-139-148

Цель исследования: *создание стеганографических алгоритмов, стойких к различным видам воздействий, для встраивания цифрового водяного знака в графический файл и извлечения его.*

Методы исследования: *быстрое дискретное преобразование Фурье, алгоритм Блюстейна, действия над комплексными числами, действия над матрицами.*

Результат исследования: *описаны этапы работы алгоритма создания стеганограммы на основе графического файла-контейнера, содержащей цифровой водяной знак (ЦВЗ), а также алгоритма извлечения ЦВЗ из созданной стеганограммы. При встраивании ЦВЗ в изображение было использовано быстрое дискретное преобразование Фурье по алгоритму Блюстейна, встраивание рекомендовано в область спектра низких или средних частот. Проведено моделирование работы алгоритма с помощью данных, полученных на имитаторе. При моделировании показаны этапы работы алгоритма и продемонстрированы особенности изменения данных контейнера и ЦВЗ при встраивании и извлечении. Представлены примеры работы алгоритмов встраивания и извлечения ЦВЗ. Проведено исследование зависимости изменений исходного контейнера от размера ЦВЗ после его встраивания и сравнение нового алгоритма с популярными стеганографическими алгоритмами. Показана стойкость алгоритма к различным видам воздействий на стеганограмму.*

Научная новизна: *заключается в разработке стеганографического алгоритма, пригодного для защиты авторского права на изображения, на основе встраивания ЦВЗ в яркостную составляющую цвета графического файла-контейнера с использованием быстрого преобразования Фурье.*

Ключевые слова: *стеганография, цифровое графическое изображение, файл-контейнер, матрица пикселей.*

Введение

Проблема защиты авторских или эксклюзивных прав на информацию является актуальной в коммерческих и государственных организациях в силу ценности как самой информации, так и прямым требованием действующего законодательства РФ. Несанкционированное/неправомерное применение данной информации может приводить для организаций как к репутационному ущербу, так и к прямым финансовым убыткам.

Предлагаемые в статье алгоритмы делают возможным защиту цифрового изображения графического файла путем встраивания в него ЦВЗ и последующей проверки ЦВЗ. Алгоритм встраивания является стойким к различным видам воздействий, при этом изображение после встраивания остается визуально не отличимым для человека от оригинального изображения. Основы предложенных алгоритмов встраивания и проверки ЦВЗ рассмотрены в [1], ниже алгоритмы описаны более подробно, а также приведены обоснование работы алгоритмов и результаты экспериментов.

Одна из задач, поставленная в рамках этой статьи, – создание простого, незаметного и надежного способа внедрения ЦВЗ в цифровое изображение. В конечном счете, результатом, достигаемым при решении вышеуказанной задачи, является обеспечение возможности автоматизированного формирования и внедрения в цифровое изображение цифровой метки, содержащей информацию, необходимую для проведения расследования и установления обстоятельств в случае несанкционированного использования графических материалов сторонними лицами.

Другая задача, которую нужно было решить в рамках данной статьи, – создание алгоритма для получения стеганограммы, стойкой к изменениям формата изображения, обрезанию изображения, сжатию, обработке шумами. Желательно, чтобы стеганограмма, созданная алгоритмом, имела стойкость ко всем перечисленным возможным изменениям одновременно [2].

В настоящий момент наиболее продвинутыми алгоритмами для встраивания ЦВЗ являются алгоритмы, основанные на методе NeRF [3]. Данный метод имеет много вариаций, например, IW-NeRF [4], RWNeRF [5], CopyNerRF [6], StegaNeRF [7]. Общий

1 Сысоев Валентин Валерьевич, аспирант кафедры «Информационная безопасность», МГТУ им. Н. Э. Баумана, г. Москва, Россия. E-mail: valsus88@mail.ru
2 Быков Александр Юрьевич, кандидат технических наук, доцент, доцент кафедры «Информационная безопасность», МГТУ им. Н. Э. Баумана, г. Москва, Россия. E-mail: abykov@bmstu.ru

принцип работы вышеназванных алгоритмов основан на том, что из объекта извлекаются его цветовые параметры, на которых обучается нейронная сеть MLP (MultiLayer Perceptron). Ключ, который представляет собой встраиваемый ЦВЗ у каждого алгоритма, создается и встраивается по его оригинальному методу. Достоинства алгоритмов семейства NeRF заключается в том, что они применимы как для 2D, так и 3D объектов и обладают относительно хорошей помехоустойчивостью. Недостатки вытекают из особенностей метода. Так как метод использует цветовые параметры объекта, при внесении в него значительных изменений информация о ключе теряется. Кроме того, создание пар «объект – нейросеть», вычислительно затратно, поскольку требует значительные вычислительные мощности для создания нейросети.

Поэтому необходим такой алгоритм, который бы, с одной стороны, работал по классическому принципу внесения ЦВЗ в изображение, но в то же время работал с цветовым представлением всего объекта, а не какой-то его части, так как это снижает эффективность инструментов стеганоанализа [8].

Одна из особенностей алгоритма, приведенного в данной статье в том, что он использует быстрое дискретное преобразование Фурье (ДПФ) в процессе создания стеганограммы. Быстрое ДПФ вычисляется посредством алгоритма Блюстейна[3]. Алгоритм Блюстейна отличается от других алгоритмов быстрого ДПФ тем, что является более гибким и быстрым алгоритмом, в котором нужно проводить меньше вычислений[4]. Алгоритм Блюстейна имеет ту же асимптотическую сложность, что и все алгоритмы быстрого ДПФ: $O(n\ \log(n))$, выбор его использования обусловлен более простой реализацией, которая позволяет задействовать его на устройствах с незначительными вычислительными ресурсами.

Имеются так же алгоритмы, схожие с предложенным, основанные на вейвлет-преобразовании [9]. Однако данная реализация сложнее, имеет сложность $O(n^2)$ и потребляет больше памяти. Другие особенности стеганографических алгоритмов рассмотрены в [10].

Математическая постановка задачи

В рамках данной статьи рассматриваются две задачи – внедрение цифрового водяного знака (ЦВЗ) в контейнер (создание стеганограммы) и извлечение из стеганограммы внедренного ЦВЗ.

3 Блеихут Р. Быстрые алгоритмы цифровой обработки сигналов. - М.: Мир, 1989. 448 с.

4 Sirin S. CZT vs FFT: Flexibility vs Speed. Embedded Systems Programming magazine. 2003. DOI: https://www.researchgate.net/publication/241978861_CZT_vs_FFT_Flexibility_vs_Speed.

1.1. Внедрение цифрового водяного знака в контейнер

Дано следующее:

Изображение C – контейнер в произвольном формате размерностью $h \times w$, где w – ширина изображения в пикселях, h – высота изображения в пикселях. Таким образом, что изображение C можно изобразить матрицей пикселей

$$C_{h\times w} = \begin{pmatrix} p_{11}^{(C)} & \cdots & p_{1w}^{(C)} \\ \vdots & \ddots & \vdots \\ p_{h1}^{(C)} & \cdots & p_{hw}^{(C)} \end{pmatrix},$$

где p_{ij} – пиксель изображения в цветовом пространстве RGB, характеризирующийся тремя цветовыми спектрами p_{ij} (r_{ij},g_{ij},b_{ij}), где r_{ij}, g_{ij}, b_{ij} – интенсивности составляющих трех цветов (красный, зеленый, синий).

Сообщение M в текстово-числовом виде предстоит внедрить в контейнер. Сообщение обычно является идентификатором автора (цифровым водяным знаком) для подтверждения авторского права.

Необходимо найти:

Стеганограмму CM, представляющую собой изображение в цветовом пространстве RGB, той же размерности $h \times w$, со встроенным сообщением M, визуально неотличимое от изображения C. Такое, что

$$CM_{h\times w} = \begin{pmatrix} p_{11}^{(CM)} & \cdots & p_{1w}^{(CM)} \\ \vdots & \ddots & \vdots \\ p_{h1}^{(CM)} & \cdots & p_{hw}^{(CM)} \end{pmatrix}.$$

1.2. Извлечение ЦВЗ из стеганоконтейнера

Дано следующее:

Изображение в цветовом пространстве RGB, являющееся стеганограммой со встроенным сообщением M, заданное матрицей $CM_{h\times w}$.

Необходимо найти: сообщение M в текстово-числовом виде.

2. Описание алгоритмов

2.1. Предварительные алгоритмы

Взаимная конверсия изображений из пикселей формата RGB в формат HSV и обратно.

По пиксельный перевод изображения из формата RGB в формат HSV обусловлен тем, что будем размещать метку не в каком-то одном цветовом диапазоне, а во всех диапазонах сразу, это повысит ее стойкость.

Пусть изображение Im представляет собой матрицу пикселей размерностью $h \times w$, где w – ширина изображения в пикселях, h – высота изображения в пикселях

$$Im_{h\times w} = \begin{pmatrix} p_{11}^{(Im)}(r_{11},g_{11},b_{11}) & \cdots & p_{1w}^{(Im)}(r_{1w},g_{1w},b_{1w}) \\ \vdots & \ddots & \vdots \\ p_{h1}^{(Im)}(r_{h1},g_{h1},b_{h1}) & \cdots & p_{hw}^{(Im)}(r_{hw},g_{hw},b_{hw}) \end{pmatrix}.$$

Для каждого пикселя $p_{ij}^{(Im)}(r_{ij},g_{ij},b_{ij})$ выполним приведение к виду $p_{ij}^{(Im')}(h_{ij},s_{ij},v_{ij})$ по общеизвестным соотношениям, где h_{ij}, s_{ij}, v_{ij} – составляющие тона, насыщенности и яркости

$$Im_{h\times w}=\begin{pmatrix} p_{11}^{(Im')}(h_{11},s_{11},v_{11}) & \cdots & p_{1w}^{(Im')}(h_{1w},s_{1w},v_{1w}) \\ \vdots & \ddots & \vdots \\ p_{h1}^{(Im')}(h_{h1},s_{h1},v_{h1}) & \cdots & p_{hw}^{(Im')}(h_{hw},s_{hw},v_{hw}) \end{pmatrix}.$$

Перевод матрицы $Imp'_{h\times w}$ изображения Im' из цветового пространства HSV в цветовое пространство RGB выполняется также по известным преобразованиям.

2.2. Преобразование Блюстейна

Основная идея вычислений заключается в замене прямого вычисления дискретного преобразования Фурье через использование свёрток. Вместо непосредственного вычисления суммы произведений, алгоритм Блюстейна преобразует её в форму, удобную для вычислений методом быстрого ДПФ. Используя за основу формулы из [8], составим формулы, наиболее применимые для нашего случая.

Пусть $X = |x_0,x_1,\ldots,x_{N-1}|$ — это входная последовательность (дискретная функция) длины N. Тогда выходная последовательность комплексных чисел $Y = |y_0,y_1,\ldots,y_{N-1}|$ такой же длины, вычисленная по ДПФ, определяется следующим образом:

$$y_k=\sum_{n=0}^{N-1} x_n\, e^{-i\frac{2\pi}{N}kn}, \qquad (1)$$

где $k = 0, 1, \ldots, N-1$ – индексы выходной последовательности.

В общем виде можно записать $Y = F(X)$, где F – оператор дискретного преобразования Фурье. Вычисление последовательности по алгоритму Блюстейна $Y = CZT(Y)$ может быть выражено через сумму свёрток, умноженную на N фазовых коэффициентов, и представлено в следующем виде:

$$y_k = b_k^* \sum_{n=0}^{N-1} a_n \times b_{|k-n|},\ k = 0, 1, \ldots, N-1, \qquad (2)$$

где $a_n = x_n\, e^{-i\frac{\pi}{N}n^2}$, $b_n = e^{-i\frac{\pi}{N}n^2}$, $b_k^* = e^{-i\frac{\pi}{N}k^2}$, $*$ – обозначение свертки.

Обратное преобразование Фурье $X' = FCZT(Y)$ вычисляется по следующей формуле:

$$x'_k=\frac{1}{N}\sum_{n=0}^{N-1} y_j\, e^{i\frac{2\pi}{N}kj},\ k = 0, 1, \ldots, N-1. \qquad (3)$$

В общем виде $X' = \mathcal{F}^{-1}Y$, где $\mathcal{F}^{-1}$ – оператор обратного дискретного преобразования Фурье.

Причина использования преобразования Блюстейна заключается в том, что оно удобно для вычисления дискретного преобразования Фурье любой входной длины. Особенность преобразования в том, что оно позволяет в уравнении свертки заполнять, при выполнении определенных условий, последовательность a_n нулями, тем самым, упрощая вычисление.

2.3. Внедрение ЦВЗ в контейнер

Алгоритм делится на несколько этапов:

ЭТАП I:

Этап I представляет собой предварительные действия над сообщением и контейнером, в который нужно внедрить сообщение.

Цепочка А – представляет собой действия над сообщением, которое нужно внедрить в контейнер.

1А: Сообщение из текстово-числового вида преобразуется в метку – изображение M произвольного размера $k \times l$, где l – ширина изображения метки, k – высота изображения метки,

$$k < h, l < w\colon M_{k\times l}=\begin{pmatrix} p_{11}^{(M)} & \cdots & p_{1l}^{(M)} \\ \vdots & \ddots & \vdots \\ p_{k1}^{(M)} & \cdots & p_{kl}^{(M)} \end{pmatrix}.$$

Метка формируется по следующему правилу:

Фон метки черный, т.е. всем пикселям $p_{ij}^{(M)}$ метки M устанавливаются значения $p_{ij}^{(M)}(0,0,0)$.

В произвольном месте метки наносится изображение, аналогичное графическому образу сообщения M, т.е. пиксели устанавливаются в белый цвет, $p_{ij}^{(M)}(1,1,1)$, если использовать интервал для задания цвета $[0, 1]$.

2А: Осуществляется попиксельный перевод изображения M из цветового пространства RGB в цветовое пространство HSV,

$$M'_{k\times l}=\begin{pmatrix} p_{11}^{(M')}(h_{11},s_{11},v_{11}) & \cdots & p_{1l}^{(M')}(h_{1l},s_{1l},v_{1l}) \\ \vdots & \ddots & \vdots \\ p_{k1}^{(M')}(h_{k1},s_{k1},v_{k1}) & \cdots & p_{kl}^{(M')}(h_{kl},s_{kl},v_{kl}) \end{pmatrix}.$$

3А: Из каждого пикселя $p_{ij}^{(M')}$ матрицы $M'_{k\times l}$ извлекаем соответствующую компоненту яркости v_{ij} и составляем соответствующую матрицу яркости метки

$$V_{k\times l}^{(M')}=\begin{pmatrix} v_{11}^{(M)} & \cdots & v_{1l}^{(M)} \\ \vdots & \ddots & \vdots \\ v_{k1}^{(M)} & \cdots & v_{kl}^{(M)} \end{pmatrix}.$$

Эта матрица состоит из 0 и 1 для черных и белых пикселей, соответственно.

Выбор значения яростной характеристики для кодирования метки обусловлено тем, что яркостная характеристика, при переводе из одного одной цветовой модели в другую, меньше всего подвергается математическим преобразованиям, при этом она покрывает все изображение, является подобием математического ожидания для распределения интенсивностей цветов в каждом пикселе изображения.

Также не маловажно, что изменение яркостной характеристики даже всего изображения на 2–4 %[5] остается незаметным для наблюдателя, что так же подтверждается и другими исследованиями [11].

5 David H. Hubel. Eye, Brain, and Vision. - Henry Holt and Company, 1995. 242 p.

Цепочка Б – представляет собой действия над контейнером.

1Б: В изображении – контейнере C переводим каждый пиксель из цветового пространства RGB в цветовое пространство HSV. Для каждого пикселя RGB $p_{ij}^{(C)}(r_{ij},g_{ij},b_{ij})$ получаем пиксель в формате HSV $p_{ij}^{(C)}(h_{ij},s_{ij},v_{ij})$. Получаем матрицу

$$C'_{h\times w}=\begin{pmatrix} p_{11}^{(C')} & \cdots & p_{1w}^{(C')} \\ \vdots & \ddots & \vdots \\ p_{h1}^{(C')} & \cdots & p_{hw}^{(C')} \end{pmatrix}.$$

Для краткости (r_{ij},g_{ij},b_{ij}) и (h_{ij},s_{ij},v_{ij}) далее опускаем.

2Б: Из каждого пикселя $p_{ij}^{(C')}$ матрицы C' извлекаем соответствующую компоненту яркости v_{ij} и составляем соответствующую матрицу яркости контейнера C':

$$V_{h\times w}^{(C')}=\begin{pmatrix} v_{11}^{(C')} & \cdots & v_{1w}^{(C')} \\ \vdots & \ddots & \vdots \\ v_{h1}^{(C')} & \cdots & v_{hw}^{(C')} \end{pmatrix}.$$

3Б: матрицу яркости контейнера $V_{h\times w}^{(C')}$ переводим в матрицу-строку $V_{hw}^{(C')}$ путем соединения всех h строк матрицы $V_{h\times w}^{(C')}$ в одну строку $V_{hw}^{(C')}=(V_1^{(C')}\ldots V_{hw}^{(C')})$, где $v_i\in[0,1]$.

4Б: над матрицей-строкой $V_{hw}^{(C')}$ проводится быстрее преобразование Фурье по формуле Блюстейна (2). В результате получаем: $CZT(V_{hw}^{(C')})=Z_{hw}^{(C')}=(z_1^{(C')}\ldots z_{hw}^{(C')})$, где $z_i^{(C')}\in C$ (множество комплексных чисел).

ЭТАП II:

На этапе II происходит непосредственное внедрение метки в контейнер:

1. Матрица-строка $Z_{hw}^{(C')}$ переводится в матрицу размерности $h\times w$ путем перевода w чисел из $Z_{hw}^{(C')}$ на новую строку:

$$Z_{h\times w}^{(C')}=\begin{pmatrix} z_{11}^{(C')} & \cdots & z_{1w}^{(C')} \\ \vdots & \ddots & \vdots \\ z_{h1}^{(C')} & \cdots & z_{hw}^{(C')} \end{pmatrix}.$$

2. Матрица $Z_{h\times w}^{(C')}$ складывается с матрицей $V_{k\times l}^{(M')}$ по определенным правилам:

$$Z_{h\times w}^{(CM')}=Z_{h\times w}^{(C')}+V_{k\times l}^{(M')}. \quad (4)$$

Складываем матрицы по следующему правилу: в матрице $Z_{h\times w}^{(C')}$ выбираем прямоугольник размерности $k\times l$, например, начиная с элемента $z_{ij}^{(C')}$, таким образом, чтобы не выйти за границы матрицы размерности $h\times w$, как правило, размерность $h\times w$ существенно выше, чем $k\times l$. К элементам этого прямоугольника прибавляем соответствующие элементы матрицы $V_{k\times l}^{(M')}$, в сложении участвуют действительные части комплексных чисел матрицы $Z_{h\times w}^{(C')}$. Выбор индексов i и j элемента, с которого начинается встраивание в матрицу контейнера, предлагается выполнять так, чтобы метка внедрялась в область спектра низких частот, в большей степени, и средних частот – в меньшей [12].

3. Полученную при сложении матрицу $Z_{h\times w}^{(CM')}$ переводим в матрицу-строку $Z_{hw}^{(CM')}$ путем соединения всех w строк матрицы $Z_{h\times w}^{(CM')}$ в одну строку $Z_{hw}^{(CM')}$.

4. Над матрицей-строкой $Z_{hw}^{(CM')}$ проводится обратное дискретное преобразование Фурье в соответствии с (3). В результате получаем матрицу-строку $V_{hw}^{(CM')}$: $FCZT(Z_{hw}^{(CM')})=V_{hw}^{(CM')}=(v_1^{(CM')}\ldots v_{hw}^{(CM')})$.

ЭТАП III:

На этапе III происходит восстановление стеганограммы из матричного вида обратно в визуальный вид.

1. Переведем матрицу-строку $V_{hw}^{(CM')}$ в матрицу размерности $w\times h$ путем перевода строки из h чисел из $V_{hw}^{(CM)}$ на новую строку:

$$V_{h\times w}^{(CM')}=\begin{pmatrix} v_{11}^{(CM')} & \cdots & v_{1w}^{(CM')} \\ \vdots & \ddots & \vdots \\ v_{h1}^{(CM')} & \cdots & v_{hw}^{(CM')} \end{pmatrix}$$

2. В матрице $C'_{h\times w}$ заменим все компоненты яркости v_{ij} в каждом пикселе $p_{ij}^{(C')}$ на компоненты яркости из $V_{h\times w}^{(CM')}$:

$$CM'_{k\times l}=\begin{pmatrix} p_{11}^{(CM')}(h_{11},s_{11},v_{11}) & \cdots & p_{1w}^{(CM')}(h_{1w},s_{1w},v_{1w}) \\ \vdots & \ddots & \vdots \\ p_{h1}^{(CM')}(h_{h1},s_{h1},v_{h1}) & \cdots & p_{hw}^{(CM')}(h_{hw},s_{hw},v_{hw}) \end{pmatrix}.$$

Если компонента $v_{ij}^{(CM')}$ выходит за интервал $[0, 1]$, то значения, меньшие 0, заменяются на 0, а значения, большие 1, заменяются на 1.

3. При необходимости переведем пиксели в матрице $C'_{h\times w}$ из цветового пространства HSV в цветовое пространство RGB по известным преобразованиям, получаем на выходе алгоритма матрицу:

$$CM_{w\times h}=\begin{pmatrix} p_{11}^{(CM)} & \cdots & p_{1w}^{(CM)} \\ \vdots & \ddots & \vdots \\ p_{h1}^{(CM)} & \cdots & p_{hw}^{(CM)} \end{pmatrix}.$$

2.4. Извлечение метки из контейнера

1. Пиксели матрицы $CM_{w\times h}$ переводятся из цветового пространства RGB в цветовое пространство HSV, получаем матрицу $CM'_{h\times w}$.

2. Из матрицы $CM'_{h\times w}$ извлекаем компоненту яркости $v_{ij}^{(CM')}$, создавая таким образом новую матрицу яркостей стегоконтейнера $V_{h\times w}^{(CM')}$.

3. Матрица $V_{h\times w}^{(CM')}$ переводится в матрицу-строку $V_{hw}^{(CM')}$ путем соединения всех строк матрицы.

4. Над матрицей-строкой $V_{hw}^{(CM')}$ проводится быстрее преобразование Фурье по алгоритму Блюстейна в соответствии с (2). В результате получаем матрицу-строку $Z_{hw}^{(CM')}$ комплексных чисел.

5. Матрицу-строку $Z_{hw}^{(CM')}$ переводят в матрицу $Z_{h\times w}^{(CM')}$ размерности $h\times w$ путем перевода каждых h элементов матрицы из $Z_{wh}^{(CM')}$ на новую строку:

6. Из матрицы $Z_{wh}^{(CM')}$ составляется матрица $D_{wh}^{(CM')}$ путем выделения действительной части чисел $z_{ij}^{(CM')}$,

$$D_{w\times h}^{(CM')} = \begin{pmatrix} d_{11}^{(CM')} & \cdots & d_{1w}^{(CM')} \\ \vdots & \ddots & \vdots \\ d_{h1}^{(CM')} & \cdots & d_{hw}^{(CM')} \end{pmatrix}.$$

7. Создаем матрицу изображения извлеченной метки из контейнера $MC'_{w\times h}$, в которой пиксели заданы в формате HSV, значения h и s в которой выставлены в 0, а значение v взято из матрицы $D_{w\times h}^{(CM')}$, при этом, если компонента $d_{ij}^{(CM')}$ выходит за интервал $[0, 1]$, то значения, меньшие 0, заменяются на 0, а значения, большие 1 заменяются на 1,

$$MC'_{h\times w} = \begin{pmatrix} p_{11}^{(MC)}(0,0,d_{11}^{(CM')}) & \cdots & p_{1w}^{(MC)}(0,0,d_{1w}^{(CM')}) \\ \vdots & \ddots & \vdots \\ p_{h1}^{(MC)}(0,0,d_{h1}^{(CM')}) & \cdots & p_{hw}^{(MC)}(0,0,d_{hw}^{(CM')}) \end{pmatrix}.$$

8. Переводим пиксели матрицы $MC'_{h\times w}$ из формата HSV в формат RGB и получаем таким образом визуальное изображение $MC_{h\times w}$.

9. По визуальному изображению находим искомое сообщение M.

2.5. Обоснование работы алгоритма

Рассмотрим пример работы алгоритмов встраивания и извлечения на имитаторе исходных данных. По сути, алгоритм встраивает данные (вектор из 0 и 1, аналог метки) в матрицу-строку (вектор) $V_{hw}^{(C')}$. Для имитации был сгенерирован вектор исходных данных для встраивания из 200 вещественных чисел в интервале $[0, 1]$, для заполнения использовался следующий код на языке C++:

```
size_t N=200;
vector<double> c(N); // Контейнер
for(size_t i=1; i<=N; i++) c[i-1]=sin(i)/2+0.5;
```

Метка представляла из себя вектор такой же размерности, в котором все 0, в начале несколько элементов 1. Результаты имитации встраивания (табл. 1).

Таблица 1.

Результаты имитации встраивания данных

Индекс элемента	Данные контейнера	Действительная часть после прямого ДПФ	Встраиваемая метка	Действительная часть после прямого ДПФ с меткой	Результат обратного ДПФ (контейнер с меткой)	Результат прямого ДПФ для извлечения метки
1	2	3	4	5	6	7
0	0,9207	100,0163	0,0000	100,0163	0,9407	1,0000
1	0,9546	0,0161	0,0000	0,0161	0,9745	0,0135
2	0,5706	0,0154	1,0000	1,0154	0,5900	0,4935
3	0,1216	0,0143	1,0000	1,0143	0,1404	0,4915
4	0,0205	0,0127	1,0000	1,0127	0,0385	0,5054
5	0,3603	0,0106	1,0000	1,0106	0,3771	0,5026
6	0,8285	0,0080	0,0000	0,0080	0,8439	0,0220
7	0,9947	0,0048	0,0000	0,0048	1,0000	0,0209
8	0,7061	0,0010	0,0000	0,0010	0,7183	0,0027
9	0,2280	–0,0035	0,0000	–0,0035	0,2384	0,0000
...						
190	0,7975	–0,0086	0,0000	–0,0086	0,8060	0,0000
191	0,3225	–0,0035	0,0000	–0,0035	0,3330	0,0000
192	0,0108	0,0010	0,0000	0,0010	0,0230	0,0027
193	0,1488	0,0048	0,0000	0,0048	0,1627	0,0209
194	0,6097	0,0080	0,0000	0,0080	0,6252	0,0220
195	0,9698	0,0106	0,0000	0,0106	0,9866	0,5026
196	0,8979	0,0127	0,0000	0,0127	0,9158	0,5054
197	0,4602	0,0143	0,0000	0,0143	0,4790	0,4915
198	0,0591	0,0154	0,0000	0,0154	0,0786	0,4935
199	0,0634	0,0161	0,0000	0,0161	0,0832	0,0135

Рис. 1. Исходный контейнер и он же со встроенной стеганограммой

В таблице 1 представлены первые и последние 10 элементов вектора, на которых показано встраивание и извлечение метки. В первом столбце заданы индексы значений вектора данных, во втором – сами данные, полученные имитатором. В третьем столбце – действительные части комплексных значений, полученных после прямого ДПФ. В четвертом столбце – данные вектора встраиваемой метки, в векторе метки элементы, равные 1, имеют индексы от 2 до 5, остальные равны 0. В пятом столбце – результат сложения третьего и четвертого столбцов, аналог выполнения операции, заданной в (4). Элементы, соответствующие значениям единиц в векторе метки, выделены цветом фона в таблице.

Следует отметить, что при сложении значений единиц из метки они складываются со значениями из третьего столбца по модулю существенно меньше 1, и большинство значений в третьем столбце, кроме константы по индексу 0, по модулю меньше единицы. Также вторая половина из третьего столбца повторяют первую, кроме константы по индексу 0, если смотреть в обратном порядке, это свойство прямого ДПФ.

После обратного ДПФ значений пятого столбца получаем имитацию стенограммы с меткой, данные ее в столбце 6. Эти данные близки с исходными данными второго столбца, они бы совпали, если бы не прибавили к данным метку. Изменения, внесённые меткой в отдельные значения, распределяются по всем данным шестого столбца. В седьмом столбце представлены данные, полученные после прямого ДПФ, для извлечения метки. Если бы данные пятого столбца не были бы изменены меткой, то седьмой столбец совпал бы с пятым. При этом элементы седьмого столбца с индексами, в которых у вектора-метки находятся 1, существенно больше соседних элементов этого столбца, что позволяет визуально отличать метку. Учитывая то, что при прямом ДПФ вторая половина полученных данных, совпадает с первой, то происходит дублирование метки во второй половине – элементы с индексами 195–198.

3. Эксперименты

3.1. Примеры работы алгоритмов

Рассмотрим результаты работы алгоритма на примере рисунка 1024 x 1024 пикселей (рис. 1).

В контейнер встроим сообщение M: «Авторское право». Визуализация матрицы после прямого ДПФ со встроенной меткой – матрица действительных частей из $Z_{h\times w}^{(CM')}$ (рис. 2).

Рис. 2. Визуализация контейнера с меткой после обратного ДПФ

Визуализация матрицы предполагает, что по ней создаем в начале изображение в формате HSV, где элементы матрицы – это значения яркости, причем, если значение элемента выходит за пределы $[0, 1]$, то значения, меньшие 0, заменяются на 0, а больше 1 – на 1, значения тона и насыщенности равны 0, при необходимости переводим в формат RGB. Как видно, метка хорошо выделяется, так как остальные пиксели в основном черные (значения яркости близки к 0), идет небольшая засветка по краям контейнера.

Далее после обратного преобразования получаем контейнер с меткой, он визуально для человека

аналогичен (рис. 1). Для извлечения метки выполняем прямое ДПФ, получаем матрицу $MC'_{h\times w}$, ей соответствует изображение (рис. 3).

Рис.3. Изображение с извлеченной меткой

Из рисунка видно, что несмотря на то, что яркость метки уменьшилась, она вполне читаема, также произошло дублирование метки во второй половине контейнера, если смотреть в обратном порядке, что является свойством прямого ДПФ.

3.2. Оценка степени изменения исходного контейнера при встраивании цифрового водяного знака

Рассмотрим оценку изменения исходного контейнера при встраивании в него ЦВЗ в зависимости от размера ЦВЗ. Размер ЦВЗ знака будем измерять отношением числа единиц в метке или изменяемых пикселей при встраивании к числу пикселей в контейнере, заданных в процентах. Изменение в контейнере будем измерять средней квадратичной ошибкой на один пиксель, вычисляемой по формуле:

$$E = \sqrt{\frac{\sum_{i=1}^{h}\sum_{j=1}^{w}(v_{ij}^{(C')} - v_{ij}^{(CM')})^2}{hw}}. \quad (5)$$

График зависимости этой ошибки от числа измененных пикселей в процентах (рис. 4). Увеличение измененных пикселей производится путем добавления символов к встраиваемой метке, измерения проводилось на контейнере (рис. 1).

Рис. 4. Зависимость ошибки от числа измененных пикселей

Из рисунка 4 можно сделать вывод о том, что зависимость изменения в контейнере от размера метки, скорее всего, является линейной или близкой к линейной.

3.3. Сравнение с другими алгоритмами

Сравнение предложенного алгоритма (в табл. 2 заголовок – алгоритм) производилось со следующими популярными стеганографическими алгоритмами: LSB [13], F3 [14], Jsteg [15], дифференциально яркостной ячеистый алгоритм (Digital Brightness Cell Algoritm, DBCA) [16], Куттера [17], Patchwork [18], преобразования Фурье в синем спектре XYZ (FXYZ).

В качестве оценки выбралось способность алгоритмов противостоять различным видам воздействий на стеганограмму. В поле стойкость ставится:

«+» – если стеганограмма, полученная данным алгоритмом, выдержала воздействие, метка была успешно извлечена.

«–» – если стеганограмма, полученная данным алгоритмом, не выдержала воздействие, метка не была успешно извлечена.

«∓» – если стеганограмма, полученная данным алгоритмом, выдержала воздействие, но при извлечении

Таблица 2.

Результаты стойкости стеганографических алгоритмов к различным преобразованиям

№ п/п	Преобразования стеганограммы	Стеганографические алгоритмы							
		LSB	F3	Jsteg	DBCA	Куттера	Patchwork	FXYZ	Алгоритм
1	Конверсия изображения в другие форматы	+	–	–	+	∓	∓	+	+
2	Вырезание частей	–	–	–	+	∓	∓	+	+
3	Обрезание части	–	–	–	+	∓	∓	-	+
4	Ухудшение качества путем сжатия	–	–	–	–	∓	∓	+	+
5	Изменение характеристик	–	–	–	–	∓	∓	–	+
6	Наложение шума	–	–	–	–	∓	∓	–	+

получились неясные результаты, метка была извлечена частично, или в некоторых изображениях она извлекается полностью, в некоторых не извлекается вообще.

Были произведены эксперименты над 10-ю изображениями с различными тематическими содержаниями. В результате получили следующие результаты (табл. 2):

Особенности преобразований стеганограммы:

1. Конверсия исходного изображения стеганограммы в другие форматы.

Перевод изображения из одного формата в другие, например, исходное изображение в формате .png в формат .jpg.

2. Вырезание частей из стеганограммы.

Изменение представляет собой вырезание части, изображения стеганограммы, но не менее 1/4 от нее, метка находится в вырезанной части. В дальнейшем предпринимается попытка извлечь из вырезанной части изображения скрытую метку.

3. Обрезание части стеганограммы.

Изменение похоже на предыдущее, но в данном случае ведется обрезание части изображения с двух сторон стеганограммы (например, обрезание изображения для перехода в формат от 16:10 к 16:9) так, чтобы от полученной стеганограммы осталось не менее 1/8 от нее оригинала, метка находится в оставшейся части. В дальнейшем предпринимается попытка извлечь из оставшейся части изображения скрытую метку.

4. Ухудшение качества стеганограммы путем сжатия изображения.

Стеганограмма подвергается алгоритмам сжатия. В дальнейшем предпринимается попытка извлечь скрытую метку.

5. Изменение характеристик стеганограммы.

В стеганограмме увеличивается и уменьшается яркость и/или контраст изображения. В дальнейшем предпринимается попытка извлечь скрытую метку.

6. Наложение шума на стеганограмму.

На стеганографию наносится цифровой шум – намеренно искажаются хаотично разбросанные пиксели. В дальнейшем предпринимается попытка извлечь скрытую метку.

Выводы

В статье был предложен алгоритм встраивания ЦВЗ в графический файл на основе быстрого ДПФ по алгоритму Блюстейна, а также был приведен алгоритм извлечения ЦВЗ из созданного контейнера.

Продемонстрированы особенности работы алгоритма на данных, полученных с помощью имитатора. Представлены результаты работы описанного алгоритма на примере встраивания и извлечения ЦВЗ в контейнер.

Проведено исследование зависимости изменений в контейнере при встраивании от размера ЦВЗ (отношение числа изменяемых пикселей при встраивании к числу пикселей в контейнере). Изменение контейнера измерялось средней квадратической ошибкой на один пиксель после встраивания ЦВЗ. По результатам эксперимента можно сделать вывод о том, что эта зависимость является близкой к линейной.

Проведено сравнение предложенного алгоритма с семью популярными стеганографическими алгоритмами. Показана стойкость алгоритма к шести различным видам изменений стеганограммы.

Работоспособность предложенных алгоритмов получила практическое подтверждение на проведенных экспериментах. Достоверность предлагаемых научных решений подтверждена анализом свойств прямого и обратного дискретных преобразований Фурье.

Литература

1. Сысоев В. В., Быков А. Ю. Защита авторского права с использованием цифровой голографии // Сборник трудов XIII всероссийской научно-технической конференции «Безопасные информационные технологии» – М.: МГТУ им. Н. Э.Баумана. 2024. С. 207–213.
2. Макаренко С.И. Эталонная модель взаимодействия стеганографических систем и обоснование на ее основе новых направлений развития теории стеганографии // Вопросы кибербезопасности. 2014. № 2(3). С. 24–32.
3. Li D., Yang Z., Jin X. Zero watermarking scheme for 3D triangle mesh model based on global and local geometric features. Multimed. Tools Appl. 2023. Vol. 82. P. 43635–43648.
4. Chen L. and another. IW-NeRF: Using Implicit Watermarks to Protect the Copyright of Neural Radiation Fields. Applied Sciences. 2024., Vol. 14. No 6184. DOI:10.3390/app14146184.
5. Sun W. and another. RWNeRF: Robust Watermarking Scheme for Neural Radiance Fields Based on Invertible Neural Networks. Computers. Materials & Continua. 2024. Vol. 80. P. 4065–4083. DOI:10.32604/cmc.2024.053115.
6. Luo Z., Guo Q., Cheung K.C., See S., Wan R. CopyRNeRF: Protecting the CopyRight of Neural Radiance Fields. 2023. DOI:10.48550/arXiv.2307.11526.
7. Li C., Feng B.Y., Fan Z., Pan P., Wang Z. StegaNeRF: Embedding Invisible Information within Neural Radiance Fields. 2022. DOI:10.48550/arXiv.2212.01602.
8. Сивачев А. В., Прохожев Н. Н., Михайличенко О. В., Башмаков Д. А. Эффективность стеганоанализа на основе методов машинного обучения // Вопросы кибербезопасности. 2017. № 2(20). С. 53–60. DOI 10.21581/2311-3456-2017-2-53-60.

9. Глинская, Е. В., Чичварин Н. В. Информационная безопасность открытых каналов передачи проектной документации, продуцируемой в САПР // Вопросы кибербезопасности. 2014. № 4(7). С. 11–22.
10. Абасова, А. М., Бабенко Л. К. Защита информационного содержания изображений в условиях наличия деструктивного воздействия // Вопросы кибербезопасности. 2019. № 2(30). С. 50–57. DOI 10.21681/2311-3456-2019-2-50-57.
11. Козачок А. В., Копылов С. А., Бочков М. В. Оценка параметров необнаруживаемости разработанного подхода к маркированию текстовых электронных документов // Вопросы кибербезопасности. 2020. № 1(35). С. 62–73. DOI 10.21681/2311-3456-2020-01-62-73.
12. Морковин С. В. Алгоритмы и программные средства человеко-машинной обработки цифровых водяных знаков в видеопоследовательности // Моделирование, оптимизация и информационные технологии. 2022. Т. 10. № 3(38). С. 30–31. DOI: 10.26102/2310-6018/2022.38.3.024.
13. Кривошеев И. А., Линник М. А. Статический способ стеганографического встраивания информации на основе LSB // Системы и средства информатики. 2020. Т. 30, № 3. С. 56–66. DOI 10.14357/08696527200306.
14. Brūzgienė Rasa and another. Enhancing Steganography through Optimized Quantization Tables. Electronics. 2024. Vol.13. No 2415. DOI: 10.3390/electronics13122415.
15. Binmin P., Qiao T., and another. Novel Hidden Bit Location Method towards JPEG Steganography. Security and Communication Networks. 2022. Vol. 2022. No. 8230263. P. 13. DOI: 10.1155/2022/8230263.
16. Крамаренко С. М. Способ внесения цифровых меток в цифровое изображение и устройство для осуществления способа. 2020. Роспатент. RU2739936C1.
17. Zhigalov I. E., Ozerova M. I., Evstigneev A. V. Application of Cutter–Jordan–Bossen method for data hiding in the image spatial area // Вестник Южно-Уральского государственного университета. Серия Компьютерные технологии, управление, радиоэлектроника. 2022. Т. 23. С. 16–23. DOI: 10.14529/ctcr230302.
18. Кушнеревич П. М. Анализ эффективности алгоритмов Patchwork и LSB для защиты графических образов с помощью водяных знаков // Интеграция науки, общества, производства и промышленности: проблемы и перспективы: сборник статей по итогам Международной научно-практической конференции. Волгоград. 2021. С. 121–124.

DIGITAL WATERMARK PROCESSING ALGORITHMS FOR COPYRIGHT PROTECTION OF GRAPHIC FILES

Sysoev V. V.[6], *Bykov A. Yu.*[7]

Keywords: *steganography, digital graphic image, container file, pixel matrix.*

The purpose of the study: *creation of steganographic algorithms for embedding a digital watermark in a graphic file and extracting it, resistant to various types of influences.*

Methods of research: *fast Fourier transform, Bluestein's algorithm, actions on complex numbers, actions on matrices.*

Result(s): *the stages of the algorithm for creating a steganogram based on a graphic container file containing a digital watermark, as well as an algorithm for extracting a digital watermark from a created steganogram, are described. When embedding a digital watermark in an image, a fast discrete Fourier transform using the Bluestein algorithm was used, embedding is recommended in the low or medium frequency spectrum. The simulation of the algorithm using the data obtained on the simulator is carried out. The simulation shows the stages of the algorithm's operation and demonstrates the features of changing container and data center data during embedding and extraction. Examples of algorithms for embedding and extracting a digital watermark are presented. A study of the dependence of changes in the original container on the size of the digital watermark after its embedding and a comparison of the new algorithm with popular steganographic algorithms has been conducted. The algorithm's resistance to various types of effects on the steganogram is shown.*

Scientific novelty: *the goal is to develop a steganographic algorithm suitable for image copyright protection based on embedding a digital image into the brightness component of the color of a graphic container file using the fast Fourier transform*

.

References

1. Sysoev V. V., Bykov A. Yu. Zashchita avtorskogo prava s ispol'zovaniem cifrovoj [Copyright Protection using Digital Holography]. Sbornik trudov XIII vserossijskoj nauchno-tekhnicheskoj konferencii «Bezopasnye informacionnye tekhnologii». M.: MGTU im. N. E. Baumana. 2024. Pp. 207–213
2. Makarenko, S. I. Etalonnaya model' vzaimodeystviya steganograficheskikh sistem i obosnovanie na ee osnove novykh napravleniy razvitiya teorii steganografii [The Reference Model of the Interaction of Steganographic Systems and the Justification based on it of New Directions in the Development of the Theory of Steganography]. Voprosy kiberbezopasnosti. 2014. No. 2(3). Pp. 24–32.
3. Li D., Yang Z., Jin X. Zero watermarking scheme for 3D triangle mesh model based on global and local geometric features. Multimed. Tools Appl. 2023. Vol. 82. P. 43635–43648.
4. Chen L. et al. IW-NeRF: Using Implicit Watermarks to Protect the Copyright of Neural Radiation Fields. Applied Sciences. 2024. Vol. 14. No. 6184. DOI:10.3390/app14146184.

6 Valentin V. Sysoev, post-graduate student, Department of Information Security, BMSTU, Moscow, Russia. E-mail: valsus88@mail.ru
7 Alexander Yu. Bykov, Ph.D. of technical Sciences, associate professor, Department of Information Security, BMSTU, Moscow, Russia. E-mail: abykov@bmstu.ru

5. Sun W. et al. RWNeRF: Robust Watermarking Scheme for Neural Radiance Fields Based on Invertible Neural Networks. Computers. Materials & Continua. 2024. Vol. 80. Pp. 4065–4083. DOI:10.32604/cmc.2024.053115.
6. Luo Z., Guo Q., Cheung K.C., See S., Wan R. CopyRNeRF: Protecting the CopyRight of Neural Radiance Fields. 2023. DOI:10.48550/arXiv.2307.11526.
7. Li C., Feng B.Y., Fan Z., Pan P., Wang Z. StegaNeRF: Embedding Invisible Information within Neural Radiance Fields. 2022. DOI:10.48550/arXiv.2212.01602.
8. Sivachev A. V., Prokhozhev N. N., Mikhailichenko O. V., Bashmakov D. A. Effektivnost' steganoanaliza na osnove metodov mashinnogo obucheniya [Effectiveness of Steganalysis based on Machine Learning Methods]. Voprosy kiberbezopasnosti. 2017. No. 2(20). Pp. 53–-60. DOI 10.21581/2311-3456-2017-2-53-60.
9. Glinskaya, E. V., Chichvarin N. V. Informatsionnaya bezopasnost' otkrytykh kanalov peredachi proektnoy dokumentatsii, produciruemoy v SAPR [Information Security of open Transmission Channels of Project Documentation produced in CAD]. Voprosy kiberbezopasnosti. 2014. No. 4(7). Pp. 11–22.
10. Abasova, A. M., Babenko L. K. Zashchita informatsionnogo soderzhaniya izobrazheniy v usloviyakh nalichiya destruktivnogo vozdeystviya [Protection of Information Content of Images in the Presence of Destructive Influence]. Voprosy kiberbezopasnosti. 2019. No. 2(30). Pp. 50–57. DOI 10.21681/2311-3456-2019-2-50-57.
11. Kozachok, A. V., Kopylov S. A., Bochkov M. V. Otsenka parametrov neobnaruzhaemosti razrabotannogo podkhoda k markirovaniyu tekstovykh elektronnykh dokumentov [Evaluation of the Undetectability Parameters of the Developed Approach to Labeling Textual Electronic Documents]. Voprosy kiberbezopasnosti. 2020. No. 1(35). Pp. 62–73. DOI 10.21681/2311-3456-2020-01-62-73.
12. Morkovin, S. V. Algoritmy i programmnye sredstva cheloveko-mashinnoy obrabotki tsifrovykh vodyanykh znakov v videoposledovatel'nosti [Algorithms and software tools for human-computer processing of digital watermarks in video sequences]. Modeling, Optimization and Information Technology. 2022. Vol. 10. No. 3(38). Pp. 30–31. DOI: 10.26102/2310-6018/2022.38.3.024.
13. Kryvoshaev I. A., Linnik M. A. Staticheskij sposob steganograficheskogo vstraivaniya informacii na osnove LSB [Static Steganographic Information Embedding Method Based on LSB]. Sistemy i sredstva informatiki [Systems and Means of Informatics]. 2020. Vol. 30. No. 3. P. 56–66. DOI 10.14357/08696527200306.
14. Brūzgienė Rasa et al. Enhancing Steganography Through Optimized Quantization Tables. Electronics. 2024. Vol. 13. No. 2415. DOI: 10.3390/electronics13122415.
15. Binmin P., Qiao T. et al. Novel Hidden Bit Location Method Towards JPEG Steganography. Secur. Commun. Netw. 2022. Vol. 2022. No. 8230263. P. 13. DOI: 10.1155/2022/8230263.
16. Kramarenko S. M. Sposob vneseniya tsifrovykh metok v tsifrovoje izobrazheniye i ustroystvo dlya osushchestvleniya sposoba [Method for embedding digital marks into digital images and device for implementing this method]. 2020. Patent RU2739936C1.
17. Zhigalov I. E., Ozerova M. I., Evstigneev A. V. Application of Cutter-Jordan-Bossen method for data hiding in the image spatial domain. Bulletin of South Ural State University. Series: Computer Technologies, Control, Radioelectronics. 2022. Vol. 23. Pp. 16–23. DOI: 10.14529/ctcr230302.
18. Kushnerevich, P. M. Analiz effektivnosti algoritmov Patchwork i LSB dlya zashchity graficheskikh obrazov s pomoshch'yu vodyanykh znakov [Efficiency Analysis of Patchwork and LSB Algorithms for Graphical Image Protection by Watermarking]. Proceedings of the International Scientific Practical Conference «Integration of Science, Society, Production and Industry: Problems and Prospects». Volgograd. Russia. 2021. Pp. 121–124.

МЕТОДИКА ЭКСПЕРТНО-АНАЛИТИЧЕСКОГО АНАЛИЗА ТЕХНИКО ЭКОНОМИЧЕСКОЙ ЭФФЕКТИВНОСТИ СИСТЕМЫ ОБЕСПЕЧЕНИЯ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ ПРЕДПРИЯТИЯ НА ОСНОВЕ СРАВНЕНИЯ С «ЛУЧШИМИ ПРАКТИКАМИ»

Гайдамакин Н. А.[1]

DOI: 10.21681/2311-3456-2025-5-149-161

Цель статьи: *рассмотреть методы анализа эффективности систем обеспечения информационной безопасности предприятий и сформировать методику экспертно-аналитического анализа их технико-экономической эффективности на основе сравнения «с лучшими практиками».*

Метод исследования: *применение методов анализа эффективности IT-сферы предприятий на основе принципов «совокупной стоимости владения».*

Результаты исследования: *рассмотрены и проанализированы проблемы двух основных подходов к анализу эффективности систем обеспечения информационной безопасности предприятий – риск-ориентированного и технико-экономического.*

На основе анализа технико-экономической эффективности по принципу «совокупной стоимости владения» в сфере информационных технологий проведена систематизация расходов (затрат) на обеспечение информационной безопасности предприятия в виде двух-уровневой иерархической схемы – капитальные затраты (по стоимости приобретения и установки средств технической защиты информации и средств обеспечения безопасности информационных технологий, затрат на проведение организационно-технических и организационно-штатных мероприятий, амортизационные потери), эксплуатационные затраты (на заработную плату и аутсорсинг, на сопровождение и техническое обслуживание, на обучение персонала, на предупредительно-профилактические мероприятия в виде аудита, пентестинга, тренировок и учений), затраты и потери, связанные с результативно-целевой стороной системы обеспечения информационной безопасности (потери от простоя корпоративной информационной системы в результате компьютерных инцидентов, затраты на ее восстановление, потери рабочего времени на организационно-технологические процедуры защиты информации в виде времени прохождения процедур идентификации и аутентификации, блокирования автоматизированных рабочих мест в результате неправильных действий пользователей).

Представлена целевая функция технико-экономической эффективности системы обеспечения информационной безопасности на основе взвешенного суммирования показателей эффективности по составляющим представленной схемы затрат, сравниваемых с «лучшими практиками» или среднестатистическими значениями по отрасли предприятия.

Сформирована методика анализа технико-экономической эффективности систем обеспечения информационной безопасности предприятий на основе представленной целевой функции и применения метода экспертных оценок для учета специфики предприятий по особенностям IT-инфраструктуры, бизнес-политики и политики информационной безопасности. По сформированной методике приведен иллюстративный пример результатов анализа технико-экономической эффективности системы обеспечения информационной безопасности предприятия.

Научная новизна: *проведена систематизация расходов, затрат и потерь на обеспечение информационной безопасности в методологии «совокупной стоимости владения», предложена суперпозиционная целевая функция и основанная на ней экспертно-аналитическая методика анализа технико-экономической эффективности систем обеспечения информационной безопасности.*

Ключевые слова: *система обеспечения информационной безопасности, эффективность, совокупная стоимость владения, риск-ориентированной анализ, технико-экономический анализ, экспертно-аналитический анализ, затраты на обеспечение информационной безопасности.*

1 Гайдамакин Николай Александрович, доктор технических наук, профессор, Уральский федеральный университет имени первого Президента России Б. Н. Ельцина, г. Екатеринбург, Россия, E-mail: n.a.gaidamakin@urfu.ru

Введение

Система обеспечения информационной безопасности (СОИБ) является неотъемлемым элементом организационно-технологической инфраструктуры любого современного предприятия и по обобщенному представлению на основе национальных стандартов 27-й серии[2], требований регуляторов[3] и других нормативно-методических документов[4] включает совокупность сил и средств, мероприятий (процессов) и локальной нормативной базы.

Соответственно в бюджет любого предприятия входит раздел расходов на создание и функционирование СОИБ и для собственника (руководства) предприятия возникает необходимость анализа (оценки) эффективности соответствующих расходов.

Под эффективностью в общем смысле понимается достижение каких-либо определенных результатов с минимальными возможными издержками, приближение к максимальному (наиболее желательному) результату при минимальных негативных последствиях и возможных издержках. Во многих случаях понятие «эффективность» связано с понятием «качество» и широко используется в сфере управления (менеджмента) качеством, которая регламентируется международными стандартами серии 9000. В ГОСТ Р ИСО 9000[5] приводится следующее «классическое» определение понятия «эффективность» – соотношение между достигнутым результатом и использованными ресурсами.

Данное определение применимо к самым разнообразным сферам деятельности. Но при этом главными проблемами являются определение понятия «результат» в соответствующей сфере деятельности (и соответственно его измерение, подсчёт) и понятия «использованные ресурсы».

Эти проблемы в полной мере относятся к оценке эффективности СОИБ.

Как и во многих других случаях, «результат деятельности» в сфере обеспечения ИБ выражается качественными понятиями – «обеспечение информационной безопасности», «нейтрализация угроз безопасности информации», «недопущение инцидентов», «своевременная ликвидация последствий инцидентов» и т.д.

Общепринятый подход экономической оценки эффективности, основанный на т.н. «принципе возврата инвестиций» [1, 2], требует определения получаемой прибыли от создания и функционирования СОИБ (экономический эффект от затрат на ИБ), что, как и в отношении других участков, обеспечивающих основную деятельность предприятия, является трудно формализуемой задачей.

Определенные трудности возникают также и с понятием и оценкой затрат на создание и функционирование СОИБ. Конечно, часть затрат имеет объективную учетную основу в системе финансово-экономического документирования деятельности предприятия. Однако в отношении таких затрат, расходов или потерь, связанных с целевой результативностью СОИБ, как, например, потери рабочего времени сотрудников из-за необходимости выполнения процедур ИБ (идентификации и аутентификации), блокирование доступа к их АРМ в результате неправильных действий или ошибок по процедурам ИБ и т.п., требуется введение сложной и затратной системы фиксации соответствующих ситуаций и событий, введение дополнительных показателей внутренней отчетности, что, в свою очередь, вызывает дополнительные расходы.

Очевидно, из-за сложностей с формализацией соответствующих понятий в сфере ИБ в настоящее время нет нормированного определения (в стандартах) понятия «эффективность обеспечения ИБ», «эффективность системы обеспечения информационной безопасности» и несмотря на большую совокупность работ по анализу эффективности обеспечения ИБ [3–9] отсутствуют общепринятые модели и методики оценки эффективности СОИБ.

1. Основные методы оценки эффективности систем обеспечения информационной безопасности

Можно выделить два подхода (метода) к анализу эффективности обеспечения ИБ – «угрозный» или иначе «риск-ориентированный»[6] [8–10] и «технико-экономический»[7] [11, 12].

В первом подходе в качестве результата обеспечения ИБ рассматривается снижение или нейтрализация угроз безопасности информации (снижение их вероятности, потока угроз, их интенсивности, «опасности») или производных от этих понятий, в частности, т.н. «риска» R от угроз безопасности информации. Соответственно эффективность СОИБ $Эфф_{СОИБ}$ рассматривается как соотношение

2 ГОСТ Р ИСО/МЭК 27001–2021 «Информационная технологии. Методы и средства обеспечения безопасности. Системы менеджмента информационной безопасности. Требования», ГОСТ Р ИСО/МЭК 27002–2021 «Информационная технологии. Методы и средства обеспечения безопасности. Свод норм и правил менеджмента информационной безопасности», Проект ГОСТ Р «Защита информации. Система организации и управления защитой информации. Общие положения».

3 Методический документ «Методика оценки состояния защиты информации в органе (организации)». Утвержден ФСТЭК России 02.05.2024.

4 Стандарт Банка России СТО БР ИББС–1.0–2014 «Обеспечение информационной безопасности организаций банковской системы Российской Федерации. Общие положения».

5 ГОСТ Р ИСО 9000–2015 «Системы менеджмента качества. Основные положения и словарь».

6 ГОСТ Р ИСО/МЭК 27005––2010 «Информационная технология. Методы и средства обеспечения безопасности. Менеджмент риска информационной безопасности».

7 См. также: Петренко С. А., Симонов С. В. Управление информационными рисками. Экономически оправданная безопасность. – М.: Компания АйТи; ДМК Пресс. 2004. 384 с.

величины уменьшения риска ΔR от угроз безопасности к стоимости затрат:

$$Эфф_{СОИБ} = \frac{R - R_{СОИБ}}{C_{СОИБ}}, \quad (1)$$

где $R_{СОИБ}$ – «остаточный» риск от угроз безопасности информации в результате создания и функционирования СОИБ; $C_{СОИБ}$ – стоимость создания и функционирования СОИБ.

В целях нормирования показателя эффективности $Эфф_{СОИБ}$ в знаменатель соотношения (1) добавляется стоимость $C_{ИТ/СОИБ}$ всей IT-системы предприятия (IT-инфраструктуры) без учета стоимости СОИБ:

$$Эфф_{СОИБ} = \frac{R - R_{СОИБ}}{C_{ИТ/СОИБ} + C_{СОИБ}}. \quad (2)$$

При этом под «риском» R понимается математическое ожидание возможного ущерба от реализации угрозы безопасности:

$$R = P_{возн}\, C_{объект}, \quad (3)$$

где $P_{возн}$ – вероятность возникновения угрозы безопасности; $C_{объект}$ – стоимость объектов защиты, на которые воздействует угроза, точнее говоря, величина потерь от реализации воздействия угрозы безопасности на соответствующие объекты.

Вероятность возникновения какой-либо угрозы безопасности информации определяется целым рядом трудно формализуемых факторов, характеризующих источник угрозы (злоумышленника, нарушителя ИБ) и объект воздействия, среди которых в 1-ю очередь следует выделить мотивацию источника *Мотив*, его подготовленность *Подг*, оснащенность *Оснащ* и осведомленность *Освед* (о СОИБ). В результате при гипотезе о независимости действия источников угроз безопасности вероятность $P_{возн_i}$ возникновения i-й угрозы определяется следующим соотношением:

$$P_{возн_i} = 1 - \prod_{k=1}^{K_i} \left(1 - P_{ист_{ik}}(Мотив_{ik}, Подг_{ik}, Оснащ_{ik}, Освед_{ik})\right), \quad (4)$$

где $P_{ист_{ik}}(Мотив_{ik}, Подг_{ik}, Оснащ_{ik}, Освед_{ik})$ – функция (функционал) определения вероятности возникновения i-й угрозы от k-го источника, в зависимости от его мотивации $Мотив_{ik}$, подготовленности $Подг_{ik}$, оснащенности $Оснащ_{ik}$ и осведомленности $Освед_{ik}$; K_i – количество идентифицированных источников i-й угрозы.

Риск от возникновения i-й угрозы:

$$R_i = \sum_{m=1}^{M} C_{объект_m} \delta^{(угр)}_{im} \times \left(1 - \prod_{k=1}^{K_i}(1 - P_{ист_{ik}}(Мотив_{ik}, Подг_{ik}, Оснащ_{ik}, Освед_{ik})\right), \quad (5)$$

где M – количество объектов воздействия угроз безопасности (объектов защиты); $C_{объект_m}$ – стоимость m-го объекта защиты; $\delta^{(угр)}_{im}$ – индикатор воздействия i-й угрозы на m-й объект защиты $\delta^{(угр)}_{im} = 1$, если i-я угроза воздействует на m-й объект защиты, и $\delta^{(угр)}_{im} = 0$ в противном случае).

Риск от возникновения всех угроз безопасности:

$$R = \sum_{i=1}^{I}\sum_{m=1}^{M} C_{объект_m}\delta^{(угр)}_{im} \times \left(1 - \prod_{k=1}^{K_i}(1 - P_{ист_{ik}}(Мотив_{ik}, Подг_{ik}, Оснащ_{ik}, Освед_{ik})\right), \quad (6)$$

где I – количество идентифицированных угроз безопасности информации.

Остаточный риск $R_{СОИБ}$ определяется как математическое ожидание ущерба от вероятности реализации угроз безопасности с учетом функционирования СОИБ. Вероятность реализации угрозы безопасности рассматривается как произведение вероятности ее возникновения $P_{возн}$ на вероятность преодоления $P_{преод}$ ею системы защиты, т.е. СОИБ:

$$R_{СОИБ} = P_{возн}\, P_{преод}\, C_{объект}. \quad (7)$$

Вероятность $P_{преод_i}$ преодоления i-й угрозой СОИБ определяется вероятностью ее нейтрализации $P_{СЗИ}$ установленными средствами технической защиты информации (СТЗИ, СЗИ, СОБИТ[8]), которая в свою очередь определяется их техническими характеристиками *Техн*:

$$P_{преод_i} = 1 - \prod_{j=1}^{J}\left(1 - P_{СЗИ_{ij}}(Техн_{ij})\right), \quad (8)$$

где J – количество СЗИ; $Техн_{ij}$ – технические характеристики j-го СЗИ, влияющие на вероятность нейтрализации i-й угрозы; $P_{СЗИ_{ij}}$ – функция (функционал) определения вероятности нейтрализации i-й угрозы j-м СЗИ.

Остаточный риск $R_{СОИБ}$ от реализации всех идентифицированных угроз:

$$R_{СОИБ} = \sum_{i=1}^{I}\sum_{m=1}^{M} C_{объект_m}\delta^{(угр)}_{im} \times \left(1 - \prod_{j=1}^{J}\left(1 - P_{СЗИ_{ij}}(Техн_{ij})\right)\right) \times \left(1 - \prod_{k=1}^{K_i}(1 - P_{ист_{ik}}(Мотив_{ik}, Подг_{ik}, Оснащ_{ik}, Освед_{ik})\right). \quad (9)$$

8 Из анализа нормативных правовых актов основного регулятора в сфере ИБ – ФСТЭК России, следует, что СТЗИ (СЗИ) – это программно-аппаратные, программные или специальные технические средства, сертификация которых осуществляется на основе утвержденных профилей защиты или специальных требований к данному виду средств по обеспечению защиты информации, издаваемых регулятором. Средства обеспечения безопасности информационных технологий (СОБИТ) – это программно-аппаратные или программные средства, применяемые в системах обеспечения безопасности корпоративных сетей, но для которых нет профилей защиты (не разработаны) или нет специальных требований ФСТЭК России (не изданы). Сертификация СОБИТ осуществляется по требованиям доверия, установленным приказом ФСТЭК России от 02.06.2020 № 76 «Требования по безопасности информации, устанавливающие уровни доверия к средствам технической защиты информации и средствам обеспечения безопасности информационных технологий».

Определение вида функций (точнее функционалов) $P_{\text{ист}_{ik}}$, $P_{\text{СЗИ}_{ij}}$, как и их аргументов (функций) *Мотив, Подг, Оснащ, Освед, Техн* представляет собой трудно формализуемую задачу, поэтому на практике используются экспертные оценки [13], различные эвристики и порядковые (качественные) шкалы оценивания, а также аппарат нечетких множеств [14–17] как в отношении самих аргументов *Мотив, Подг, Оснащ, Освед, Техн*, так величин $P_{\text{возн}}$, $P_{\text{преод}}$, $P_{\text{СЗИ}}$. В качестве примера можно привести попытку создания количественно-качественной методики оценки т.н. «потенциала» нарушителя ИБ, которая была представлена в проекте «Методики определения угроз безопасности информации в информационных системах» ФСТЭК России в 2015 г.

Отдельную и также трудно-формализуемую проблему составляет оценка стоимости объектов защиты $C_{\text{объект}_m}$, поскольку в большинстве случаев даже эвристическим образом не удается полностью идентифицировать и формализовать все факторы, влияющие (определяющие) на стоимость объектов защиты, в т.ч. с учетом потерь, например, репутационных, в результате воздействия на объект защиты угроз безопасности. В итоге опять-таки для оценивания $C_{\text{объект}_m}$ широко применяются экспертные оценки, порядковые шкалы и в т.ч. аппарат нечетких множеств.

Несмотря на указанные проблемы риск-ориентированные подходы к оценке эффективности систем защиты информации широко используются на практике и реализованы в методиках и поддерживающих их программно-инструментальных средствах ряда известных зарубежных и отечественных ИБ компаний – методики CRAMM (Великобритания), RickWatch (США), ГРИФ (Россия)[9]. В некоторых последних исследованиях вопросов современного состояния информационной безопасности предприятий[10] отмечается, что 89 % компаний проводят оценку критических рисков ИБ, но только 5% компаний ориентируются на оценку рисков при бюджетировании ИБ.

Технико-экономический подход к анализу эффективности СОИБ базируется на принципах и методике анализа эффективности сферы IT-предприятий [18, 19][11], разработанных в конце 80-х – середине 90-х годов известной исследовательской консалтинговой компанией Gartner Group, и включает определение и анализ двух ключевых параметров по СОИБ – совокупной стоимости владения (*TCO* – total cost of ownership) и возврата инвестиций (*ROI* – return of investment).

Наиболее показательным в смысле экономической эффективности является параметр *ROI*, определяемый как отношение суммы прибыли или убытков к сумме инвестиций (затрат). Вместе с тем, как уже отмечалось, вычленить степень непосредственного влияния функционирования СОИБ на прибыль предприятия в структуре инвестиций среди других «обеспечивающих» направлений деятельности (внедрение новых технологий и оборудования, организационно-технологические изменения и т.д.) практически невозможно.

Подход *TCO* включает анализ первоначальных и эксплуатационных, прямых и косвенных расходов на СОИБ ($TCO_{\text{СОИБ}}$) обычно в процентах по отношению к *TCO* всего предприятия ($TCO_{\text{Предпр}}$). Такой подход, прежде всего, обеспечивает понимание (представление) относительной величины и картины расходов на ИБ предприятия. При этом с учетом того, что затраты на ИБ предприятия в целом рассматриваются как необходимые (неизбежные) издержки (потери) бизнеса, то в качестве критерия эффективности СОИБ принимается снижение (минимизация) этих издержек или сравнение с т.н. «лучшими практиками» (best practices), т.е. со средними $TCO_{\text{СОИБ}}$ предприятий такой же сферы, такого же масштаба. По критерию сравнения с «лучшими практиками, чем ближе $TCO_{\text{СОИБ}}$ предприятия к средним показателям $TCO_{\text{СОИБ}}$ по отрасли, тем больше эффективность СОИБ.

Практическое применение методик анализа эффективности СОИБ на основе *TCO* включает, прежде всего, определение структуры расходов и создание системы их фиксации в финансово-производственной отчетности и документации предприятия (организации).

2. Обобщенная структура затрат на систему обеспечения информационной безопасности

В известной работе Петренко С. А. и Симонова С. В. «Управление информационными рисками. Экономически оправданная безопасность»[12] выделена следующая структура т.н. «систематических» затрат на функционирование СОИБ (после того, как она создана), состоящая из 3-х групп – затраты на обслуживание СОИБ (управление СОИБ, регламентное обслуживание СЗИ, аудит, обучение персонала и т.д.), затраты на контрольные мероприятия (плановые и внеплановые испытания и проверки, пересмотр политики безопасности и моделей угроз, контрольные мероприятия внешних структур, обусловленные регламентацией и надзорной деятельностью регуляторов и т.д.), затраты на ликвидацию

9 См.: Современные методы и средства анализа и контроля рисков информационных систем компаний. [Электронный ресурс]. URL: https://www.ixbt.com/cm/informationsystem-risks012004.shtml (Дата обращения 12.08.25).

10 Инвестиции в информационную безопасность России. Июнь 2025 г. [Электронный ресурс]. URL: https://www.csr.ru/upload/iblock/448/yzkej5xg955rh52v2eyakztjekt0rqc7.pdf (Дата обращения 12.08.25).

11 См. также: Аншина М. Л. Оценка эффективности ИТ // Системы управления бизнес-процессами". [Электронный ресурс]. URL: https://journal.itmane.ru/node/591?ysclid=me29t5rk6t214544958 (Дата обращения 12.08.25).

12 Петренко С. А., Симонов С. В. Управление информационными рисками. Экономически оправданная безопасность. – М.: Компания АйТи, ДМК Пресс. 2004. 384 с.

последствий нарушения ИБ (восстановление IT-инфраструктуры и СОИБ, их элементов в результате инцидентов ИБ, ликвидация последствий инцидентов ИБ).

Обобщая подходы к категориям и перечням затрат на ИБ, отталкиваясь в т.ч. от методики расчета TCO в IT-сфере компании Gartner Group и учитывая современные подходы к анализу экономической эффективности IT-сферы [18, 19], можно сформировать следующую систему затрат на СОИБ, определяющих $TCO_{\text{СОИБ}}$.

1. ***Капитальные затраты*** (первоначальные и текущие) на создание и развитие СОИБ – $TCO_{\text{СОИБ}}^{(\text{Кап})}$, включая:
   - затраты (первоначальные и текущие) на приобретение и установку СТЗИ-СОБИТ – $TCO_{\text{СОИБ}}^{(\text{Кап})_{\text{СТЗИ}}}$;
   - затраты (первоначальные и текущие) по организационно-техническим и организационно-штатным мероприятиям на создание и развитие СОИБ (на получение лицензий/разрешений, аттестацию по требования ИБ корпоративной информационной системы и/или ее элементов (ИС, АС. АСУТП), на создание ИБ-подразделений и/или штатных должностей по ИБ в других подразделениях, на возможный реинжиниринг бизнес-процессов в связи с созданием или модернизацией СОИБ) – $TCO_{\text{СОИБ}}^{(\text{Кап})_{\text{Оргтех}}}$;
   - амортизационные потери по СОИБ (за счет амортизации стоимости СТЗИ-СОБИТ, устаревания и/или потери перспективности тех или иных видов приобретенных СТЗИ-СОБИТ и технологий, снижения их надежности и соответственно их стоимости в балансе предприятия) – $TCO_{\text{СОИБ}}^{(\text{Кап})_{\text{Аморт}}}$.
2. ***Эксплуатационные затраты*** на СОИБ (годовой бюджет ИБ без учета затрат на приобретение и установку СТЗИ-СОБИТ, на восстановление и ликвидацию последствий по инцидентам) – $TCO_{\text{СОИБ}}^{(\text{Экспл})}$, включая:
   - годовой бюджет на содержание штатных подразделений ИБ и должностей ИБ в других подразделениях (заработная плата персонала, другие расходы без учета стоимости приобретения СТЗИ-СОБИТ и обучения персонала по ИБ), в т.ч. затраты на аутсорсинг ИБ (функции SOC[13] внешними структурами и т.д.)[14] – $TCO_{\text{СОИБ}}^{(\text{Экспл})_{\text{ЗП}}}$;
   - затраты на сопровождение, регламентное техническое обслуживание, ремонт СТЗИ-СОБИТ (программных и программно-аппаратных СТЗИ-СОБИТ, составляющих программно-техническую основу СОИБ)[15] – $TCO_{\text{СОИБ}}^{(\text{Экспл})_{\text{Сопр}}}$;
   - затраты на обучение (получение квалификационных категорий, повышение квалификации) персонала предприятия по вопросам ИБ (подразделений ИБ и других подразделений) в специализированных внешних и/или внутренних учебных структурах – $TCO_{\text{СОИБ}}^{(\text{Экспл})_{\text{Обуч}}}$;
   - затраты на предупредительно-профилактические мероприятия ИБ (аудит ИБ, в т.ч. активный аудит внешними структурами, т.е. проведение тестирования на проникновение, тренировки/учения по ИБ, проводимые внешними структурами) – $TCO_{\text{СОИБ}}^{(\text{Экспл})_{\text{Предупр}}}$.
3. ***Затраты (потери), связанные с целевой результативностью СОИБ*** (из-за инцидентов ИБ, ситуаций и процедур ИБ) – $TCO_{\text{СОИБ}}^{(\text{Рез})}$:
   - потери предприятия из-за остановки функционирования корпоративной информационной системы (ИС, АС, АСУТП, компьютерной сети, отдельных АРМ) в результате инцидентов ИБ – $TCO_{\text{СОИБ}}^{(\text{Рез})_{\text{Остан}}}$;
   - затраты на восстановление функционирования корпоративной информационной системы (ее элементов) после инцидентов ИБ и ликвидацию других негативных последствий (репутационных, нормативных, социальных) – $TCO_{\text{СОИБ}}^{(\text{Рез})_{\text{Вост}}}$;
   - стоимость (издержки) из-за потерь рабочего времени сотрудников и должностных лиц предприятия, связанных с административно-технологическими процедурами и ситуациями ИБ (время получения/изменения учетной записи и/или прав доступа, идентификационных атрибутов, т.е. паролей, электронных ключей в ИС, АС, АСУТП корпоративной информационной системы[16], время блокирования учетных записей пользователей в результате несоблюдения или игнорирования ими требований и процедур ИБ, их небрежности, невнимательности и т.п.) – $TCO_{\text{СОИБ}}^{(\text{Рез})_{\text{РабВр}}}$.

13 Security Operation Center.

14 Во многих источниках отдельно выделяют затраты на планирование мероприятий ИБ, на контроль состояния ИБ, на поддержание соответствия нормативным требованиям регуляторов, на пересмотр политик/угроз ИБ, локальной нормативной базы ИБ. Вместе с тем данные работы в основной своей массе выполняются сотрудниками подразделений ИБ и входят в группу затрат на их содержание (заработная плата персонала, годовые затраты на деятельность соответствующих подразделений без учета стоимости приобретения СЗИ-СОБИТ), т.е. отражаются в годовом бюджете по расходам, связанным с подразделением ИБ.

15 Причем как внешними организациями, так и внутренними структурными подразделениями (в рамках составляющих их бюджета по соответствующим расходам).

16 Следует отметить, что вообще говоря данные издержки следует относить к TCO_{IT} IT-сферы предприятия, поскольку они связаны с эффективностью деятельности системных администраторов, входящих в состав не ИБ, а IT-подразделений, но во многих случаях в соответствующих процедурах участвуют администраторы безопасности, что является основанием для учета соответствующих издержек и в $TCO_{\text{СОИБ}}$.

Таким образом:

$$TCO_{\text{СОИБ}} = TCO_{\text{СОИБ}}^{(\text{Кап})} + TCO_{\text{СОИБ}}^{(\text{Экспл})} + TCO_{\text{СОИБ}}^{(\text{Рез})} =$$
$$= TCO_{\text{СОИБ}}^{(\text{Кап})_{\text{СТЗИ}}} + TCO_{\text{СОИБ}}^{(\text{Кап})_{\text{Оргтех}}} + TCO_{\text{СОИБ}}^{(\text{Кап})_{\text{Аморт}}} +$$
$$+ TCO_{\text{СОИБ}}^{(\text{Экспл})_{\text{ЗП}}} + TCO_{\text{СОИБ}}^{(\text{Экспл})_{\text{Сопр}}} + TCO_{\text{СОИБ}}^{(\text{Экспл})_{\text{Обуч}}} +$$
$$+ TCO_{\text{СОИБ}}^{(\text{Экспл})_{\text{Предупр}}} + TCO_{\text{СОИБ}}^{(\text{Рез})_{\text{Остан}}} + TCO_{\text{СОИБ}}^{(\text{Рез})_{\text{Вост}}} +$$
$$TCO_{\text{СОИБ}}^{(\text{Рез})_{\text{РабВр}}}. \quad (10)$$

Представленная структура затрат на ИБ предприятия охватывает все составляющие системы обеспечения информационной безопасности, наглядна как для технических специалистов, так и для финансово-экономических работников, и дает полную детализированную картину расходов и затрат на ИБ для собственника (руководства) предприятия.

Важно отметить, что представленная структура затрат включает количественные показатели, которые в большинстве своем, как это показано работах Лукацкого А. В.[17], можно объективно фиксировать в системе внутриотчетной статистики и финансово-экономической документации предприятия.

Также следует отметить различие применяемого в анализах и обзорах по состоянию обеспечения ИБ на предприятиях и в организациях понятия «бюджет ИБ» $BD_{\text{ИБ}}$ от $TCO_{\text{СОИБ}}$ по соотношению (10). Как правило, бюджет ИБ включает эксплуатационные затраты на СОИБ (в нашем случае – $TCO_{\text{СОИБ}}^{(\text{Экспл})}$ и затраты на приобретение новых СТЗИ-СОБИТ, а также затраты по организационно-техническим мероприятиям на развитие СОИБ, осуществляемые в текущем году (в нашем случае – $TCO_{\text{СОИБ}}^{(\text{Кап})_{\text{СТЗИ}_{\text{Текущ}}}}$ и $TCO_{\text{СОИБ}}^{(\text{Кап})_{\text{Оргтех}_{\text{Текущ}}}}$.

Таким образом бюджет предприятия на ИБ:

$$BD_{\text{ИБ}} = TCO_{\text{СОИБ}}^{(\text{Кап})_{\text{СТЗИ}_{\text{Текущ}}}} + TCO_{\text{СОИБ}}^{(\text{Кап})_{\text{Оргтех}_{\text{Текущ}}}} + TCO_{\text{СОИБ}}^{(\text{Экспл})}. \quad (11)$$

Кроме того, в целях нивелирования различий предприятий по масштабу бюджет ИБ $BD_{\text{ИБ}}$ как правило анализируется не столько в абсолютных значениях, а в относительной доле бюджета на IT-сферу предприятия.

Детальное рассмотрение параметров (метрик), отражающих, прежде всего результативно-целевую сторону создания и функционирования СОИБ ($TCO_{\text{СОИБ}}^{(\text{Рез})_{\text{Остан}}}$, $TCO_{\text{СОИБ}}^{(\text{Рез})_{\text{Вост}}}$, $TCO_{\text{СОИБ}}^{(\text{Рез})_{\text{РабВр}}}$) представлено в отмеченных работах Лукацкого А. В. и работе Филиппова М. Г.[18]

17 Лукацкий А. В. Как посчитать эффективность информационной безопасности? [Электронный ресурс]. URL: https://www.cisco.com/c/dam/global/ru_ua/assets/securityforum/presentations/10_security_measurement.pdf (Дата обращения 12.08.25).

18 Филиппов М. Г. Сколько стоит безопасность. Анализ процессов обеспечения ИБ в российских компаниях. [Электронный ресурс]. URL: https://pt-corp.storage.yandexcloud.net/upload/corporate/ru-ru/analytics/IS-Cost-rus.pdf (Дата обращения 12.08.25).

3. Целевая функция и экспертно аналитическая методика анализа технико экономической эффективности СОИБ по принципу сравнения с «лучшими практиками»

Оценка технико-экономической эффективности СОИБ предприятия на основе количественного сравнения $TCO_{\text{СОИБ}}$ с «лучшими практиками» или усредненными по отрасли значениями несмотря на свою логичность, информативность и наглядность, тем не менее характеризуется рядом проблем.

Во-первых, в интегральном показателе $TCO_{\text{СОИБ}}$ нивелируется «вклад» различных составляющих в эффективность СОИБ предприятия, поэтому требуется сравнивать с «лучшими практиками» или средними по отрасли показателями не только интегральный показатель $TCO_{\text{СОИБ}}$, но и осуществлять сравнительный анализ всех его слагаемых – $TCO_{\text{СОИБ}}^{(\text{Кап})}$, $TCO_{\text{СОИБ}}^{(\text{Экспл})}$, $TCO_{\text{СОИБ}}^{(\text{Рез})}$, и, в свою очередь, их составляющих – $TCO_{\text{СОИБ}}^{(\text{Кап})_{\text{СТЗИ}}}$, $TCO_{\text{СОИБ}}^{(\text{Кап})_{\text{Оргтех}}}$, $TCO_{\text{СОИБ}}^{(\text{Кап})_{\text{Аморт}}}$, $TCO_{\text{СОИБ}}^{(\text{Экспл})_{\text{ЗП}}}$, $TCO_{\text{СОИБ}}^{(\text{Экспл})_{\text{Сопр}}}$, $TCO_{\text{СОИБ}}^{(\text{Экспл})_{\text{Обуч}}}$, $TCO_{\text{СОИБ}}^{(\text{Экспл})_{\text{Предупр}}}$, $TCO_{\text{СОИБ}}^{(\text{Рез})_{\text{Остан}}}$, $TCO_{\text{СОИБ}}^{(\text{Рез})_{\text{Вост}}}$, $TCO_{\text{СОИБ}}^{(\text{Рез})_{\text{РабВр}}}$.

Во-вторых, количественная оценка эффективности СОИБ по принципу сравнения» с «лучшими практиками» или усредненными значениями может осуществляется различным образом, самым очевидным из которых является использование относительных значений[19] –

$$\left(\frac{TCO_{\text{СОИБ}}}{TCO_{\text{Предпр}}}\right) \Big/ \overline{\left(\frac{TCO_{\text{СОИБ}}}{TCO_{\text{Предпр}}}\right)}_{\text{Отрасл}}$$

или с учетом существенных различий $TCO_{\text{Предпр}}$ по отраслям, масштабу и специфике предприятий на основе сравнения с TCO_{IT} IT-сферы (IT-инфраструктуры) предприятия –

$$\left(\frac{TCO_{\text{СОИБ}}}{TCO_{\text{IT}}}\right) \Big/ \overline{\left(\frac{TCO_{\text{СОИБ}}}{TCO_{\text{IT}}}\right)}_{\text{Отрасл}}.$$

В этом случае наибольшая эффективность СОИБ предприятия характеризуется единичными значениями соответствующих отношений, но величина отклонения от единицы в большую или меньшую стороны трудно поддается интерпретации в качестве метрики эффективности.

В-третьих, как абсолютные, так и относительные значения $TCO_{\text{СОИБ}}$ и его составляющих по соотношению (10) не могут отражать специфику предприятия, в т.ч. одной отрасли, с точки зрения особенностей внутренней и внешней среды, т.н. «зрелости» бизнес-процессов, инфраструктуры и т.д. Иначе говоря отклонение $TCO_{\text{СОИБ}}$ от среднего значения на 20 % для одного (специфичного) предприятия может означать высокую эффективность СОИБ, а для другого (типового) недостаточную эффективность.

19 *Обозначение $(\overline{x})_{\text{отрасл}}$ означает среднее значение величины x по отрасли.*

В-четвертых, несмотря на то, что представленная структура затрат на СОИБ по соотношению (10), как отмечалось, включает количественные показатели, все же в отношении таких составляющих, как стоимость ликвидации репутационных (нормативных, социальных) негативных последствий ($ТСО^{(Рез)_{Вост_{Репутац}}}_{СОИБ}$) от компьютерных инцидентов, имеется существенная неопределенность в количественных методиках их подсчета. В этом же плане имеются трудности и неопределенности с определением величины $ТСО^{(Рез)_{РабВр}}_{СОИБ}$, поскольку несмотря на показанную Лукацким А. В. возможность создания систем внутренней отчетности, которые фиксируют и учитывают потери рабочего времени из-за процедур ИБ (блокирование учетных записей пользователей в результате несоблюдения или игнорирования ими требований и процедур ИБ и соответствующие обращения в подразделения ИБ или Service Desk для их разрешения и т.п.), все же часть потерь рабочего времени связана с самостоятельным решением (попытками) пользователями соответствующих проблем и трудно фиксируется в системах учета и отчетности.

Разрешением отмеченных проблем может стать использование экспертных оценок в процессах сравнения составляющих $ТСО_{СОИБ}$ с «лучшими практиками» или средними по отрасли значениями с учетом специфики конкретного предприятия и введение на этой основе суперпозиционной целевой функции технико-экономической эффективности СОИБ.

Суть представляемого подхода заключается в следующем.

В качестве целевой функции $ТЭЭ_{СОИБ}$ технико-экономической эффективности СОИБ рассматривается взвешенное суммирование показателей технико-экономической эффективности по составляющим затрат на СОИБ, установленным соотношением (10):

$$ТЭЭ_{СОИБ} = c_{(Кап)} ТЭЭ^{(Кап)}_{СОИБ} + c_{(Экспл)} ТЭЭ^{(Экспл)}_{СОИБ} + c_{(Рез)} ТЭЭ^{(Рез)}_{СОИБ} = c_{(Кап)} \left(c_{(Кап)_{СТЗИ}} ТЭЭ^{(Кап)_{СТЗИ}}_{СОИБ} + c_{(Кап)_{Оргтех}} ТЭЭ^{(Кап)_{Оргтех}}_{СОИБ} + c_{(Кап)_{Аморт}} ТЭЭ^{(Кап)_{Аморт}}_{СОИБ} \right) + c_{(Экспл)} \left(c_{(Экспл)_{ЗП}} ТЭЭ^{(Экспл)_{ЗП}}_{СОИБ} + c_{(Экспл)_{Сопр}} ТЭЭ^{(Экспл)_{Сопр}}_{СОИБ} + c_{(Экспл)_{Обуч}} ТЭЭ^{(Экспл)_{Обуч}}_{СОИБ} + c_{(Экспл)_{Предупр}} ТЭЭ^{(Экспл)_{Предупр}}_{СОИБ} \right) + c_{(Рез)} \left(c_{(Рез)_{Остан}} ТЭЭ^{(Рез)_{Остан}}_{СОИБ} + c_{(Рез)_{Вост}} ТЭЭ^{(Рез)_{Вост}}_{СОИБ} + c_{(Рез)_{РабВр}} ТЭЭ^{(Рез)_{РабВр}}_{СОИБ} \right), \quad (12)$$

где

- $ТЭЭ^{(Кап)}_{СОИБ}$, $ТЭЭ^{(Экспл)}_{СОИБ}$, $ТЭЭ^{(Рез)}_{СОИБ}$, $ТЭЭ^{(Кап)_{СТЗИ}}_{СОИБ}$, $ТЭЭ^{(Кап)_{Оргтех}}_{СОИБ}$, $ТЭЭ^{(Кап)_{Аморт}}_{СОИБ}$, $ТЭЭ^{(Экспл)_{ЗП}}_{СОИБ}$, $ТЭЭ^{(Экспл)_{Сопр}}_{СОИБ}$, $ТЭЭ^{(Экспл)_{Обуч}}_{СОИБ}$, $ТЭЭ^{(Экспл)_{Предупр}}_{СОИБ}$, $ТЭЭ^{(Рез)_{Остан}}_{СОИБ}$, $ТЭЭ^{(Рез)_{Вост}}_{СОИБ}$, $ТЭЭ^{(Рез)_{РабВр}}_{СОИБ}$ – величины технико-экономической эффективности СОИБ по составляющим структуры затрат на СОИБ первого и второго уровня детализации, соответственно;
- $(c_{(Кап)} + c_{(Экспл)} + c_{(Рез)} = 1)$ – весовые коэффициенты относительной значимости влияния капитальных, эксплуатационных и результативно-целевых составляющих затрат на интегральный показатель технико-экономической эффективности СОИБ $ТЭЭ_{СОИБ}$;
- $(c_{(Кап)_{СТЗИ}} + c_{(Кап)_{Оргтех}} + c_{(Кап)_{Аморт}}) = 1)$ – весовые коэффициенты относительной значимости влияния составляющих капитальных затрат на $ТЭЭ^{(Кап)}_{СОИБ}$;
- $(c_{(Экспл)_{ЗП}} + c_{(Экспл)_{Сопр}} + c_{(Экспл)_{Обуч}} + c_{(Экспл)_{Предупр}}) = 1$ – весовые коэффициенты относительной значимости влияния составляющих эксплуатационных затрат на $ТЭЭ^{(Экспл)}_{СОИБ}$;
- $(c_{(Рез)_{Остан}} + c_{(Рез)_{Вост}} + c_{(Рез)_{РабВр}})$ – весовые коэффициенты относительной значимости влияния составляющих затрат, связанных с результативно-целевой стороной СОИБ, на $ТЭЭ^{(Рез)}_{СОИБ}$.

Введение целевой функции (12) позволяет сформировать следующую Методику экспертно-аналитического анализа технико-экономической эффективности системы обеспечения информационной безопасности предприятия по принципу сравнения с «лучшими практиками».

1. На основе внутренней статистической отчетности и финансово-экономической документации предприятия определяются количественные значения составляющих капитальных, эксплуатационных затрат и затрат, связанных с результативно-целевой стороной по функционированию СОИБ, – $ТСО^{(Кап)}_{СОИБ}$, $ТСО^{(Экспл)}_{СОИБ}$, $ТСО^{(Рез)}_{СОИБ}$, $ТСО^{(Кап)_{СТЗИ}}_{СОИБ}$, $ТСО^{(Кап)_{Оргтех}}_{СОИБ}$, $ТСО^{(Кап)_{Аморт}}_{СОИБ}$, $ТСО^{(Экспл)_{ЗП}}_{СОИБ}$, $ТСО^{(Экспл)_{Сопр}}_{СОИБ}$, $ТСО^{(Экспл)_{Обуч}}_{СОИБ}$, $ТСО^{(Экспл)_{Предупр}}_{СОИБ}$, $ТСО^{(Рез)_{Остан}}_{СОИБ}$, $ТСО^{(Рез)_{Вост}}_{СОИБ}$, $ТСО^{(Рез)_{РабВр}}_{СОИБ}$.

2. На основе экспертных оценок специалистов, хорошо представляющих специфику деятельности предприятия, ее IT-сферы (IT-инфраструктуры), приоритеты бизнеса и общую политику безопасности предприятия, включая политику информационной безопасности, определяются весовые коэффициенты $c_{(Кап)}$, $c_{(Экспл)}$, $c_{(Рез)}$, $c_{(Кап)_{СТЗИ}}$, $c_{(Кап)_{Оргтех}}$, $c_{(Кап)_{Аморт}}$, $c_{(Экспл)_{ЗП}}$, $c_{(Экспл)_{Сопр}}$, $c_{(Экспл)_{Обуч}}$, $c_{(Экспл)_{Предупр}}$, $c_{(Рез)_{Остан}}$, $c_{(Рез)_{Вост}}$, $c_{(Рез)_{РабВр}}$.

3. Эксперты на основе сравнения количественных значений составляющих $ТСО_{СОИБ}$ (см. п. 1) с аналогичными значениями «лучших практик» или со средними значениями по отрасли соответствующих показателей и учитывая специфику предприятия с точки зрения особенностей внутренней и внешней среды, «зрелости» бизнес-процессов и инфраструктуры и т.д. в порядково-вербальной шкале Харрингтона[20] (см. табл. 1) выставляют качественные

20 Мацкевич А. А. Управленческие шкалы. Часть 1. Шкалы показателей без планового значения // Управление предприятием. [Электронный ресурс]. URL: https://upr.ru/article/upravlencheskie-shkaly-chast-1-shkaly-pokazateley-bez-planovogo-znacheniya/ (Дата обращения 12.08.25).

(порядковые) оценки соответствующих показателей технико-экономической эффективности СОИБ[21] – $[\widehat{ТЭЭ}]^{(Кап)_{СТЗИ}}_{СОИБ}$, $[\widehat{ТЭЭ}]^{(Кап)_{Оргтех}}_{СОИБ}$, $[\widehat{ТЭЭ}]^{(Кап)_{Аморт}}_{СОИБ}$, $[\widehat{ТЭЭ}]^{(Экспл)_{ЗП}}_{СОИБ}$, $[\widehat{ТЭЭ}]^{(Экспл)_{Сопр}}_{СОИБ}$, $[\widehat{ТЭЭ}]^{(Экспл)_{Обуч}}_{СОИБ}$, $[\widehat{ТЭЭ}]^{(Экспл)_{Предупр}}_{СОИБ}$, $[\widehat{ТЭЭ}]^{(Рез)_{Остан}}_{СОИБ}$, $[\widehat{ТЭЭ}]^{(Рез)_{Вост}}_{СОИБ}$, $[\widehat{ТЭЭ}]^{(Рез)_{РабВр}}_{СОИБ}$.

Другим вариантом может быть использование шкалы экспертного сравнения альтернатив Т. Саати[22] (см. табл. 2). Эксперты в контексте технико-экономической эффективности СОИБ сравнивают две альтернативы – количественные значения соответствующих составляющих $ТСО_{СОИБ}$ предприятия (первая альтернатива) и усредненных значений по отрасли соответствующих значений составляющих $ТСО_{СОИБ}$ (вторая альтернатива).

4. Формируются усредненные экспертные оценки составляющих $ТЭЭ_{СОИБ}$ в порядковой шкале путем взятия медианы ряда индивидуальных экспертных оценок, выстроенных в неубывающей последовательности[23] – $\left([\widehat{ТЭЭ}]^{(Кап)_{СТЗИ}}_{СОИБ}\right)_{Усред}$, $\left([\widehat{ТЭЭ}]^{(Кап)_{Оргтех}}_{СОИБ}\right)_{Усред}$, $\left([\widehat{ТЭЭ}]^{(Кап)_{Аморт}}_{СОИБ}\right)_{Усред}$, $\left([\widehat{ТЭЭ}]^{(Экспл)_{ЗП}}_{СОИБ}\right)_{Усред}$, $\left([\widehat{ТЭЭ}]^{(Экспл)_{Сопр}}_{СОИБ}\right)_{Усред}$, $\left([\widehat{ТЭЭ}]^{(Экспл)_{Обуч}}_{СОИБ}\right)_{Усред}$, $\left([\widehat{ТЭЭ}]^{(Экспл)_{Предупр}}_{СОИБ}\right)_{Усред}$, $\left([\widehat{ТЭЭ}]^{(Рез)_{Остан}}_{СОИБ}\right)_{Усред}$, $\left([\widehat{ТЭЭ}]^{(Рез)_{Вост}}_{СОИБ}\right)_{Усред}$, $\left([\widehat{ТЭЭ}]^{(Рез)_{РабВр}}_{СОИБ}\right)_{Усред}$.

5. На основе соотношения порядковых и балльных оценок, установленных в шкале Харрингтона или шкале Т. Саати, усредненные экспертные оценки составляющих технико-экономической эффективности СОИБ в порядковой шкале переводятся в числовую шкалу [0,1] по шкале Харрингтона (см. табл. 1) или в бальные оценки по шкале Т. Саати[24] (см. табл. 2) – $\left(\overline{ТЭЭ}^{(Кап)_{СТЗИ}}_{СОИБ}\right)_{Усред}$, $\left(\overline{ТЭЭ}^{(Кап)_{Оргтех}}_{СОИБ}\right)_{Усред}$, $\left(\overline{ТЭЭ}^{(Кап)_{Аморт}}_{СОИБ}\right)_{Усред}$, $\left(\overline{ТЭЭ}^{(Экспл)_{ЗП}}_{СОИБ}\right)_{Усред}$, $\left(\overline{ТЭЭ}^{(Экспл)_{Сопр}}_{СОИБ}\right)_{Усред}$, $\left(\overline{ТЭЭ}^{(Экспл)_{Обуч}}_{СОИБ}\right)_{Усред}$, $\left(\overline{ТЭЭ}^{(Экспл)_{Предупр}}_{СОИБ}\right)_{Усред}$, $\left(\overline{ТЭЭ}^{(Рез)_{Остан}}_{СОИБ}\right)_{Усред}$, $\left(\overline{ТЭЭ}^{(Рез)_{Вост}}_{СОИБ}\right)_{Усред}$, $\left(\overline{ТЭЭ}^{(Рез)_{РабВр}}_{СОИБ}\right)_{Усред}$.

6. По соотношению (12) на основе весового суммирования с соответствующими весовыми коэффициентами по количественным усредненным экспертным оценкам, полученным по п. 5, определяются величины технико-экономической эффективности СОИБ в разрезе ее трех составляющих – $\overline{ТЭЭ}^{(Кап)}_{СОИБ}$, $\overline{ТЭЭ}^{(Экспл)}_{СОИБ}$, $\overline{ТЭЭ}^{(Рез)}_{СОИБ}$ и интегральный показатель $\overline{ТЭЭ}_{СОИБ}$.

С учетом единичной нормировки весовых коэффициентов максимальное значение величин $\overline{ТЭЭ}^{(Кап)}_{СОИБ}$, $\overline{ТЭЭ}^{(Экспл)}_{СОИБ}$, $\overline{ТЭЭ}^{(Рез)}_{СОИБ}$ и $\overline{ТЭЭ}_{СОИБ}$ в шкале Харрингтона равняется единице (100 %), в шкале Т. Саати равняется девяти (100 %).

Таблица 1.

Шкала Харрингтона

Порядково-вербальное значение оцениваемой величины	Числовое значение в шкале [0,1]
Очень высокое	[0,8 – 1]
Высокое	[0,64 – 0,8]
Среднее	[0,37 – 0,64]
Низкое	[0,2 – 0,37]
Очень низкое	[0 – 0,2]

Следует отметить, что на стратегическом уровне анализа эффективности деятельности предприятия оценки интегральных показателей технико-экономической эффективности $\overline{ТЭЭ}^{(Кап)}_{СОИБ}$, $\overline{ТЭЭ}^{(Экспл)}_{СОИБ}$, $\overline{ТЭЭ}^{(Рез)}_{СОИБ}$ и $\overline{ТЭЭ}_{СОИБ}$ помимо количественной (бальной, процентной) шкалы могут требоваться и обратно в порядко-вербальной шкале. Такие оценки формируются обратным преобразованием по шкалам Харрингтона или Т. Саати.

Проиллюстрируем на примере гипотетического предприятия результаты экспертно-аналитической оценки технико-экономической эффективности СОИБ по представленной методике[25], которые для наглядности сведены в табл. 3.

Приведенный пример-иллюстрация наглядно раскрывает «картину» составляющих технико-экономической эффективности СОИБ и особенности экспертно-аналитического характера представленной методики.

В первую очередь следует отметить необходимость создания и ведения баз данных по используемым в методике среднестатистическим показателям –

$$\left(\overline{ТСО}^{(Кап)_{СТЗИ}}_{СОИБ}\right)_{Отрасл}, \left(\overline{ТСО}^{(Кап)_{Оргтех}}_{СОИБ}\right)_{Отрасл}, \left(\overline{ТСО}^{(Кап)_{Аморт}}_{СОИБ}\right)_{Отрасл}, \left(\overline{ТСО}^{(Экспл)_{ЗП}}_{СОИБ}\right)_{Отрасл}, \left(\overline{ТСО}^{(Экспл)_{Сопр}}_{СОИБ}\right)_{Отрасл}, \left(\overline{ТСО}^{(Экспл)_{Обуч}}_{СОИБ}\right)_{Отрасл}, \left(\overline{ТСО}^{(Экспл)_{Предупр}}_{СОИБ}\right)_{Отрасл}, \left(\overline{ТСО}^{(Рез)_{Остан}}_{СОИБ}\right)_{Отрасл}, \left(\overline{ТСО}^{(Рез)_{Вост}}_{СОИБ}\right)_{Отрасл}, \left(\overline{ТСО}^{(Рез)_{РабВр}}_{СОИБ}\right)_{Отрасл}$$

в разрезе разных отраслей предприятий и организаций. Насколько известно, такие базы данных, но в отношении $ТСО_{IT}$, ведут известные консалтинго-исследовательские в сфере IT-компании, в частности Gartner Group. На российском рынке исследованиями

21 Обозначение $[\hat{x}]$ в данном случае означает оценку величины x в порядковой (вербальной) шкале Харрингтона.

22 Саати Т. Принятие решений. Метод анализа иерархий. – М.: Радио и связь, 1993. 278 с.

23 Усреднение индивидуальных экспертных оценок в порядковой шкале путем взятия медианы их неубывающего ряда является корректной математической операцией усреднения. Пример: пусть 7 экспертов величине x дали следующие индивидуальные оценки в порядковой шкале – $[\hat{x}]_1$ = «низкая», $[\hat{x}]_2$ = «высокая», $[\hat{x}]_3$ = «средняя», $[\hat{x}]_4$ = «низкая», $[\hat{x}]_5$ = «низкая», $[\hat{x}]_6$ = «низкая», $[\hat{x}]_7$ = «средняя»; выстраивая индивидуальные оценки в неубывающий ряд – («низкая», «низкая», «низкая», «низкая», «средняя», «средняя», «высокая») по его медиане (середине) получаем усредненную экспертную оценку в порядковой шкале $[\hat{x}]_{усред}$ = «низкая».

24 Обозначение $\hat{x}$ означает оценку величины x в количественной (интервальной) шкале значений.

25 Среднестатистические показатели в табл. 3 не имеют отношения к реальной действительности, являются гипотетическими и используются исключительно для иллюстрации. Все весовые коэффициенты для слагаемых целевой функции $ТЭЭ_{СОИБ}$ одинаковы, т.е. равны 1/3 или 1/4 для коэффициентов $\overline{ТЭЭ}^{(Экспл)}_{СОИБ}$.

Таблица 2.

Шкала сравнения альтернатив Т. Саати

Порядково-вербальная оценка предпочтительности альтернативы	Соответствующая балльная оценка	Пояснение (квалиметрическая характеристика)
Абсолютное превосходство (гораздо хуже)	9 (1/9)	Альтернатива абсолютно (неоспоримо) предпочтительнее, весь опыт эксперта и многочисленная практика в высшей степени убедительно свидетельствуют об этом
Значительное превосходство (значительно хуже)	7 (1/7)	Альтернатива значительно (убедительно, совершенно очевидно) предпочтительнее, опыт эксперта и известная эксперту практика свидетельствуют об этом
Существенное (сильное) превосходство (хуже)	5 (1/5)	Альтернатива существенно (явно) предпочтительнее, имеются свидетельства и надежные данные в пользу этого решения
Среднее (слабое) превосходство (чуть хуже)	3 (1/3)	Альтернатива незначительно (немного) предпочтительнее, имеется некоторый опыт и некоторые свидетельства в пользу такого решения, но они недостаточно убедительны
Несравнимы	0	По анализируемому свойству альтернативы несравнимы (разной природы) или эксперт затрудняется в сравнении альтернатив, поскольку не имеется никакого опыта, никаких свидетельств в пользу превосходства одной альтернативы над другой
Промежуточные решения	2, 4, 6, 8 (1/2, 1/4, 1/6, 1/8)	

Таблица 3.

Иллюстративный пример экспертно-аналитической оценки технико экономической эффективности СОИБ

Частные показатели эффективности СОИБ		Среднестатистические показатели затрат на СОИБ по группе однородных предприятий	Показатели СОИБ конкретного предприятия	Усредненная экспертная оценка эффективности СОИБ по соответствующим показателям на основе сравнения со среднестатистическими (с «best practices»)	
				Порядковые оценки по шкале Т. Саати	Баллы по шкале Т. Саати
1		2	3	4	5
$\widehat{ТЭЭ}^{(Кап)}_{СОИБ}$	По капитальным затратам на создание СОИБ				
	$\widehat{ТЭЭ}^{(Кап)_{СТЗИ}}_{СОИБ}$	По затратам на приобретение и установку СТЗИ-СОБИТ, % от стоимости программно-технических средств КИС (корпоративная информационная система)			
		15 %	20 %	Существенно лучше	5
	$\widehat{ТЭЭ}^{(Кап)_{Оргтех}}_{СОИБ}$	По затратам на организационно-технические и организационно-штатные мероприятия в рамках создания СОИБ, % от общих расходов предприятия			
		3 %	1 %	Значительно хуже	1/7
	$\widehat{ТЭЭ}^{(Кап)_{Аморт}}_{СОИБ}$	По годовым показателям амортизации стоимости СТЗИ-СОБИТ и технологий СОИБ (% ежегодного уменьшения стоимости основных средств, устаревание технологий)			
		30 %	25 %	Чуть лучше	3
ИТОГО (баллы) $\widehat{ТЭЭ}^{(Кап)}_{СОИБ} = c_{(Кап)_{СТЗИ}} \widehat{ТЭЭ}^{(Кап)_{СТЗИ}}_{СОИБ} + c_{(Кап)_{Оргтех}} \widehat{ТЭЭ}^{(Кап)_{Оргтех}}_{СОИБ} + c_{(Кап)_{Аморт}} \widehat{ТЭЭ}^{(Кап)_{Аморт}}_{СОИБ}$					**2,7**
ИТОГО (%) $\widehat{ТЭЭ}^{(Кап)}_{СОИБ}$ = **30 %** (от максимальной или оптимальной по отрасли) ИТОГО (по шкале Саати) $[\widehat{ТЭЭ}]^{(Кап)}_{СОИБ}$ = «**Средняя**» (относительно оптимальной по отрасли)					

1		2	3	4	5
$\widehat{ТЭЭ}^{(Экспл)}_{СОИБ}$	По эксплуатационным затратам на СОИБ				
	$\widehat{ТЭЭ}^{(Экспл)_{ЗП}}_{СОИБ}$	По заработной плате персонала подразделений СОИБ, включая затраты на аутсорсинг ИБ (SOC и т.д. кроме пентестинга), % от ФОТ предприятия			
		3 %	5 %	Существенно лучше	5
	$\widehat{ТЭЭ}^{(Экспл)_{Сопр}}_{СОИБ}$	По затратам на сопровождение, регламентное техническое обслуживание, ремонт СТЗИ-СОБИТ, % от стоимости СТЗИ-СОБИТ			
		12 %	20 %	Значительно лучше	7
	$\widehat{ТЭЭ}^{(Экспл)_{Обуч}}_{СОИБ}$	По затратам на обучение персонала предприятия по вопросам ИБ, % от годового дохода предприятия			
		0,035 %	0,03 %	Средне	3
	$\widehat{ТЭЭ}^{(Экспл)_{Предупр}}_{СОИБ}$	Затраты на предупредительно-профилактические мероприятия ИБ (аудит, пентестинг, учения/тренировки), % от бюджета на ИБ (от $BD_{ИБ}$)			
		10 %	15 %	Значительно лучше	7
ИТОГО (баллы) $\widehat{ТЭЭ}^{(Экспл)}_{СОИБ} = c_{(Экспл)_{ЗП}} \widehat{ТЭЭ}^{(Экспл)_{ЗП}}_{СОИБ} + c_{(Экспл)_{Сопр}} \widehat{ТЭЭ}^{(Экспл)_{Сопр}}_{СОИБ} + c_{(Экспл)_{Обуч}} \widehat{ТЭЭ}^{(Экспл)_{Обуч}}_{СОИБ} + c_{(Экспл)_{Предупр}} \widehat{ТЭЭ}^{(Экспл)_{Предупр}}_{СОИБ}$					**5,5**
ИТОГО (%) $\widehat{ТЭЭ}^{(Экспл)}_{СОИБ}$ = **61 %** (от максимальной или оптимальной по отрасли) ИТОГО (по шкале Саати) $[\widehat{ТЭЭ}]^{(Экспл)}_{СОИБ}$ = «**Значительно лучше средней**» (по отрасли)					
$\widehat{ТЭЭ}^{(Рез)}_{СОИБ}$	По затратам/потерям в отношении результативно-целевой составляющей СОИБ				
	$\widehat{ТЭЭ}^{(Рез)_{Остан}}_{СОИБ}$	По потерям из-за остановки функционирования корпоративной информационной системы (ИС, АС, АСУТП, компьютерной сети, отдельных АРМ) в результате инцидентов ИБ, % от годового дохода предприятия			
		2 %	1 %	Существенно лучше	5
	$\widehat{ТЭЭ}^{(Рез)_{Вост}}_{СОИБ}$	По затратам на восстановление функционирования корпоративной информационной системы после инцидентов ИБ и ликвидацию других негативных последствий (репутационных, нормативных, социальных), % от годового дохода предприятия			
		1 %	1,5 %	Существенно хуже	1/5
	$\widehat{ТЭЭ}^{(Рез)_{РабВр}}_{СОИБ}$	Потери рабочего времени сотрудников предприятия, связанные с административно-технологическими процедурами ИБ (из-за блокирования учетных записей и т.п.), % от годового бюджета рабочего времени предприятия			
		2 %	1 %	Значительно лучше	7
ИТОГО (баллы) $\widehat{ТЭЭ}^{(Рез)}_{СОИБ} = c_{(Рез)_{Остан}} \widehat{ТЭЭ}^{(Рез)_{Остан}}_{СОИБ} + c_{(Рез)_{Вост}} \widehat{ТЭЭ}^{(Рез)_{Вост}}_{СОИБ} + c_{(Рез)_{РабВр}} \widehat{ТЭЭ}^{(Рез)_{РабВр}}_{СОИБ}$					**4,07**
ИТОГО (%) $\widehat{ТЭЭ}^{(Рез)}_{СОИБ}$ = **45 %** (от максимальной или оптимальной по отрасли) ИТОГО (по шкале Саати) $[\widehat{ТЭЭ}]^{(Рез)}_{СОИБ}$ = «**Значительно лучше средней**» (по отрасли)					
ИТОГО по общей (интегральной) технико-экономической эффективности СОИБ предприятия, баллы: $\widehat{ТЭЭ}_{СОИБ} = c_{(Кап)} \widehat{ТЭЭ}^{(Кап)}_{СОИБ} + c_{(Экспл)} \widehat{ТЭЭ}^{(Экспл)}_{СОИБ} + c_{(Рез)} \widehat{ТЭЭ}^{(Рез)}_{СОИБ}$					**4,39**
ИТОГО общая (интегральная) технико-экономическая эффективность СОИБ предприятия (%): $\widehat{ТЭЭ}_{СОИБ}$ = **49 %** (от максимальной или оптимальной по отрасли)					
ИТОГО (по шкале Саати) общая (интегральная) технико-экономическая эффективность СОИБ предприятия: $[\widehat{ТЭЭ}]_{СОИБ}$ = «**Немного лучше средней**» (по отрасли)					

в сфере эффективности затрат (бюджетов) на информационную безопасность предприятий помимо отдельных известных специалистов (Лукацкий А. В.) занимается также ряд ведущих в сфере ИБ компаний (Positive Technologies, Лаборатория Касперского, InfoWatch, РТК-Солар и др.). Источниками создания и ведения таких баз данных могут быть как материалы государственной статистики, так и специализированные опросы-исследования предприятий.

Кроме того, в представленной методике принципиальным являются экспертные оценки показателей технико-экономической эффективности, осуществляемые экспертами на основе количественных сравнений.

У многих специалистов-практиков наблюдается устойчивый скепсис в отношении экспертных оценок из-за их якобы неотъемлемой субъективности. Очевидно, это происходит в результате непонимания сути и особенностей организационно-процедурной стороны экспертных оценок. При правильно сформированной экспертной группе, качественно разработанных оценочных инструментах (опросных листах), профессионально организованных процедурах опроса экспертов и корректной математической обработке результатов индивидуальных оценок процедуры экспертных оценок позволяют решать «нерешаемые» (не формализуемые или трудно-формализуемые) задачи и проблемы, обеспечивают высокую степень объективности итоговых оценок, что широко известно и активно применяется в сфере выработки и принятия управленческих решений [20].

В отношении данной методики экспертные оценки не только позволяют свести все показатели технико-экономической эффективности СОИБ в единую шкалу, но при сравнении количественных показателей предприятия и среднестатистических отраслевых показателей обеспечивают учет существенной во многих случаях специфики или наоборот типичности инфраструктуры конкретного предприятия, кадрового обеспечения, бизнес-политики и политики ИБ.

Заключение

Представленные целевая функция и экспертно-аналитическая методика анализа технико-экономической эффективности СОИБ базируются на классической структуре широко известного и применяемого технико-экономического показателя «TCO» (total cost of ownership), но специфицированного в отношении обеспечивающего характера сферы ИБ в отношении в свою очередь обеспечивающей основную деятельность предприятия IT сферы. Исходя из этого, общий смысл и детализация показателей технико-экономической эффективности СОИБ, как представляется, должны быть хорошо понимаемыми и наглядными для руководителей и финансово-экономических работников предприятий, что является, в свою очередь информативной основой для руководителей СОИБ при формировании и «отстаивании» бюджетов ИБ.

Часть показателей $TCO_{\text{СОИБ}}$ второго уровня детализации фиксируется и ведется в системе внутренней отчетности и финансово-экономической документации предприятий – $TCO_{\text{СОИБ}}^{(\text{Кап})_{\text{СТЗИ}}}$, $TCO_{\text{СОИБ}}^{(\text{Экспл})_{\text{ЗП}}}$, $TCO_{\text{СОИБ}}^{(\text{Экспл})_{\text{Сопр}}}$, $TCO_{\text{СОИБ}}^{(\text{Экспл})_{\text{Обуч}}}$, $TCO_{\text{СОИБ}}^{(\text{Экспл})_{\text{Предупр}}}$. Другая часть ($TCO_{\text{СОИБ}}^{(\text{Рез})_{\text{Остан}}}$, $TCO_{\text{СОИБ}}^{(\text{Рез})_{\text{Вост}}}$, $TCO_{\text{СОИБ}}^{(\text{Рез})_{\text{РабВр}}}$), как это показано в обстоятельных работах Лукацкого А. В., может формироваться в специально создаваемой на предприятии системе учета в сфере функционирования СОИБ.

Вместе с тем, как уже указывалось, требуется создание и ведение баз данных по используемым в методике среднестатистическим показателям $TCO_{\text{СОИБ}}$ в разрезе отраслей предприятий и организаций, что может быть одной из специализаций консалтингово-исследовательских компаний в сфере ИБ.

Литература

1. Паршина И. С. Рентабельность инвестиций (ROI) в проекты внедрения исполнительных производственных систем (MES) на российских предприятиях // Наукоемкие технологии в машиностроении. 2020. № 3 (105). С. 37–43.
2. Zelezinskii et al. Modern Methods of Evaluating the Effectiveness of the Organization // Экономический вектор 2021. № 4(27). pp. 65–70.
3. Zegzhda D. P., Saurenko T. N., Anisimov V. G., Anisimov E. G. Assessment of the Effectiveness of an Information Security System // Automatic Control and Computer Sciences. 2023. Volume 57, № 8. pp. 855–861. https://doi.org/10.3103/S0146411623080345.
4. Митяков Е. С., Артемова С. В., Бакаев А. А., Душкин А. В., Вегера Ж. Г. Модель оценки эффективности систем защиты информации // Безопасность информационных технологий. 2024. Том 31, № 4. С. 56–66. doi: 10.26583/bit.2024.4.03.
5. Belov V., Belova N., Pestunova T., Kosov D. Technique for Evaluating the Effectiveness of the Information Security Department. IEEE XVI International Scientific and Technical Conference Actual Problems of Electronic Instrument Engineering (APEIE). 2023. pp. 1130–1133. DOI: 10.1109/APEIE59731.2023.10347645.
6. Сухов А. М., Крупенин А. В., Якунин В. И. Метод вычисления показателя эффективности процесса функционирования системы обеспечения информационной безопасности // Автоматизация процессов управления. 2022. № 1(67). С. 33–42.
7. Добрышин М. М. Подход к формированию обобщенного критерия оценки эффективности системы обеспечения информационной безопасности // Известия тульского государственного университета. Технические науки. 2021. № 9. С. 113–121.
8. Sow M. et al. Evaluating Information Security System Effectiveness for Risk Management, Control, and Corporate Governance // Business and Economic Research. 2019. Vol. 9, № 1. pp.164–172.
9. Громов Ю. Ю., Карасев П. И., Губсков Ю. А., Котюкова В. О. Оценка эффективности систем защиты и анализ рисков информационной безопасности // Информация и безопасность. 2022. Т. 25, Вып. 2. С. 187–192.

10. Пашков Н. Н., Дрозд В. Г. Анализ рисков информационной безопасности и оценка эффективности систем защиты информации на предприятии // Современные научные исследования и инновации. 2020. № 1. [Электронный ресурс]. URL: https://web.snauka.ru/issues/2020/01/90380 (дата обращения: 26.08.2025).
11. Краузе Р. П. Исследование методических подходов к оценке эффективности ИТ-проектов на предприятиях // Бизнес-образование в экономике знаний. 2020. № 3. С.87–92.
12. Шабуров А. С., Шлыков А. И. Разработка метода оценки экономической эффективности системы защиты информации для коммерческих предприятий // Вестник ПНИПУ. Электротехника, информационные технологии, системы управления. 2020. № 36. С. 193–213.
13. Курило А. П., Паршин И. С., Симачков С. А., Потапов Г. Д. Определение эффективности комплексных систем обеспечения информационной безопасности методом экспертных оценок / Актуальные проблемы защиты информации: современность и перспективы. Материалы II Научно-практической конференции. Москва, 2025. С. 43–48.
14. Бутусов И. В., Нащекин П. А., Романов А. А. Теоретико-семантические аспекты организации комплексной системы защиты информационных систем // Вопросы кибербезопасности. 2016. № 1(14). С. 9–16.
15. Ziro A., Gnatyuk S., Toibayeva S. Investigation of the Method of Evaluating the Effectiveness of the Information Security System Based on Fuzzy Inference // Scientific Journal of Astana IT University. 2023. Volume 13. pp. 52–63. DOI: 10.37943/13dzev3953.
16. Братченко А. И., Бутусов И. В., Кобелян А. М., Романов А. А. Применение методов теории нечетких множеств к оценке рисков нарушения критически важных свойств защищаемых ресурсов автоматизированных систем управления // Вопросы кибербезопасности. 2019. № 1(29). С. 18–24.
17. Ермаков С. А., Чурсин А. Г., Болгов А. А. Нечетко-множественная методика оценки рисков автоматизированной системы «умный дом» с динамической топологией // Информация и безопасность. 2022. Т. 25. Вып. 4. С. 495–500.
18. Wojtaszek H. et al. Methods for Assessing the Economic Efficiency of IT Projects // European research studies journal. 2024. Volume XXVII (Issue 3). pp. 637–651. DOI:10.35808/ersj/3457.
19. Kasim M. K. M. et al. A systematic literature review on the effect of information systems on the performance of government officials International // Journal of Advanced and Applied Sciences, 11(3) 2024, Pages: 46–54.
20. Иванова Л. Н., Луговской В. Д. Экспертные оценки в принятии управленческих решений // Современные научные исследования и инновации. 2020. № 10 [Электронный ресурс]. URL: https://web.snauka.ru/issues/2020/10/93677 (дата обращения: 26.08.2025).

METHODOLOGY OF EXPERT-ANALYTICAL ANALYSIS OF TECHNICAL AND ECONOMIC EFFICIENCY OF THE INFORMATION SECURITY SYSTEM OF AN ENTERPRISE BASED ON COMPARISON WITH «BEST PRACTICES»

Gaydamakin N. A.[26]

Keywords: *: information security management system, effectiveness, total cost of ownership, risk-based analysis, technical and economic analysis, expert and analytical analysis, costs of ensuring information security.*

Purpose of the study: *to consider methods for analyzing the effectiveness of information security systems of enterprises and to develop a methodology for expert-analytical analysis of their technical and economic efficiency based on comparison «with best practices».*

Methods of research: *application of methods for analyzing the efficiency of the IT sphere of enterprises based on the principles of «total cost of ownership».*

Result(s): *The problems of two main approaches to the analysis of the effectiveness of information security systems of enterprises – risk-based and techno-economic - are considered and analyzed.*

Based on the analysis of technical and economic efficiency according to the principle of «total cost of ownership» in the field of information technology, a systematization of expenses (costs) for ensuring the information security of the enterprise was carried out in the form of a two-level hierarchical scheme – capital costs (according to the cost of acquiring and installing technical means of information protection and means of ensuring the security of information technologies, costs of carrying out organizational-technical and organizational-staffing measures, depreciation losses), operating costs (for wages and outsourcing, for support and technical maintenance, for personnel training, for preventive and preventive measures in the form of audit, pentesting, training and exercises), costs and losses associated with the result-target side of the information security system (losses from downtime of the corporate information system as a result of computer incidents, costs of its restoration, loss of working time on organizational and technological procedures for information protection in the form of time spent on identification and authentication procedures, blocking of automated workstations as a result of incorrect actions of users).

The objective function of technical and economic efficiency of the information security system is presented based on the weighted summation of efficiency indicators for the components of the presented cost scheme, compared with «best practices» or average statistical values for the enterprise industry.

26 Nikolay A. Gaydamakin, Dr.Sc. (of Tech.), Professor, Ural Federal Boris Yeltsin University. Yekaterinburg, Russia. E-mail: n.a.gaidamakin@urfu.ru

A methodology has been developed for analyzing the technical and economic efficiency of enterprise information security systems based on the presented objective function and the application of the expert assessment method to take into account the specifics of enterprises in terms of IT infrastructure, business policy and information security policy. According to the formed methodology, an illustrative example of the results of the analysis of the technical and economic efficiency of the information security system of the enterprise is given.

Scientific novelty: *the systematization of costs, expenses and losses for ensuring information security in the methodology of «total cost of ownership» was carried out, a superposition objective function was proposed and an expert-analytical methodology based on it for analyzing the technical and economic efficiency of information security systems was proposed.*

References

1. Parshina I. S. Rentabel'nost' investitsiy (ROI) v proyekty razvitiya ispolnitel'nykh proizvodstvennykh sistem (IPS) na rossiyskikh predpriyatiyakh // Naukoyemkiye tekhnologii v ma-shinostroyenii. 2020. № 3(105). S. 37–43.
2. Zelezinskii et al. Modern Methods of Evaluating the Effectiveness of the Organization // Экономический вектор 2021. № 4(27). С. 65–70.
3. Zegzhda D. P., Saurenko T. N., Anisimov V. G., Anisimov E. G. Assessment of the Effectiveness of an Information Security System // Automatic Control and Computer Sciences. 2023. Volume 57, № 8. Pp. 855–861. https://doi.org/10.3103/S0146411623080345.
4. Mityakov E. S., Artemova S. V., Bakaev A. A., Dushkin A. V., Vegera Zh. G. Model for assessing the effectiveness of information security systems // Information Technology Security. 2024. Vol. 31, No. 4. Pp 56–66. doi: 10.26583/bit.2024.4.03.
5. Belov V., Belova N., Pestunova T., Kosov D. Technique for Evaluating the Effectiveness of the Information Security Department. IEEE XVI International Scientific and Technical Conference Actual Problems of Electronic Instrument Engineering (APEIE). 2023. Pp. 1130–1133. DOI: 10.1109/APEIE59731.2023.10347645.
6. Sukhov A. M., Krupenin A. V., Yakunin V. I. Metod rascheta effektivnosti effektivnogo protsessa preobrazovaniya obespecheniya informatsionnoy bezopasnosti // Avtomatizatsiya protsessov upravleniya. 2022. № 1(67). S. 33–42.
7. Dobryshin M. M. Podkhod k formirovaniyu obobshchennogo kriteriya effektivnosti effektivno-sti sistemy obespecheniya informatsionnoy bezopasnosti // Izvestiya tul'skogo gosu-darstvennogo universiteta. Tekhnicheskiye nauki. 2021. № 9. S. 113–121.
8. Sow M. et al. Evaluating Information Security System Effectiveness for Risk Management, Control, and Corporate Governance // Business and Economic Research. 2019. Vol. 9, № 1. Pp. 164–172.
9. Gromov YU. YU., Karasev P. I., Gubskov YU. A., Kotyukova V. O. Otsenka effektivnosti si-stem zashchity i analiz riskov informatsionnoy bezopasnosti // Informatsiya i bezopasnost'. 2022. T. 25, Vyp 2. S. 187–192.
10. Pashkov N. N., Drozd V. G. Analiz riskov informatsionnoy bezopasnosti i otsenki effektivnosti sistem zashchity informatsii na predpriyatii // Sovremennyye nauchnyye issledovaniya i innovatsii. 2020. № 1. [Elektronnyy resurs]. URL: https://web.snauka.ru/issues/2020/01/90380 (data obrashcheniya: 26.08.2025).
11. Krauze R. P. Issledovaniye metodicheskikh podkhodov k effektivnosti effektivnosti IT proyektov na predpriyatiyakh // Biznes-obrazovaniye v ekonomike znaniy. 2020. № 3. S. 87–92.
12. Shaburov A. S., Shlykov A. I. Razrabotka metoda otsenki ekonomicheskoy effektivnosti sistemy zashchity informatsii dlya kommercheskikh predpriyatiy // Vestnik PNIPU. Elektrotekhnika, informatsionnyye tekhnologii, sistemy upravleniya. 2020. № 36. S. 193–213.
13. Kurilo A. P., Parshin I. S., Simachkov S. A., Potapov G. D. Opredeleniye effektivnosti kompleksnykh sistem informatsionnoy bezopasnosti metodom ekspertnykh otsenok / Aktual'nyye problemy zashchity informatsii: sovremennost' i perspektivy. Materialy II Nauchno-prakticheskoy konferentsii. Moskva, 2025. S. 43–48.
14. Butusov I. V., Nashchekin P. A., Romanov A. A. Teoretiko-semanticheskiye aspekty organizatsii kompleksnoy sistemy zashchity informatsionnykh sistem // Voprosy kiberbez-opasnosti. 2016. № 1(14). S. 9–16.
15. Ziro A., Gnatyuk S., Toibayeva S. Investigation of the Method of Evaluating the Effectiveness of the Information Security System Based on Fuzzy Inference // Scientific Journal of Astana IT University. 2023. Volume 13. Pp. 52–63. DOI: 10.37943/13dzev3953.
16. Bratchenko A. I., Butusov I. V., Kobelyan A. M., Romanov A. A. Metody primeneniya teorii nechetkikh mnozhestv k snizheniyu riska vozniknoveniya vazhneyshikh svoystv zashchitnykh resursov upravlencheskikh sistem upravleniya // Voprosy kiberbezopasnosti. 2019. № 1(29). S. 18–24.
17. Yermakov S. A., Chursin A. G., Bolgov A. A. Nechetko-mnozhestvennaya metodika otsenki riska dorozhnoy sistemy «umnyy dom» s dinamicheskoy topologiyey // Informatsiya i bezopasnost'. 2022. T. 25. Vyp. 4. S. 495–500.
18. Wojtaszek H. et al. Methods for Assessing the Economic Efficiency of IT Projects // European research studies journal. 2024. Volume XXVII (Issue 3). Pp :637–651. DOI:10.35808/ersj/3457.
19. Kasim M. K. M. et al. A systematic literature review on the effect of information systems on the performance of government officials International // Journal of Advanced and Applied Sciences, 11(3) 2024, Pages: 46–54.
20. Ivanova L. N., Lugovskoy V. D. Ekspertnyye otsenki v upravlencheskikh resheniyakh // Sovremennyye nauchnyye issledovaniya i innovatsii. 2020. № 10 [Elektronnyy resurs]. URL: https://web.snauka.ru/issues/2020/10/93677 (data obrashcheniya: 26.08.2025).

ADAPTIVE CUMULATIVE ENTROPY THRESHOLD: A NOVEL APPROACH TO DDOS ATTACK DETECTION IN IOT DEVICES AND SMART HOMES SYSTEMS

Amit Kumar Jaiswal[1]

DOI: 10.21681/2311-3456-2025-5-162-171

Abstract. *The rising prevalence of smart home systems in everyday life, attacks such as cyber flooding on these interconnected devices have become critical. The present research talks about the innovative model using adaptive threshold, which applies cumulative entropy analysis of time series data to detect and mitigate flood attacks more effectively in the smart home environment. The model sets dynamic thresholds adaptable to changes in data fluctuations in real-time by utilizing cumulative entropy, a measure that identifies the unpredictability and variance of network traffic patterns. Advanced machine learning techniques will be further explored to refine the threshold process that will eventually lead to higher accuracy in detecting anomalies. In fact, essential factors including temporal patterns, types of protocols, and actions of users will be analyzed concerning their impact on objective metrics. Research aims at validating proposed adaptive threshold framework effectiveness in response toward significantly reducing false positives while improving responsiveness against emerging threats; hence contributing overall resilience of smart-home systems under flood attacks towards detected attacks. Anterior work shall focus on adapting algorithms and exploring scalability over diverse smart home architectures as an extension of this work. Research also intends to tackle questions linked with data privacy as well as system efficiency.*

Keywords: *Adaptive Threshold, Cumulative Entropy, Time Series Analysis, Flood Attack Mitigation, Smart Home Security, Anomaly Detection, Network Traffic An al ysis, Temporal Data Patterns.*

1. Introduction

The increasing penetration of smart home systems in everyday life has brought about the enormous advantages of convenience and automation. However, this evolutional process also expose the shortcomings by vulnerability issues primarily focusing on possible cybersecurity threats, one of which is the Distributed Denial of Service (DDoS) attack that can immensely menace the operation of smart devices with potential risks to users' safety and privacy. Therefore, it is extremely imperative to develop efficient countermeasures to timely detect and react upon this type of cyberattack [1,2]. A promising direction towards ameliorating smart home security is to adopt adaptive thresholding techniques relying on cumulative entropy based time series analysis. Entropy, a fundamental concept taken from inFormation theories, measures uncertain or random characteristics in given dataset. In network traffic analysis domain for instance, monitoring inherent entropy regimes facilitates distinguishing normal patterns from unusual behaviors underlying DDoS attacks. By employing cumulative entropy measures, researchers can develop adaptive methods that adjust thresholds dynamically based on realtimedata, thereby improving detection accuracy and reducing false positives [3].

2. Background on Topic

In recent years, the increasing popularity of smart home systems has raised serious concerns about cybersecurity, particularly related to distributed denial of service (DDoS) attacks that can severely overflow network resources, incapacitate smart devices, and pose threats to user security and privacy [4,5]. Therefore, defense mechanisms must be put into places to ensure the resiliency of smart home systems towards this type of attack. A promising solution can be perceived by adopting adaptive thresholding techniques based on entropy measures to analyze time-series data that originate from network traffic [13,14]. Entropy is a measure that represents the uncertainty or randomness when characterizing a certain data-set. Specifically, utilizing entropy within the context of network traffic an alysis al lows re searchers to determine how different normal network requests (which are considered as non-malicious) are from exploitative counterparts illustrating DDoS traits (which are deemed malicious)[15][17]. Researchers have consistently shown how cumulative entropy measures enable detection algorithms to become more adaptive and accurate due to their capability in adjusting dynamically with respect to current network conditions [18], [19].

1 Amit Kumar Jaiswal, Ph.D., Student/Researcher, Department of Radio Engineering and Cybernetics, Moscow Institute of Physics and Technology (MIPT). E-mail: dzhaisval.a@phystech.edu

3. Related Works

Several works have studied thresholding techniques for anomaly detection innetwork traffic. For example, Sahoo and Arora (2004) proposed a thresholding technique based on two-dimensional Renyi entropy that achieved a much better segmentation performance in image processing applications, indicating the potential of entropy-based techniques to discriminate normal patterns from anomalies [8].

Dragos et al. (2020) investigated some entropy-based metrics for uncertainty evaluation in Bayesian networks designed for cyber threat detection and concluded that the entropy measurement is important both in performance estimation of a model and as an added value to decision-making under uncertainty [5].

This work paves the way for applying two-pronged on-line entropy based defense mechanism at DDoS attack by defending attack traffic in path [7].

Recent improvements in adaptive thresholding techniques show the potential of such methods in many domains. A machine learning-aided entropy-based anomaly detection framework for dynamic network adaptations was proposed by Timcenko and Gajin (2021) [6].

They elevate the need for adaptivity that relies on threshold adjustments by real-time data analysis, which is essential in combating DDoS attacks in smarthomes [9].

The use of cumulative entropy in time series analysis has been presented in some previous works. In particular, some researches have focused on using cumulative residual entropy as a risk measure and they have proven that it is a useful tool in many different situations. This is consistent with our research goal to apply cumulative entropy for adaptive thresholding in the analysis of time series data of DDoS attacks [10,11]. Zhang et al. [12] conducted a comprehensive survey on network anomaly detection frameworks based on kinds of entropy measures such as Shannon and Renyi entropies and concluded that using many kinds of features can improve the accuracy of model to against various anomalies types.

4. Detailed Raw Dataset Description Used in this Research

The UCM_FibIoT2024 dataset gathers substantial data to understand better Distributed Denial of Service (DDoS) attacks against smart home central control units, namely the Fibaro Home Center 3. This dataset records many types of DDoS assaults, such as TCP SYN floods, I CMP fl oods, an d HT TP floods, to provide light on how they influence the operation and availability of IoT devices [16]. Data was collected on a local network using the hping3 tool for SYN and ICMP flood attacks, and the LOIC tool for HTTP flood assaults. Wireshark software was used to gather network traffic, and the information is available in PCAP and CSV formats for future analysis. The collected data includes critical details such as timestamps, source and destination IP addresses, protocols, packet lengths, and port numbers [16]. The major purpose of this dataset is to make it easier to simulate and analyze DDoS attacks on smart home central control units, hence serving as a resource for cybersecurity and IoT device protection researchers. Researchers can discover attack patterns, understand the dynamics of various forms of DDoS attacks, and design effective mitigation systems by inspecting network traffic records and packet captures [16]. The collection is structured to provide comprehensive logs for each attack, such as start and finish timings, frame numbers, and the total number of assault packets. For simplicity of usage, the data is sorted into folders, and the SYN flood attack data is further split by the ports targeted (80, 443, and 500)[16]. The UCM_FibIoT2024 dataset serves as a profitable instrument for analyzing and creating resistance against DDoS attacks on IoT gadgets. It gives a viable asset to analysts and cybersecurity experts to successfully reenact, analyze, and moderate DDoS attacks [16]. For more information about the dataset, refer to the UCM_FibIoT2024 dataset is available at https:// doi.org/10.17632/p42xjtv8pv.1. However, for this study, we will be only using HTTP flood and ICMP flood data.

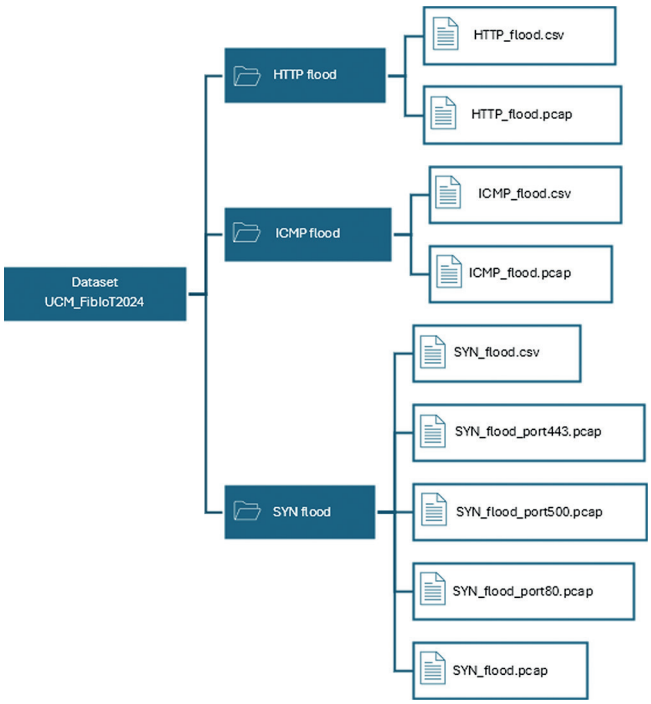


Figure 1: Flow Chart of the Raw Data Structure Files capture [17]

Table 1.

Overview of Raw CSV Dataset Columns[16]

Column Name	Description
No.	Frame number.
Time	Date and time of capture (dd.mm.yyyy hh:mm:ss).
Source	Source IP address of the packet.
Destination	Destination IP address of the packet.
Protocol	Protocol type identifying the network protocol used For each packet.
Length	Packet length in bytes.
Source port	Source port of the packet.
Destination port	Destination port of the packet.

5. Research Objective

We note here that the previous authors have utilized several thresholding techniques in different fields and datasets to research different purposes. However, no authors have used our novel approach, that is, network traffic detection with time series analysis using the cumulative entropy method with thresholding, to detect such attacks most likely in DDoS on smart home systems and on IoT devices specifically, which will ultimately help future research scope growth.

6. Methodologies Used in this research

6.1 Raw Data Preprocessing

Table 1 represents column names and descriptions of these featured columns in raw data preprocessing plays a very crucial and vital role and methods for this research. There were many steps performed on raw dataset for data preprocessing. First step was data cleaning, in this step we identified and rectified errors, inconsistencies, and inaccuracies in the raw dataset. We found source port and destination port were having huge number of missing, inconsistent and inaccurate values. We used techniques like handling missing data and removing duplicates to clean the dataset. Later we analyzed and removed source port and destination port due to their high inaccuracy and irrelevancy to this research outcome.

6.2 Exploratory Data Analysis

In this section, we will demonstrate the exploratory data analysis performed by us to analyze dataset more deeply. Firstly, we applied many data analysis codes and functions, we checked the dataset size, the dataset description in terms of counts, min and max values, and different percentiles of the dataset in each column. There are two types of attack files used, one is HTTP flood, a type of attack that targets web servers by overwhelming them with high HTTP requests. Another file is ICMP Flood, an attack that sends a large number of ICMPv6 packets (often ping requests, but in this research, data hping was used for a more aggressive attack) to a target, consuming bandwidth and resources. In Table 2, we found that HTTP flood has a higher number of packet counts but the lowest frequency, whereas ICMPv6 flood has a higher frequency in comparison to HTTP flood. These findings results in ICMPv6 floods having a higher effect in the consumption of storage and bandwidth uses, which can result in a DDoS attack in IoT devices and smart home systems.

In Table 3, we analyzed and found that IP address 10.0.1.22 has the highest number of traffic as an IP source. HTTP floods have 132 unique values whereas, ICMPv6 have the highest number of uniqueness in the traffic. As a finding, higher uniqueness in traffic are to have unknown distributed sources which is to result into a DDoS attack (Distributed Denial of service attack) in the network.

Table 2.

Summary of HTTP and ICMPv6 Flood Protocol Features

Feature	HTTP Flood	ICMPv6 Flood
Count	22,780,665	11,203,031
Unique	121	26
Top Protocol	TCP	ICMPv6
Frequency	8,060,484	11,172,532

Table 3.

Overview of HTTP and ICMPv6 Flood Source Features

Feature	HTTP Flood	ICMPv6 Flood
Count	10,799,707	11,203,031
Unique	132	11,214,481
Top Source	10.0.1.22	10.0.1.22
Frequency	4,859,392	6,602

In table 4, we analyzed and found the destination IP address request traffic. Also we found that the count of packet traffic had the same count. We discovered ICMPv6 requests, indicating a slightly higher volume compared to HTTP floods, with much higher frequency than source IP requests. Finally higher percentage of requests indicates more aggressive and sustained attack on target.

Table 4.

Overview of HTTP and ICMPv6 Flood Destination Features

Feature	HTTP Flood	ICMPv6 Flood
Count	10,799,707	11,203,031
Unique	127	1,747
Top Destination	10.0.1.22	10.0.1.22
Frequency	5,911,972	11,174,499

7. Feature Engineering

We did feature engineering task to enrich the features and applied stratified sampling technique on HTTP flood dataset and ICMPv6 flood dataset, stratified sampling, which involves dividing the population into subpopulations (strata) based on one or more common attributes; strata membership is determined by some factor(s) that are hypothesized to be related to the process being measured, such as class labels- to reduce business and increase the performance of model learning and testing.

The UCM_FibIoT2024 dataset records are enormous (millions of records), processing whole CSV file at once requires more time-consuming and computation resources. Therefore, we applied stratified random sampling to HTTP flood and ICMPv6 flood CSV fles for our experiment. We have considered the sample of frac = 0.02 for each file. In algorithm 1, we have demonstrated pseudo code representation to our code and method used on stratified sampling of high volume dataset. We have also used the time-based feature engineering method to extract each feature from the time column in separate columns (year, month day, hour, minutes, seconds, microseconds).

Algorithm 1 Sampling from Dataset by Protocol

```
BEGIN
    DATASET ← Load dataset
    GROUPED_DATASET ← GROUP DATASET BY 'Protocol'
    SAMPLED_DATA ← []
    For each GROUP in GROUPED_DATASET DO
        SAMPLE ← SAMPLE 20% FROM GROUP
        SAMPLED_DATA.APPEND(SAMPLE)
    END For
    FINAL_SAMPLED_DATA ← Convert SAMPLED_DATA TO DataFrame
END
```

7.1 Cyclical Time Encoding

Further we used Cyclical Encoding technique to create more features on time. To handle the cyclical nature of time (e.g., hours in a day, days in a week), we have converted time into a circular representation using sine and cosine functions. Let hour be the hour of the day (in 24-hour format). The angle can be calculated as:

$$\text{angle} = \left(\frac{\text{hour}}{24}\right) \times 2\pi \quad (1)$$

The sine and cosine transFormations are defined as follows:

$$X = \sin(\text{angle}) = \sin\left(\left(\frac{\text{hour}}{24}\right) \times 2\pi\right) \quad (2)$$

$$Y = \cos(\text{angle}) = \cos\left(\left(\frac{\text{hour}}{24}\right) \times 2\pi\right) \quad (3)$$

These transformations allow the model to capture the cyclical nature of time, effectively treating 23 : 00 and 00 : 00 as close to each other.

7.2 Seconds Since Epoch

The term seconds since the epoch represents the representation of time by counting the aggregate number of seconds elapsed, from a particular starting point in time, is called the epoch. Seconds since epoch are prominently used to detect and analyze Distributed Denial of Service (DDoS) attacks. Here, we highlight its usage concerning time-stamping and network traffic monitoring. In DDoS detection systems, every packet of network traffic can be timestamped in seconds since epoch format to keep the record of exactly when it was received. Accurate timestamps can be used to track trends, such as a traffic spike over an extended period of time indicating a possible DDoS attack. Using the timestamps from incoming packets detection systems can measure the number of packets or amount of traffic within a given (40 seconds) window, if too many packets arrive within that time span. Systems can compare the number of packets received in an epoch to a threshold value and then generate an alert if the packet total is above a pre-determined threshold baseline, indicative of DDoS. So, we created new feature called seconds since epoch. To do this we have created a mathematical formulae calculation to calculate seconds since epoch on each packet traffic. Let T represent the timestamp from the data sample, and let T_0 denote the epoch time defined as:

$$T_0 = \text{Timestamp}(2024, 1, 1, 0, 0, 0). \quad (4)$$

The Seconds Since Epoch can be calculated as follows:

$$\text{SecondsSinceEpoch} = (T - T_0) \div 1 \text{ second}. \quad (5)$$

Where:

- T = ['Time'] (the dataset timestamp features);
- T_0 = Timestamp representing the epoch;
- The division by 1 second effectively converts the time difference from a Timedelta object into an integer representing seconds.

The formula presented provides a clear method to calculating Seconds Since Epoch, which is fundamental in various applications involving time series analysis and event logging.

8. Mathematical Formulations For calculating entropy and detecting anomalies in packet data

8.1 Entropy Calculation

The Shannon entropy $H(X)$ for a discrete random variable X is defined as:

$$H(X) = -\sum_{i=1}^{n} p_i \log_2(p_i) \quad (6)$$

Where:

- p_i is the probability of occurrence of the i-th outcome.
- n is the total number of distinct outcomes.

In this context, the entropy is calculated for packet lengths over a rolling window of size 10:

$$H(\text{Length}_{\text{window}}) = -\sum_{j=1}^{m} p_j \log_2(p_j) \quad (7)$$

Where:

- m is the number of distinct packet lengths in the current window.

8.2 Cumulative Entropy Calculation

The cumulative entropy at time t can be expressed as:

$$C(t) = -\sum_{i=1}^{t} H(\text{Length}_i). \quad (8)$$

Where:

- $C(t)$ is the cumulative entropy up to time index t.
- $H(\text{Length}_i)$ is the entropy calculated For packet lengths at time index i.

8.3 Anomaly Detection

Anomaly detection is performed using a simple thresholding method. The threshold T is defined as:

$$T = \mu + 3\sigma. \quad (9)$$

Where:

- μ is the mean of the cumulative entropy values.
- σ is the standard deviation of the cumulative entropy values.

An anomaly occurs when:

$$A(t) = \begin{cases} 1 & \text{if } C(t) > T, \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

Where:

- $A(t)$ indicates whether an anomaly is detected at time index t.

9. Experiments

We extracted time components from time feature. We have written our own code for date-time feature with (Year, Month, Day, Hour, Minutes, Seconds, Microseconds). We used python prebuilt library called DATETIME. In algorithm 2 Label of the algorithm is "Extract Time Components from Date-time which tells us that this algorithm is responsible for extracting specific-time-related features from date-time. In the beginning of the algorithm there is a comment saying that: dataset is a data-structure (like table or Data Frame) that has a column called Time which contains date-time values. A separate loop goes throw all rows in dataset one by one and extract time column value and stores it into new created separate column named (Year, Month, Day, Hour, Minutes, Seconds, Microseconds). The algorithm is finished with an «END» statement then. In general, this algorithm is designed to extract all the possible individual time components including year, month, day, hour, minute, second and microsecond of a date-time object separately in order to analyze or process them individually. It can be very helpful for data analysis purposes when we may want to analyze/visualize some patterns at year/month/day/hour/minute/second/microsecond level or want to filter/group by these individual time components etc. while performing some machine learning tasks over timeseries like feature engineering.

We have used mathematical formulae for sine and cosine calculations for hour, for each row, it retrieves the value of hour and assigns it to hour value. Then it fetches the value of hour and stores it in an hour variable. It subsequently calculates sine and cosine of this hour value using above mentioned Formulas 1, 2 and 3 as before. By doing these calculations, it maps the respective hour into a form of cyclic representation which helps to present time-concept to the models. Sine and cosine calculations for month, similarly, it fetches the value of month and stores it in a month variable. Then, it calculates sine and cosine for month value with formulae similar to hours but divided by 12. The algorithm ends with an «END» statement representing that all calculations have been made here. This algorithm essentially performs conversion of cyclical time data (hours & months) into simple sine-cosine way. This whole code is classically inspired from https://en.wikipedia.org/wiki/Besselpublication by Don E. Knuth which approximates values of sin() & cos().

In algorithm 3, the algorithm name is «Calculate Seconds Since Epoch». The algorithm defines the constant EPOCH_TIMESTAMP as a string, representing this epoch: «2024-01-01 00:00:00». It loops through each data row assuming that there is a column with date-times Time in the dataset. For each row, it assigns the current timestamp from column Time to CURRENT_TIMESTAMP. It calculates the difference CURRENT_TIMESTAMP minus EPOCH_TIMESTAMP as TIME_DIFFERENCE. This difference represents how much time passed between the epoch and that timestamp. The algorithm converts this value then into seconds by dividing it by one second (which might be implicit for many programming languages if you handle simply date objects). The resulting number of seconds since the epoch SECONDS_SINCE_EPOCH, it saves in an additional column named SecondsSinceEpoch, defined in memory for the dataset data_sample at corresponding row. Finally, there is an «END» after which we know that all these operations end. The main aim of converting date-time values into such a standardized numeric Format (seconds since epoch) is facilitating their usage for various operations and especially mathematical analyses during which we want to help computer somehow understand how timestamps are big/small

or older/newer than other timestamps. For example when comparing them during some model learning.

Algorithm 2 Extract Time Components from Datetime

```
BEGIN
// Assume data_sample is a data structure (like a table or DataFrame) with a column 'Time' of datetime type
// Extract year from the 'Time' column
for each row in data_sample do
    row['Year'] ← EXTRACT_YEAR(row['Time'])
end for
// Extract month from the 'Time' column
for each row in data_sample do
    row['Month'] ← EXTRACT_MONTH(row['Time'])
end for
// Extract day from the 'Time' column
for each row in data_sample do
    row['Day'] ← EXTRACT_DAY(row['Time'])
end for
// Extract hour from the 'Time' column
for each row in data_sample do
    row['Hour'] ← EXTRACT_HOUR(row['Time'])
end for
// Extract minute from the 'Time' column
for each row in data_sample do
    row['Minute'] ← EXTRACT_MINUTE(row['Time'])
end for
// Extract second from the 'Time' column
for each row in data_sample do
    row['Second'] ← EXTRACT_SECOND (row['Time'])
end for
// Extract microsecond from the 'Time' column
for each row in data_sample do
    row['Microsecond'] ← EXTRACT_MICRO-SECOND(row['Time'])
end for
END
```

Algorithm 3 Calculate Seconds Since Epoch

```
BEGIN
EPOCH_TIMESTAMP ← "2024-01-01 00:00:00"
    For each row in data_sample DO
    CURRENT_TIMESTAMP ← data_sample ['Time'] [row]
        TIME_DIFFERENCE ← CURRENT_ TIMESTAMP -
        EPOCH_TIMESTAMP SECONDS_SINCE_ EPOCH ← TIME_DIFFERENCE // 1 second
        data_sample['SecondsSinceEpoch'][row] ← SECONDS_SINCE_EPOCH
    END For
END
```

9.1 Anomaly Detection Using Threshold and Cumulative Entropy

In algorithm 4, we experimented on sample data using cumulative entropy and different threshold values. The algorithm takes sample data as a Dataframe having several columns as input, displays the cumulative entropy and which packets are considered an anomaly. A list of required column names (required columns) is created, it consists of the attributes, for example, Length, Year, Month etc to ensure that the dataset contains all the necessary information For analysis. Then it checks if all provided columns exist in sample data if any of required column is missing from dataset, then raise Value Error with suitable message. A function calculate_entropy(data) is created to compute Shannon's entropy of given data it calculates normalized value counts of unique values in the data. It returns the entropy using the Formula mentioned in 6, 7, 8, 9, and 10 earlier in sections of this paper. A new column, PacketLengthEntropy, is created in sample data. This column stores the rolling entropy calculated over the last 10 entries of the Length column, using the previously defined f unction. The cumulative sum of the PacketLengthEntropy column is calculated and stored in a new column CumulativeEntropy. This serves as the cumulative entropy over time. Any NaN value in the CumulativeEntropy column is replaced with 0, such that subsequent calculations do not fail. The threshold to determine anomalies is computed as mean(CumulativeEntropy) + 3 * standard_deviation(CumulativeEntropy), where an anomaly represents an entry being seen after which its cumulative entropy becomes larger than this threshold. Also, another new column Anomaly among the sample dataset constructed by replicate indicating if each packet's cumulative entropy exceeds the determined threshold (TRUE for anomaly; FALSE otherwise). Finally, we print columns year, month, day, hour, minute, second, Cumulative_ Entropy and Anomaly from our sample datasets.

Algorithm 4 Anomaly Detection in Packet Data

```
BEGIN
// Input: sample_data (DataFrame containing packet data with required columns)
// Output: Display of cumulative entropy and detected anomalies
// Step 1: Define required columns
required_columns ← [Length, Year, Month, Day, Hour, Minute, Second, Microsecond, Protocols...]
// Step 2: Check if all required columns are present NOT ALL(col ∈ sample_data.columns For col in required_columns)
RAISE ValueError("Missing required columns in the dataset.")
```

8: // Step 3: Define function calculate_entropy (data)
9: **Function** calculate_entropy(data)
10: // Calculate value counts of data normalized to probabilities
11: *value_counts* ← COUNT(*occurrences* of each unique value in *data*)
12: RETURN $-\sum(p_i * \log_2(p_i + \epsilon))$ where $\epsilon = 1e - 9$
13: // Step 4: Create new column *PacketLength Entropy*
14: *sample_data*['*PacketLengthEntropy*'] ← APPLY calculate_entropy ON ROLLING WINDOW OF SIZE 10 OVER *icmp_sample_data*['*Length*'] WITH min_periods = 1
15: // Step 5: Calculate cumulative entropy
16: *sample_data*['*CumulativeEntropy*'] ← CUMULATIVE SUM OF *sample_ data*['*PacketLengthEntropy*']
17: // Step 6: Fill NaN values in Cumulative Entropy with zero
18: 18: FILL NaN VALUES IN *sample_data* ['*CumulativeEntropy*'] WITH 0
19: // Step 7: Define threshold for anomaly detection
20: *threshold* ← MEAN(*sample_data* ['*CumulativeEntropy*']) + 3 $*STD$ (*sample_data*['*CumulativeEntropy*'])
21: // Step 8: Create new column *Anomaly*
22: *sample_data*['*Anomaly*'] ←TRUE IF *icmp_ sample_data*['*CumulativeEntropy*'] > *threshold* ELSE FALSE
23: // Step 9: Display results
24: PRINT SELECTED COLUMNS (*Year*,*Month*,*Day*, *Hour*,*Minute*,*Second*,*CumulativeEntropy*, *Anomaly*)
25: **END**

10. Results and Findings

10.1 Time series analysis research findings on dataset comparing HTTP flood a ttacks and ICMP flood attacks

In this research, we have created a graph for both HTTP flood attack and ICMP flood attack IOT datasets. We used the most important length and time features indicated in the dataset. Then we compared these two graphs and we discovered that ICMP traffic was much higher in the ICMP flood dataset. In Fig. 2, we also found that in the HTTP flood dataset, the other protocol traffic was higher and stable, which indicates a much lower risk. In Fig. 3, however, in the ICMP flood dataset, the other protocol traffic was less and unstable, which indicates a much higher risk. In this time series analysis, we have also found that having higher ICMP traffic in ICMP floods would have resulted in disrupting other protocol traffic in the system, creating a traffic congestion in IOT devices and smart home systems. We also discovered a very important understanding with this research analysis that, if both ICMP flood attack and HTTP flood attack have been initiated simultaneously, if both ICMP and HTTP traffic increase simultaneously, this may suggest a multi-vector attack strategy and would have and in future can have much more higher risk of traffic congestion and will result in more successful DDoS attack.

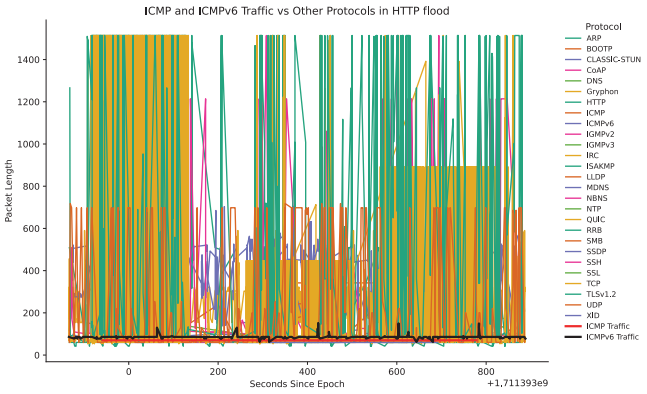


Figure 2. Time series analysis on HTTP Flooded attack traffic

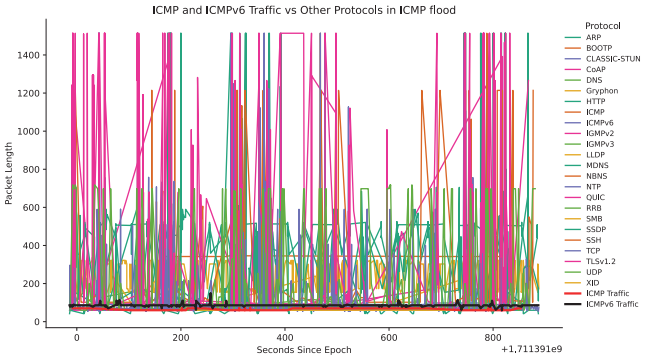


Figure 3: Time series analysis on ICMP Flooded attack traffic

10.2 Anomaly Detection using Cumulative Entropy and Thresholding Comparison

In both graphs, you will find a common chart, «Anomaly Detection Threshold». It is generally obtained by statistically calculating (taking average and standard deviation of cumulative entropy values) from the historical data. Now when your cumulative entropy exceeds that threshold, then there is an indication that an anomaly has been detected, i.e. some malicious activity (here DDoS attack) might be going on. Then, in both graphs, you can see a few explicit points marked where anomalies were detected throughout the time frame for which the analysis was done. So those peaks in Cumulative Entropy give an idea of which explicit timings during that period traffic was abnormal with respect to other timing instances.

In fig. 4, we have found that there was an unstable traffic anomaly detected after the threshold mentioned

in the ICMP flooded attack IoT dataset. We have also discovered, that there was a sudden, unstable traffic change, and an anomaly was detected after a certain point of threshold calculated. In fig. 5, however, we have investigated the HTTP flooded attack IoT dataset with the same threshold algorithm and we found no anomaly detection of any unstable traffic in comparison to the ICMP flooded attack. We have also discovered that the traffic was not able to even touch the threshold mark at any point. Hence, after all the investigation and analysis of both the graphs, we have concluded that, ICMP flooded attack poses much higher risk in IoT devices and smart systems than the HTTP flooded attack.

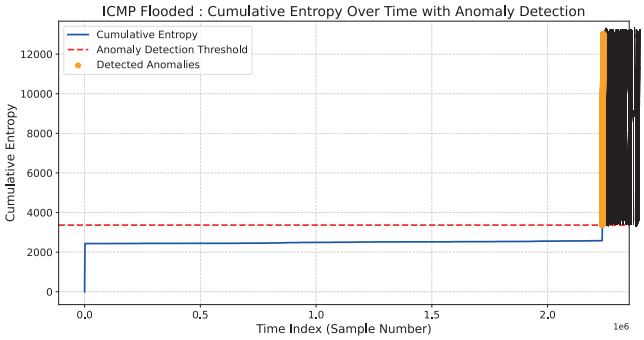


Figure 4: Anomaly Detection in ICMP Flooded attack dataset

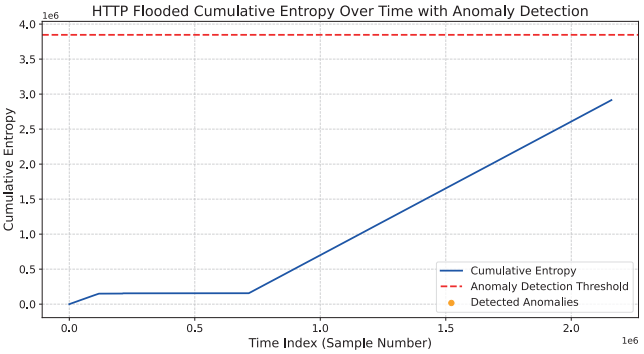


Figure 5: Anomaly Detection in HTTP Flooded attack dataset

11. Conclusion

The research findings reported in this article rovide substantial information on the impact of DDoS attacks, with an emphasis on ICMP and HTTP flood attacks in IoT contexts. We created comparison graphs of network traffic flow during different assault scenarios by analyzing time series data. Our results indicated that the level of ICMP traffic was significantly higher than HTTP in the ICMP flood dataset with the highest risk among other threats as reflected in our dataset. However, HTTP at its constant and highest rate used much traffic of other protocols within the HTTP flood dataset. Therefore, the risks associated with these types of attacks are less in general because they do not strictly use a specific protocol. By investigating cumulative entropy graphs, we noted the presence of peak sections or varying windows where anomalies occur along time signatures which function as reliable indicators to identify a potential malicious behavior and provide required knowledge on duration; how long and how network can be exposed to DDoS attacks harm. Notably, we discovered that simultaneous surges in both ICMP and HTTP traffic might indicate a multi-vector assault approach. This scenario increases the likelihood of network congestion and may lead to more effective DDoS attacks on IoT devices and smart home systems. In conclusion, our findings clearly show that ICMP flood attacks offer a far higher danger to IoT devices and smart home systems than HTTP flood assaults. The findings of this study will help to design effective countermeasures for strengthening the security of smart home systems against growing DDoS attacks. Further research is required to develop anomaly detection tools and study adaptive responses for increasing resilience to these sorts of intrusions.

12. Future Work and Ideas

In the future, we will use the same datasets, UCM_fiblo for hybrid model training on IoT devices and smart home systems datasets. This can be made possible by exploiting more powerful machine learning methods like deep learning models to improve anomaly detection accuracy in different kinds of network domains. Also, the combination with real-time big data analytics and edge computing aids in immediate threat intelligence-driven response and minimizes the latency. It is worth collaborating with IoT device manufacturers where adaptive security functionalities should be directly deployed in devices so that smart homes will have the capabilities of proactive defense against emerging DDoS attack vectors. In addition, multiple field experiments on real-world smart home deployment should be more intense, as these deployments provide a wealth of practical data that can facilitate rapid iteration on performance optimizations based on both user feedback and scientific measurements.

Acknowledgements

I am grateful to my wife, Darina Olegovna Ershova, For her professional advice and calm hand help during these challenging times. My study was helped by her expertise and vision, as well as her meticulous attention to detail. I am grateful to my department for their valuable contributions to this research, including their amazing experience of guidance. We appreciate the valuable feedback and recommendations from the faculty of Artificial I ntelligence a nd Cybersecurity at Moscow Institute of Physics and Technology (MIPT), Department of Radio Engineering and Cybernetics, (Institutsky Lane, 9, Moscow, Russian Federation, 141701). Their contributions helped shape an intellectually diverse conversation about my work.

Amit Kumar Jaiswal

АДАПТИВНЫЙ ПОРОГ КУМУЛЯТИВНОЙ ЭНТРОПИИ: НОВЫЙ ПОДХОД К ОБНАРУЖЕНИЮ DDOS-АТАК В УСТРОЙСТВАХ ИНТЕРНЕТА ВЕЩЕЙ И СИСТЕМАХ УМНЫХ ДОМОВ

Амит Кумар Джайсвал

Цель исследования: *предложить инновационную модель с использованием адаптивного порога, которая применяет кумулятивный энтропийный анализ временных рядов данных для более эффективного обнаружения и смягчения атак флудинга в среде «умного дома.*

Метод: *системный анализ, математические модели. Результат: с ростом популярности систем «умного дома» в повседневной жизни, атаки типа кибер-флудинга на эти взаимосвязанные устройства стали критически важными. В настоящем исследовании предложена инновационная модель с использованием адаптивного порога, которая применяет кумулятивный энтропийный анализ временных рядов данных для более эффективного обнаружения и смягчения атак флудинга в среде «умного дома». Модель устанавливает динамические пороги, адаптируемые к изменениям колебаний данных в режиме реального времени, используя кумулятивную энтропию – показатель, который определяет непредсказуемость и дисперсию моделей сетевого трафика. Изучены передовые методы машинного обучения для уточнения процесса установления пороговых значений, что в итоге приведет к более высокой точности обнаружения аномалий. Фактически будут проанализированы такие важные факторы, как временные паттерны, типы протоколов и действия пользователей, с точки зрения их влияния на показатели целей.*

Научная новизна: *подтверждена эффективность предлагаемых адаптивных пороговых рамок в ответ на значительное сокращение ложных срабатываний при одновременном улучшении реагирования на возникающие угрозы, что в целом повышает устойчивость систем умного дома к обнаруженным атакам типа «флуд».*

Ключевые слова: *анализ временных рядов, смягчение последствий атак типа «флуд», безопасность умного дома, обнаружение аномалий, анализ сетевого трафика, временные паттерны данных.*

References/Литература

1. *Lee S-H, Shiue Y-L, Cheng C-H, Li Y-H, Huang Y-F. Detection and Prevention of DDoS Attacks on the IoT. Applied Sciences. 2022; 12 (23): 12407. https://doi.org/10.3390/app122312407.*
2. *Shrahili, M.; Kayid, M. Cumulative Entropy of Past Lifetime for Coherent Systems at the System Level. Axioms 2023, 12, 899. https://doi.org/ 10.3390/axioms12090899*
3. *M. Tharun Kumar, G. Sesha Phaneendra babu, D. Lakshmi Narayana Reddy, «A Novel Framework for Mitigating DDoS Attacks in IoT Based Smart Network Environments using Machine Learning», Industrial Engineering Journal, ISSN: 0970-2555 Volume: 53, Issue 5, May: 2024. http://www.journal-iiie-india.com/1_may_24/125_online_may.pdf.*
4. *A. K. Jaiswal, «Deep Comparison Analysis: Statistical Methods and Deep Learning For Network Anomaly Detection», 2024. https://doi.org/10. 5281/zenodo.14051107*
5. *J. Dragos, J. P. Ziegler, A. de Villiers, A.-L. Jousselme, and E. Blasch, «Entropy-Based Metrics For URREF Criteria to Assess Uncertainty in Bayesian Networks For Cyber Threat Detection», in 2019 22nd International Conference on InFormation Fusion (FUSION), Ottawa, ON, Canada, 2019, pp. 1–8. DOI: 10.23919/FUSION43075.2019.9011276.*
6. *V. Timcenko and S. Gajin, «Machine Learning Enhanced Entropy- Based Network Anomaly Detection», Advances in Electrical and Computer Engineering, vol. 21, no. 4, pp. 51–60, 2021. DOI: 10.4316/AECE.2021.04006*
7. *P. Verma, S. Tapaswi, and W. W. Godfrey, «An Adaptive Threshold-Based Attribute Selection to Classify Requests Under DDoS Attack in Cloud- Based Systems», Arab Journal of Science and Engineering, vol. 45, pp. 2813–2834, 2020. DOI: 10.1007/s13369-019-04178-x.*
8. *P. Sahoo and Gurdial Arora, «A Thresholding Method Based on Two-Dimensional Renyi's Entropy», Pattern Recognition, vol. 37, no. 6, pp. 1149–1161, 2004. DOI: 10.1016/j.patcog.2003.10.008.*
9. *H. Lin and N.Bergmann, «IoT Privacy and Security Challenges For Smart Home Environments,"InFormation, vol. 7, no. 44, 2016. DOI: 10.3390/ info7030044.*
10. *M.C. Dani et al., «Adaptive Threshold For Anomaly Detection Using Time Series Segmentation», in Neural InFormation Processing, S. Arik et al., Eds., vol 9491 of Lecture Notes in Computer Science., Springer Cham., 2015.*
11. *Amit Jaiswal., «DOS Attack Network Traffic Monitoring in Software Defined Networking Using Mininet and RYU Controller». 2022. DOI: 10.21203/ rs.3.rs-2282189/v1.*
12. *Berezin'ski P, Jasiul B, Szpyrka M. An Entropy-Based Network Anomaly Detection Method. Entropy. 2015; 17(4): 2367–2408. DOI: https://doi.org/ 10.3390/e17042367.*
13. *Rong Lan and Lekang Zhang. 2023. Image Thresholding Segmentation Algorithm Based on Two-parameter Cumulative Residual Masi Entropy. In Proceedings of the 2022 5th International Conference on Artificial Intelligence and Pattern Recognition (AIPR '22). Association For Computing Machinery, New York, NY, USA, pp.406–411. DOI: https://doi.org/10.1145/3573942.3574041.*

14. J.Assfalg et al., «Time Series Analysis Using the Concept of Adaptable Threshold Similarity», in 18th International Conference on Scientific and Statistical Database Management (SSDBM'06), Vienna, Austria, pp. 251–260, 2006. https://www.dbs.ifi.lmu.de/Publikationen/Papers/ssdbm06. threshold.pdf.
15. D. Shang and P. Shang , «Analysis of Time Series in the Cumulative Residual Entropy Plane Based on Oscillation Roughness Exponent, Nonlinear Dynamics, vol.100, pp.,2167–2186 , 2020. DOI:10.1007/s11071-020-05646-y.
16. A. Patharkar et al., «Eigen-entropy Based Time Series Signatures to Support Multivariate Time Series Classification», Scientific Reports, vol.14, no.1, Article16076, 2024. DOI:10.1038/s41598-024-66953-7.
17. Huraj, Ladislav; Lietava, Jakub; Šimon, Marek (2024), «UCM_FibIoT2024», Mendeley Data, V1. DOI: 10.17632/p42xjtv8pv.1.
18. Yu, H., Yang, W., Cui, B. et al. Renyi entropy-driven network traffic anomaly detection with dynamic threshold. Cybersecurity 7, 64 (2024). https://doi.org/10.1186/s42400-024-00249-1.
19. M. Thakur and R.K. Sharma, «Anomaly Detection in Smart Home Networks Using Adaptive Thresholding Techniques Based on Cumulative Entropy», International Journal of Computer Applications, 2022. https: //doi.org/10.5120/ijca2016911955.
20. D. G. Narayan, W. Heena, and K. Amit, «A Collaborative Approach to Detecting DDoS Attacks in SDN Using Entropy and Deep Learning», Journal of Telecommunications and InFormation Technology, vol. 3, no. 3, 2024. https://doi.org/10.26636/jtit.2024.3.1609.

CYBERSECURITY ISSUES

SCIENTIFIC PEER-REVIEWED JOURNAL

2025, № 5 (69)

Cybersecurity Issues is a research periodical scientific and practical publication specializing in information security.
Published six times a year

https://cyberrus.info

The journal is being published from 2013
(Registration Certificate PI No. FS 77-75239).
CrossRef number (DOI): 10.21681/2311-3456

The journal is included in the Russian list of peer-reviewed academic publications of the Higher Attestation Commission (VAK), it is registered in the Russian Science Citation Index (RSCI/RINTs) on the Web of Science (WoS) platform and holds the 1st place in its cyber security rating. The journal's articles are available in full text

Founder and publisher
JSC «NPO «Echelon»
Postal address: Elektrozavodskaya str., 24, bld. 1, 107023, Moscow, Russia
E-mail: editor@cyberrus.info

CONTENTS

Вышла в свет уникальная книга, посвященная наиболее актуальной проблематике в области информационной безопасности – разработке безопасного ПО: КАК ИЗБЕЖАТЬ ОШИБОК ПРИ БЕЗОПАСНОЙ РАЗРАБОТКЕ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ / В.В. Вареница, А.С. Марков, В.Л. Цирлов и др. М.: Квант-Медиа, 2025. 344 с. Переплет твердый. ISBN 978-5-6053386-1-1

Книга подготовлена как учебное пособие коллективом авторов учебного центра «Эшелон» под редакцией доктора технических наук Маркова Алексея Сергеевича и группы его сотрудников на основе 20-летнего опыта работы в названной области. Впервые в нашей стране выпущено практическое пособие с учетом выполнения всех требований ГОСТ Р 56939-2024 — «Защита информации. Разработка безопасного программного обеспечения. Общие требования», являющегося национальным стандартом.

Содержание книги рассматривает все актуальные вопросы на всех стадиях жизненного цикла разработки безопасного программного обеспечения, каждый раздел сопровождается проверочными листами, в Приложение вынесены Типичные ошибки при внедрении процессов разработки.

Это учебное пособие станет настольной книгой специалистов, занимающихся вопросами организации и непосредственного внедрения процедур разработки безопасного программного обеспечения – как для молодых слушателей, так и для практикующих профессионалов, которым это пособие сэкономит много времени на поиск решений и существующих требований по безопасности.

CYBERSECURITY ISSUES
VOPROSY KIBERBEZOPASNOSTI

№ 5 2025

DOI: 10.21681/2311-3456

- Improvement of trade secret protection systems
- Systematization of information security tools
- Assessment of the security of big data management systems

www.cyberrus.info
editor@cyberrus.info