На правах рукописи

СУЛАВКО АЛЕКСЕЙ ЕВГЕНЬЕВИЧ

ВЫСОКОНАДЕЖНАЯ БИОМЕТРИЧЕСКАЯ АУТЕНТИФИКАЦИЯ НА ОСНОВЕ ЗАЩИЩЕННОГО ИСПОЛНЕНИЯ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ И АЛГОРИТМОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Специальность:

2.3.6. Методы и системы защиты информации, информационная безопасность

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
доктора технических наук

Омск – 2023

Работа выполнена в федеральном государственном автономном образовательном учреждении высшего образования «Омский государственный технический университет (ОмГТУ)» на кафедре «Комплексная защита информации»

Научный консультант: **Ложников Павел Сергеевич**
доктор технических наук, доцент

Официальные оппоненты: **Котенко Игорь Витальевич**
доктор технических наук, профессор,
ФГБУН «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук», главный научный сотрудник лаборатории проблем компьютерной безопасности

Катасёв Алексей Сергеевич
доктор технических наук, профессор,
ФГБОУ ВО «Казанский национальный исследовательский технический университет им. А.Н. Туполева-КАИ», профессор кафедры систем информационной безопасности

Вульфин Алексей Михайлович
доктор технических наук,
ФГБОУ ВО «Уфимский университет науки и технологий», доцент кафедры вычислительной техники и защиты информации

Ведущая организация: ФГАОУ ВО «Южный федеральный университет», г. Таганрог

Защита диссертации состоится «15» сентября 2023 г. в 10^{00} часов на заседании диссертационного совета 24.2.479.07, созданного на базе ФГБОУ ВО «Уфимский университет науки и технологий», по адресу:
450008, г. Уфа, ул. К. Маркса, 12.

С диссертацией можно ознакомиться в библиотеке ФГБОУ ВО «Уфимский университет науки и технологий» и на сайте *www.uust.ru*.

Автореферат разослан «___» _____ 2023 года.

Ученый секретарь
диссертационного совета Виноградова Ирина Леонидовна

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Сегодня мировой рынок биометрии проходит фазу активного роста (по данным MarketsAndMarkets к 2025 г. его объем составит 68 млрд. $). Биометрические системы внедряются повсеместно: на объектах критической информационной инфраструктуры, в банковской сфере, государственном секторе (более 80 стран используют биометрические паспорта), в сфере управления транспортом и городом. Рост рынка биометрических систем обусловлен новыми тенденциями и вызовами, с которыми столкнулось общество и государство: увеличение объемов данных о действиях пользователей в сети Интернет, которые могут быть использованы в злоумышленных целях (обострились проблемы приватности, анонимности пользователей и защищенности биометрических шаблонов от компрометации); применение технологий искусственного интеллекта (ИИ) для реализации хакерских атак, дезинформации, мошенничества, фальсификации биометрических образов человека (например, при помощи deepfake, голосовых синтезаторов); замена традиционных биометрических образов отпечатка пальца на более удобные образы голоса, лица и др., пригодные для бесконтактной аутентификации, но в большей степени подверженные дрейфу (изменчивости). В связи с этим современная высоконадежная биометрическая система должна строиться на основе доверенного ИИ, устойчивого к деструктивным факторам (дрейф биометрических данных, компьютерные атаки) и обладающего поддержкой защищенного режима исполнения. Под «защищенным исполнением» понимается невозможность анализа логики работы ИИ, управления ИИ и извлечения знаний из памяти ИИ любым неавторизованным субъектом.

Настоящее диссертационное исследование посвящено решению **научно-технической проблемы**, которая заключается в повышении надежности многофакторной биометрической аутентификации и защищенности биометрических систем от компьютерных атак на основе технологии защищенного исполнения нейросетевых моделей и алгоритмов искусственного интеллекта.

Степень разработанности темы исследования. На данный момент действует ряд международных стандартов, связанных с вопросами защиты биометрических систем от компьютерных атак (ISO/IEC 19792:2009, ISO/IEC 24761:2019, ISO/IEC 24745:2022, ISO/IEC 30107). Однако эти стандарты не позволяют устранить ряд актуальных угроз (извлечение знаний моделей ИИ, компрометация открытых биометрических образов, состязательные атаки). В России действует серия национальных стандартов ГОСТ Р 52633, не имеющих международных аналогов. Стандарты ГОСТ Р 52633 регламентируют особенности разработки, обучения и тестирования систем высоконадежной биометрической аутентификации, которые должны строиться на базе нейросетевых преобразователей биометрия-код (НПБК), позволяющих связать криптографический ключ или пароль пользователя с его биометрическим образом. Тем не менее, из-за наличия ряда недостатков применимость данных стандартов ограничена (высокая вероятность ошибок, малая длина ключа, подверженность атакам).

В мировой практике сложилось несколько подходов к повышению надежности биометрических систем аутентификации с обеспечением конфиденциальности биометрических данных, которые основаны на использовании нечетких экстракторов, искусственных нейронных сетей, искусственных иммунных систем, применении шифрования (в том числе, гомоморфного). Развитию аппарата искусственных нейронных сетей и искусственных иммунных систем, а также вопросам

создания доверенного ИИ посвящены работы многих ведущих российских и зарубежных ученых. Среди них Брюхомицкий Ю.А., Вульфин А.М., Гарбук С.В., Галушкин А.И., Иванов А.И., Котенко И.В., Николенко С. И., Baker B., Bengio Y., De Castro L. N., Fung C., Greensmith J., Hinton G.E., Kurkova V., LeCun Y., Mishra P.K., Schapire R.E., Stanley K.O., Timmis J. и другие. Вопросам высоконадежной биометрической аутентификации, оценки изменчивости биометрических параметров, обеспечения конфиденциальности биометрических данных, а также защите биометрических систем от компьютерных атак посвятили множество своих работ Ахметов Б.С., Бабенко Л.К., Безяев А.В., Васильев В.И, Волчихин В.И., Епифанцев Б.Н., Еременко А. В., Иванов А.И., Катасёв А.С., Ложников П.С., Маршалко Г.Б., Akkermans T.H., Catak F.O., Dodis Y., Hao F., Hafemann L.G., Jain A.K., Kumar A., Maiorana E., Mulionoa Y., Roy N.D., Wang L., Yuan L. и другие. Анализ этих работ позволил определиться с направлением диссертационного исследования и выявить перспективные подходы к решению обозначенной научно-технической проблемы. Эти подходы связаны с разработкой концепции защищенного исполнения нейросетевых алгоритмов ИИ, моделей искусственных нейронов и НПБК на их основе, изначально устойчивых к деструктивным воздействиям и атакам, адаптивных моделей ИИ, способных подстраиваться под изменяющиеся данные, снижая влияние концептуального дрейфа в задачах высоконадежной биометрической аутентификации, а также алгоритмов их обучения. Из проведенного анализа следует, что на основе предложенных концепции, моделей и алгоритмов необходимо разработать методы, технологию и программный комплекс для создания систем высоконадежной многофакторной биометрической аутентификации с обеспечением защиты биометрических данных от компрометации.

Объект исследования: системы биометрической аутентификации на основе методов, моделей и алгоритмов доверенного ИИ.

Предмет исследования: нейросетевые модели и алгоритмы машинного обучения на малых выборках для высоконадежной биометрической аутентификации и защиты биометрических данных от компрометации.

Цель диссертационной работы: повысить надежность многофакторной биометрической аутентификации на основе защищенного исполнения нейросетевых моделей доверенного ИИ и алгоритмов их автоматического синтеза и обучения на малых выборках биометрических данных.

Для достижения цели были выполнены следующие **задачи**:

1. Разработка концепции защищенного исполнения нейросетевых алгоритмов ИИ.
2. Разработка моделей искусственных нейронов и нейросетевого преобразователя биометрия-код, потенциально устойчивых к деструктивным воздействиям, и алгоритмов их робастного автоматического обучения на малых выборках.
3. Разработка адаптивной модели ИИ и алгоритмов ее обучения, позволяющих предупредить или снизить влияние концептуального дрейфа данных в системах биометрической аутентификации.
4. Разработка методов многофакторной аутентификации на базе тайных биометрических образов с обеспечением конфиденциальности биометрических данных.
5. Разработка технологии автоматического синтеза и обучения нейросетевых моделей для высоконадежной многофакторной биометрической аутентификации.

Основные результаты, выносимые на защиту:

1. Концепция защищенного исполнения нейросетевых алгоритмов ИИ, основанная на преобразовании корреляционных связей между признаками в мета-признаки, позволяющая снизить количество ошибок классификации образов и повысить защищенность систем ИИ от извлечения знаний.
2. Модель корреляционных нейронов и модель нейросетевого преобразователя биометрия-код на их основе, а также алгоритм их автоматического синтеза и обучения на малых выборках, которые позволяют повысить длину ключа, связываемого с биометрическими образами субъектов, и устойчивость биометрических систем к состязательным атакам и извлечению знаний.
3. Адаптивная нейро-иммунная модель ИИ и алгоритмы ее обучения с учителем и с подкреплением, позволяющие предупредить или снизить влияние концептуального дрейфа в системах биометрической аутентификации.
4. Методы и алгоритм высоконадежной многофакторной аутентификации на основе рукописных и голосовых образов, а также акустических образов уха с обеспечением защиты биометрических данных от компрометации, позволяющие повысить защищенность информации от неавторизованного доступа.
5. Технология автоматического синтеза и обучения нейросетевых моделей доверенного ИИ на малых выборках, а также программный комплекс на ее основе, позволяющие создавать системы высоконадежной биометрической аутентификации и другие ответственные приложения ИИ, обладающие повышенной устойчивостью к деструктивным воздействиям.

Научная новизна результатов:

1. Предложена концепция защищенного исполнения нейросетевых алгоритмов ИИ, *позволяющая* обеспечить устойчивость моделей и алгоритмов ИИ к извлечению знаний в задачах классификации образов, которая *в отличие* от существовавших ранее концепций основана на преобразовании корреляционных связей между признаками в высокоинформативные мета-признаки Байеса-Минковского с помощью предложенного для этой цели отображения. Экспериментально установлено, что корреляция между признаками увеличивает количество информации об образе (один мета-признак Байеса-Минковского может содержать в 2-3 раза больше информации, чем содержится суммарно в паре исходных признаков, от которых он порожден), что повышает надежность распознавания образов.
2. Разработаны модель корреляционных нейронов и модель НПБК на их основе, *отличающиеся* тем, что они анализируют корреляционные связи между признаками вместо признаков, а также робастной алгоритм автоматического синтеза и обучения этих моделей на малых выборках, что *позволяет* повысить защищенность биометрических данных от компрометации, длину ключа, связываемого с биометрическими образами субъектов, и устойчивость систем биометрической аутентификации к состязательным атакам.
3. Разработана адаптивная нейро-иммунная модель ИИ, *отличающаяся* от существовавших ранее использованием предложенной гибкой архитектуры искусственных иммунных детекторов (антител и клеток памяти), использованием в основе детекторов ядерных функций, сочетанием ансамблевых методов машинного обучения и метода обучения с подкреплением, что *позволяет* ей устойчиво обучаться на малых выборках и адаптироваться к изменению биометрических данных в процессе функционирования. Предложенные нейро-иммунная модель и алгоритмы ее обучения *в отличие* от существовавших ранее *позволяют* снизить

влияние концептуального дрейфа и вероятность ошибок биометрической аутентификации, даже если исходная обучающая выборка недостаточно репрезентативна или незначительна в объеме.

4. Разработаны методы и алгоритм высоконадежной многофакторной биометрической аутентификации, *отличающиеся* использованием новых акустических биометрических параметров, характеризующих внутреннее строение ушного канала, комплексированием динамических и статических признаков с учетом их приоритезации, информативности и стабильности, а также совместным использованием НПБК и нейро-иммунной модели, что *позволяет* обеспечить более высокую надежность аутентификации, робастность дрейфующих характеристик, защиту биометрических образов от компрометации, снизить вероятность ошибок «ложного допуска» и «ложного отказа» по сравнению с известными аналогами.

5. Разработана технология синтеза нейросетевых моделей доверенного ИИ, которая *позволяет* снизить объем тренировочной выборки, повысить надежность и защищенность биометрических систем аутентификации и других приложений ИИ, *отличающаяся* наличием режимов автоматического обучения нейросетевых моделей ИИ, защищенного исполнения нейросетевых алгоритмов классификации образов и применением процедур автоматической оценки информативности признаков.

Теоретическая значимость работы заключается в предложенной концепции, моделях и алгоритмах обучения. В совокупности они образуют математический аппарат, позволяющий создавать нейросетевой ИИ, который будет устойчив к различным деструктивным воздействиям на уровне архитектуры. Хотя в настоящей работе в качестве ключевой научной задачи выбрана задача высоконадежной многофакторной биометрической аутентификации, предложенный аппарат может применяться в других приложениях ИИ, для которых актуальны вопросы обеспечения защиты от компьютерных атак, извлечения знаний и обучения/дообучения на малых выборках. Решены важнейшие задачи автоматизации машинного обучения с использованием малых выборок биометрических данных и онлайн-обучения нейросетевых моделей (обучения модели в процессе ее исполнения в реальной практике). Это позволяет снизить негативное влияние таких факторов, как дрейф биометрических данных, а также в некоторых случаях успешно обучать модели, даже если биометрических данных мало, а тренировочная выборка недостаточна репрезентативна. Полученные результаты вносят значительный вклад в теорию машинного обучения, так как впервые предлагается использовать корреляционные связи между признаками в качестве новых метапризнаков и дается количественная оценка информативности этих метапризнаков.

Практическая значимость работы. На базе предложенной технологии синтеза нейросетевых моделей ИИ под руководством соискателя на базе ОмГТУ разработана первая редакция государственного национального стандарта ГОСТ Р «Искусственный интеллект. Нейросетевые алгоритмы в защищённом исполнении. Автоматическое обучение нейросетевых моделей на малых выборках в задачах классификации». Это первый стандарт, который регламентирует особенности создания и обучения нейросетевых моделей ИИ, исполняемых в защищенном от исследования режиме. Стандарт направлен на использование на объектах критической информационной инфраструктуры при разработке ответственных приложений ИИ. Стандарт прошел экспертизу технических комитетов Росстандарта и

включен в программу стандартизации технического комитета «Искусственный интеллект» (ТК164).

Результаты работы легли в основу линейки программных продуктов AIConstructor (AIC), научным руководителем разработки которых является Сулавко А.Е. AIC desktop – программный комплекс для проведения научных исследований по машинному обучению. AIC ModelOps Platform – корпоративная среда управления жизненным циклом ИИ, может использоваться для автоматизации, отслеживания и контроля рабочих процессов на всех этапах: от исследования до внедрения в бизнес среду.

Практическую значимость представляют методы высоконадежной многофакторной биометрической аутентификации по особенностям ушного канала, рукописным и голосовым образам с показателями FRR=0,12 при $FAR<10^{-14}$ и FRR=0,03 при $FAR<10^{-10}$ и программные продукты на их основе.

Методы исследования. Применялись методы распознавания образов, машинного обучения, кодирования информации и защиты данных от компрометации, аппарат искусственных нейронных сетей (ИНС), ансамблевые методы, биоинспирированные алгоритмы и модели классификации образов, методы теории вероятностей и математической статистики, спектрального и корреляционного анализа, обеспечения дифференциальной конфиденциальности данных и знаний, идентификации и аутентификации.

Достоверность полученных результатов обусловлена корректным применением методов исследования, использованием признанных методик статистической обработки данных, математически строгим выполнением расчетов и подтверждается результатами практического использования и актами внедрения. Вводимые допущения мотивировались фактами, известными из практики. Предложенные в работе концепция, модели, методы и алгоритмы теоретически обоснованы и не противоречат известным достоверно подтвержденным результатам исследований других авторов.

Реализация и внедрение результатов работы. Результаты работы внедрены на предприятиях: ООО «Открытый код», ООО «Системы информационной безопасности», ООО «АИ ЗИОН», ООО «Джемс Девелопмент», БУЗОО «Медико-санитарная часть № 4», где они использовались в проектно-конструкторской деятельности, и в учебный процесс ФГАОУ ВО СПбГЭТУ «ЛЭТИ» и ФГАОУ ВО «ОмГТУ». **Результаты применялись при разработке первой редакции национального стандарта** ГОСТ Р «Искусственный интеллект. Нейросетевые алгоритмы в защищённом исполнении. Автоматическое обучение нейросетевых моделей на малых выборках в задачах классификации» под руководством соискателя на базе ОмГТУ, который поставлен в план стандартизации техническим комитетом № 164 «Искусственный интеллект».

Результаты работы связаны с научными программами, руководителем которых являлся соискатель: государственное задание Минобрнауки России на 2023-2025 годы № FSGF-2023-0004, Грант ИБ №6 от МИРЭА и Минобрнауки РФ; Гранты РНФ 17-71-10094 «Разработка технологии широких нейронных сетей сверхбыстрого обучения и ее применение для надежной аутентификации субъектов на основе тайных биометрических образов», РФФИ 18-41-550002, РФФИ 16-37-50005; НИР «Защищенный режим исполнения искусственного интеллекта на базе автоматически обучаемых сетей автокорреляционных нейронов»; Грант Фонда Содействия Инновациям на проведение НИОКР по теме «Разработка

ModelOps платформы для оптимизации процесса цифровой трансформации при создании и внедрении доверенного искусственного интеллекта с использованием сетей корреляционных нейронов»; НИР «Способы распознавания субъектов и их психофизиологического состояния по динамическим биометрическим признакам», НИР «Гибкие нейросетевые алгоритмы для анализа биометрических образов», НИР «Защита информационных и компьютерных систем на базе предиктивного анализа биометрических и поведенческих характеристик оператора».

Также соискатель участвовал в статусе исполнителя в Госзадании 2.9314.2017/БЧ и следующих проектах РФФИ: 13-07-00246, 15-07-09053, 16-07-01204, 18-37-00399, 15-37-50269, 16-37-50045.

Апробация результатов. Результаты работы регулярно докладывались и обсуждались на научных конференциях: Международная IEEE научно-техническая конференция «Динамика систем, механизмов и машин», г.Омск (2014, 2016, 2017, 2018); Научно-практическая конференция «Безопасность информационных технологий», г.Пенза (2014, 2016, 2020); Международная IEEE Сибирская конференция по управлению и связи SIBCON, г.Омск, 2015, г.Москва, 2016, г.Астана, 2017; Международная конференция «Аппроксимация логических моделей, алгоритмов и задач», г.Омск, 2015; IEEE Международная конференция по использованию информационно-коммуникационных технологий г.Баку, Азербайджан, 2016; Международная научно-практическая конференция «Научно-технический прогресс: актуальные и перспективные направления будущего», г.Кемерово, 2016; Международная научно-практическая конференция «Инфографика и информационный дизайн: визуализация данных в науке», г.Омск, 2017; IFAC Conference on Technology, Culture and International Stability (TECIS), г.Баку, Азербайджан, 2018, г.Созополь, Болгария, 2019; Межвузовская научно-практическая конференция «Информационная безопасность: современная теория и практика», г.Омск (2018, 2019, 2020); Международная научно-техническая конференция «Актуальные проблемы электронного приборостроения (АПЭП)», г.Новосибирск, 2018; Всероссийская научно-практическая конференция с международным участием им. В.В.Губарева «Интеллектуальный анализ сигналов, данных и знаний: методы и средства», г.Новосибирск, 2018; Международная научно-техническая конференция «Проблемы машиноведения», г.Омск (2018, 2019, 2020); Международная научно-практическая конференция «Цифровизация и кибербезопасность: современная теория и практика», г.Омск, 2021.

Соответствие паспорту специальности. Результаты диссертационной работы соответствуют следующим пунктам паспорта научной специальности 2.3.6 – Методы и системы защиты информации, информационная безопасность: п. 9. «Модели противодействия угрозам нарушения информационной безопасности для любого вида информационных систем, позволяющие получать оценки показателей информационной безопасности»; п. 12. «Технологии идентификации и аутентификации пользователей и субъектов информационных процессов. Системы разграничения доступа»; п. 15. «Принципы и решения (технические, математические, организационные и др.) по созданию новых и совершенствованию существующих средств защиты информации и обеспечения информационной безопасности».

Публикации. Соискателем опубликовано 80 работ, содержащих результаты диссертационного исследования, в том числе 38 статей в журналах из Перечня рецензируемых научных изданий, рекомендованных ВАК, либо в научных изда-

ниях, индексируемых базой RSCI, 21 научная работа в изданиях, включенных в базы Web of Science и Scopus, 11 научных работ в других изданиях и 1 коллективная монография. Получен 1 патент на изобретение и 8 свидетельств о регистрации программ.

Структура и объём диссертации. Диссертация состоит из введения, 5 глав, заключения, списка сокращений, списка литературы и приложений. Диссертация содержит 391 страница машинописного текста, включая 108 рисунков, 28 таблиц, список литературы из 362 наименования.

Личный вклад автора состоит в постановке задач исследования, разработке экспериментальных и теоретических методов, разработке, тестировании и реализации предложенных концепции, моделей, методов, алгоритмов и компьютерных программ, анализе и обобщении полученных результатов и формулировке выводов. **Все результаты и положения, выносимые на защиту, а также научная новизна получены лично автором.** Подготовка к публикации некоторых результатов проводилась совместно с соавторами, но вклад диссертанта был определяющим. Участие научного консультанта заключалось в оказании методической и организационной помощи в формулировании задач, представлении результатов и оценке их корректности.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** изложена актуальность направлений исследования, сформулированы его цель и задачи, перечислены применяемые для их решения методы. Изложены элементы научной новизны и практической ценности результатов исследования.

В **первой главе** даны определения ключевых понятий. Описаны основные критерии доверенного ИИ. В частности, отметим, что под *надежностью* понимается способность биометрической системы сохранять во времени требуемый уровень точности аутентификации и защищенности от компьютерных атак в изменяющихся условиях функционирования. Согласно ГОСТ Р 52633.0 система биометрической аутентификации является высоконадежной, если показатель вероятности ошибки «ложного допуска» составляет менее 10^{-12}. Приведены угрозы и атаки («на решающий бит», «извлечение знаний», «ключ под ковриком», «состязательные», «представления», рисунок 1), перед которыми уязвимы биометрические системы и другие приложения ИИ. Указаны недостатки методов защиты знаний ИИ и биометрических шаблонов на основе гомоморфного шифрования, а также ограничения концепции федеративного обучения. Проведен анализ существующих стандартов и научных публикаций в области защиты ИИ и биометрических систем от обозначенных угроз.

В рамках деятельности Международных технических комитетов (ТК) по стандартизации ISO/IEC JTC 1/SC 42 «Artificial intelligence», ISO/IEC JTC 1/SC 37 Biometrics и национальных ТК 164 «Искусственный интеллект» и ТК 098 Биометрия и биомониторинг пока не разработано стандартов в области защиты решающих правил от обозначенных атак. Наиболее детально эти вопросы проработаны для приложений биометрии. Национальные стандарты по нейросетевой биометрии серии ГОСТ Р 52633, закрепленные за ТК 362 «Защита информации», основаны на концепции НПБК.

Концепцию защищенного исполнения нейросетевых алгоритмов ИИ в задачах биометрической аутентификации позволяют реализовать: нечёткие экстракторы; гибриды нечеткого экстрактора и многослойных ИНС; автоматически обу-

чаемые НПБК на базе ИНС; НПБК на базе квадратичных нейронов. Выявлены принципиальные недостатки методов защиты знаний путем гомоморфного шифрования. Анализ работ показал, что для задач классификации в защищенном режиме архитектуру ИИ можно разделить на два блока: блок извлечения признаков и блок классификации образов.

Длина ключа для нечеткого экстрактора и НПБК ограничена из-за ряда принципиальных недостатков и подверженности атакам, соответственно. Для сетей квадратичных нейронов длина ключа может быть примерно в 4 раза выше, чем для классических НПБК. Ограничения длины ключа не позволяют использовать такие нейроны для широкого спектра классификационных задач. Применение методов глубокого обучения и аппарата многослойных нейронных сетей, по всей видимости, ограничивается блоком извлечения признаков. Блок классификации должен иметь простую архитектуру, такую, чтобы его синтез и обучение могли выполняться в автоматическом режиме.

В биометрических системах крайне важно обеспечить актуальность знаний ИИ. С течением времени возрастает количество сбоев и ошибок из-за изменчивости биометрических данных, например, в зависимости от ПФС. Данная проблема является частным случаем дрейфа концепций и данных.

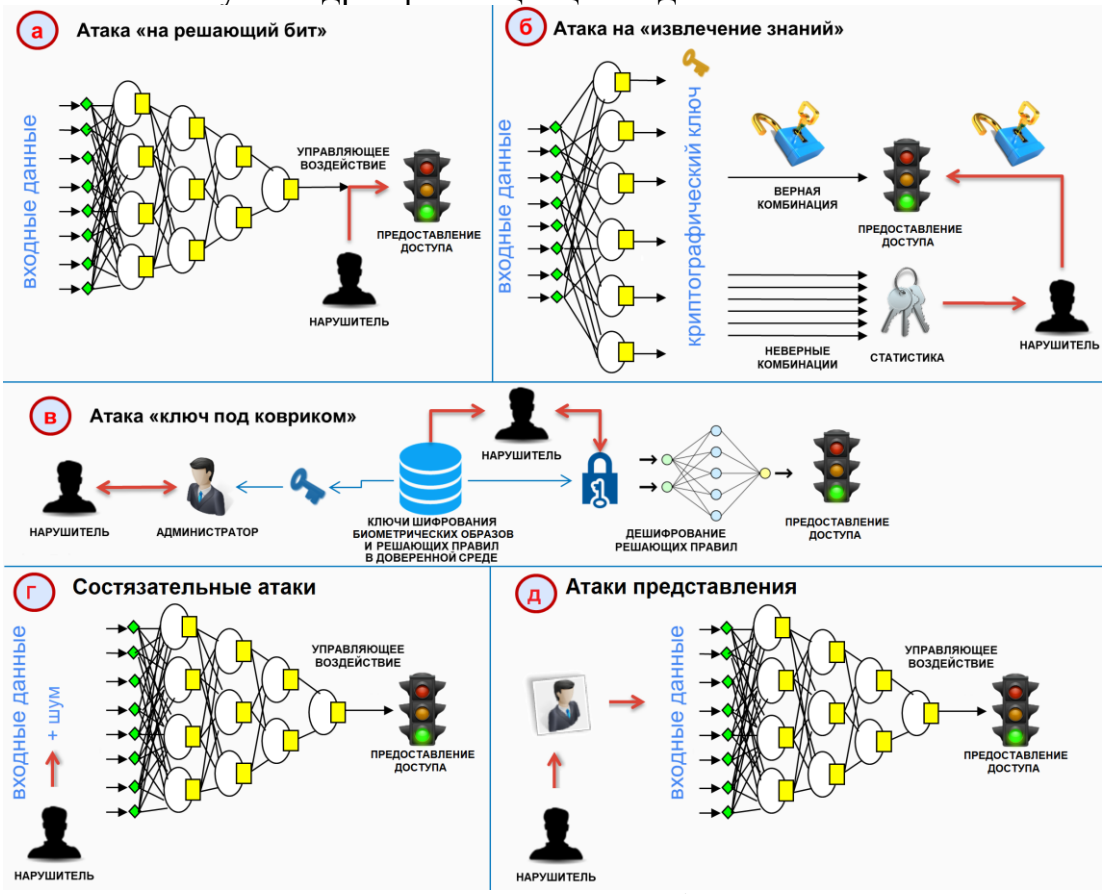

Рисунок 1. – Атаки на высоконадежные биометрические системы.

Выводы: для решения обозначенных проблем требуется разработать иные модели искусственных нейронов и сетей, позволяющих повысить длину ключа в соответствии с текущими требованиями (например, ГОСТ Р 34.10-2012) или выше (на перспективу), а также позволяющих предотвратить или снизить влияние концептуального дрейфа. Безопасность блока классификации может быть дополнительно усилена криптографическими методами.

В конце главы сформулированы цели и задачи исследования.

Во **второй главе** разработана концепция защищенного исполнения нейросетевых алгоритмов ИИ и модель корреляционных нейронов для высоконадежной биометрической аутентификации – это новый класс нейронов, анализирующих корреляционные связи между признаками вместо признаков. Анализ внутренних корреляционных связей образов и их классификация происходит без необходимости хранения информации о корреляционных связях или признаках, характерных для биометрических образов пользователей (биометрические шаблоны не компрометируется при хранении). Корреляционные нейроны используются для формирования высоконадежного блока классификации, устойчивого к деструктивным воздействиям. Предложена модель НПБК на основе корреляционных нейронов и алгоритм ее автоматического обучения на малых выборках биометрических данных.

Показано, что пространство признаков искривляется из-за корреляционных связей между измерениями. Классы образов имеют уникальные матрицы коэффициентов корреляции $C_{j,t}$ между признаками. Поэтому относительно различных классов пространство признаков искривлено по-разному. Корреляция переносит часть информации об образах, касающейся уровней искривления, в «скрытые» измерения. Чтобы извлечь эту информацию предложено множество метрик Байеса-Минковского, в частности (1).

$$y = \sum_{j=1}^{n} \left| \left| \frac{a_t}{\delta_t} \right|^p - \left| \frac{a_j}{\delta_j} \right|^p \right|, j \neq t, \qquad y = \sqrt[p]{\sum_{j=1}^{n} \left| \left| \frac{a_t}{\delta_t} \right|^p - \left| \frac{a_j}{\delta_j} \right|^p \right|}, j \neq t, \qquad (1)$$

где y – оценка близости образа к классу «Свой», a_j – значение j-го признака из вектора $\bar{a}$, представляющего распознаваемый образ; n – количество признаков, δ_j – нормирующие коэффициенты, вычисляемые как среднеквадратичное отклонение признака для класса «Чужие», представляющего множество обезличенных образов, поэтому δ_j не компрометируют данные класса «Свой» (*обеспечивается дифференциальная конфиденциальность*).

На рисунок 2 а,б видно, что $AUC_{|C|>0,95}(\Phi_G(y), \Phi_I(y)) < AUC_{|C|<0,3}(\Phi_G(y), \Phi_I(y))$, где AUC – площадь, ограниченная функциями плотности вероятности (ФПВ) «Свой» $\Phi_G(a_j)$, «Чужие» $\Phi_I(a_j)$, и осью абсцисс. Чем выше корреляция между признаками, тем меньше неверных решений дает мера (1). Размерность пространства Байеса-Минковского составляет: $n' = 0{,}5(n(n\text{-}1)) = 0{,}5n^2 - 0{,}5n$.

Под мета-признаками подразумеваются разности вида (2):

$$a'_{j^*} = a'_{t,j} = f(a_t, a_j) = \left| \left| \frac{a_t}{\delta_t} \right|^p - \left| \frac{a_j}{\delta_j} \right|^p \right|, j > t, j^* = \sum_{j^*=1}^{t-1} (n - t^*) + j - t \qquad (2)$$

которые фактически являются грубыми оценками корреляционной зависимости между двумя исходными признаками под номерами j и t, сделанными всего по одному примеру, но при наличии некоторых априорных знаний, полученных в процессе обучения на выборке небольшого объема.

Экспериментально установлено (на больших объемах сгенерированных данных), что один мета-признак Байеса-Минковского может содержать в 2-3 раза больше информации, чем содержится суммарно в паре исходных признаков, от которых он порожден, если они сильно коррелированы.

Корреляционный нейрон предлагается строить на метрике взвешенного среднеквадратичного отклонения (3), которая позволяет отделить как положительно, так и отрицательно коррелированные данные:

$$y=\sqrt{\frac{1}{\eta}\sum_{j^*=1}^{n'} w_{j^*}\left(a'_{j^*}-m'\right)^2}=\sqrt{\frac{1}{\eta}\sum_{\iota=1}^{\eta} w_{\iota}\left(a'_{\iota}-m'\right)^2},\ m'=\frac{1}{\eta}\sum_{\iota=1}^{\eta} a'_{\iota} \qquad (3)$$

где η – количество входов нейрона, w_{j^*} – вес синапса под номером j^* ($w_{j^*}\geq 0$, если $w_{j^*}=0$, то j^*-й мета-признак не влияет на сумму, т.е. не соединяется с нейроном), ι – номер мета-признака без учета синапсов с нулевым весом (для сквозной нумерации). Вес синапса рассчитывается по формуле (4):

$$w_{\iota}=\frac{\left|m''_{(G),\iota}-m''_{(I),\iota}\right|}{\sigma''_{(G),\iota}\cdot\sigma''_{(I),\iota}}, \qquad (4)$$

где $m_{(G),\iota}''$, $m_{(I),\iota}''$ – математические ожидания, а $\sigma_{(G),\iota}''$, $\sigma_{(I),\iota}''$ – среднеквадратичные отклонения значений ι-го мета-признака второго порядка ($a''_{\iota}=(a'_{\iota}-m')^2$) для образов «Свой» и «Чужие», соответственно, рассчитанные по данным обучающей выборки. После обучения нейрона параметры $m_{(G),\iota}''$, $m_{(I),\iota}''$, $\sigma_{(G),\iota}''$, $\sigma_{(I),\iota}''$ должны быть удалены. В качестве функции активации предлагается использовать многоуровневую пороговую функцию квантования (5):

$$\phi(y)=\begin{cases}3, y<T_{left}\\ 2, T_{left}\leq y<T_{middle}\\ 1, T_{middle}\leq y<T_{right}\\ 0, y\geq T_{rightt}\end{cases}, \qquad (5)$$

где T_{left}, T_{middle} и T_{right} – левый, средний и правый пороговые значения активации нейрона (рисунок 2 в,г). В соответствии с предлагаемой моделью нейрон имеет четыре варианта активации {0, 1, 2, 3} и только одно из них соответствует гипотезе «Свой», остальные – гипотезе «Чужие». *О том, какое состояние активации соответствует гипотезе «Свой» (далее ϕ_G), известно только на этапе синтеза и обучения НПБК, злоумышленник не обладает этой информацией, так как она не сохраняется после настройки нейрона.*

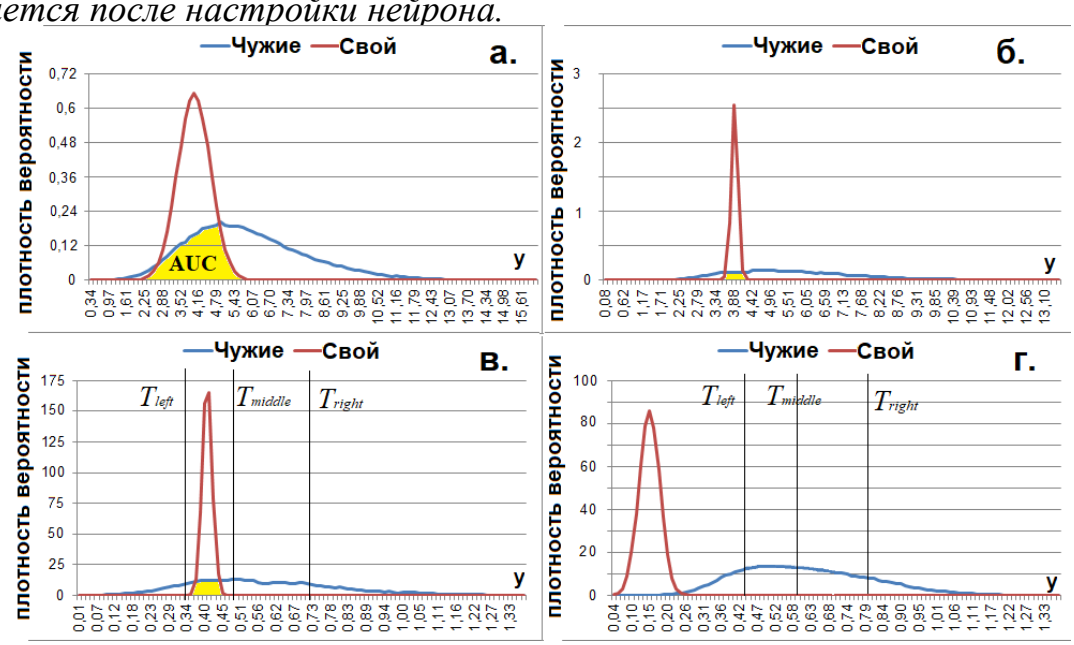

Рисунок 2. – Графики плотностей вероятности:

- значений меры (1) при p=1: а. $|C_{j,t}|<0{,}3$, n'=5; б. $1>C_{j,t}>0{,}95$, n'=5;
- значений меры (3): в. $1>C_{j,t}>0{,}95$, n'=10; г. $-1<C_{j,t}<-0{,}95$, n'=10.

При $0{,}1<P(\phi(y))<0{,}4$ обеспечивается достаточно высокая энтропия выходов нейронов в ответ на образы «Чужие», где $P(\phi(y))$ – относительная частота появления $\phi(y)$ при поступлении на вход образа «Чужой». При вычислении порогов сна-

чала рассчитываются граничные значения откликов нейрона y на обучающие примеры «Свой» (y_{Gmin}, y_{Gmax}) и «Чужие» (y_{Imin}, y_{Imax}), а также значения их функций распределения $F_G(y)$ и $F_I(y)$, исходя из гипотезы нормального распределения y (подтверждено методом хи-квадрат).

Предложен алгоритм настройки порогов (рисунок 3).

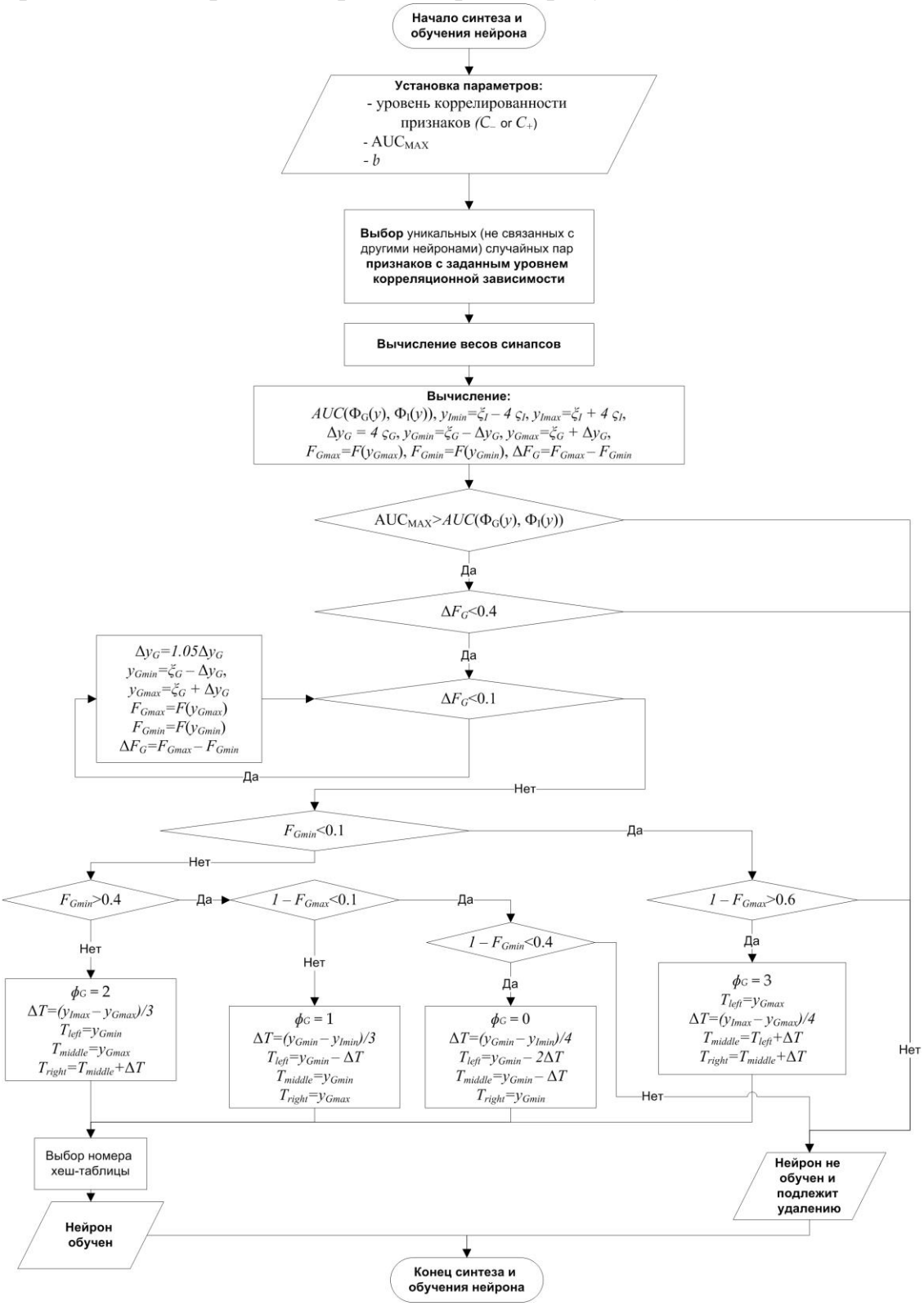

Рисунок 3. – Схема алгоритма синтеза и обучения корреляционного нейрона.

Введем коэффициент AUC_{MAX}, равный максимально допустимому показателю $AUC(\Phi_G(y), \Phi_I(y))$ для нейрона, чтобы исключить «слабые» нейроны, кото-

рые дают близкие отклики на образы «Свой» и «Чужие». К значению функции активации применяется одна из возможных таблиц перевода состояний {0, 1, 2, 3} в двухбитный код. При обучении нейрона хеш-таблица выбирается случайно, но с учетом того, на какие два ключевых бита (далее b) настраивается нейрон.

Разработана модель НПБК на базе корреляционных нейронов для реализации блока классификации, устойчивого к деструктивным воздействиям (рисунок 4). Перед поступлением образов на входы НПБК, они должны быть обработаны блоком извлечения признаков. Эксперименты с использованием синтетических данных показали, что при параметре p=0,9 отображения (2) в большинстве случаев удается достичь наименьшего количества ошибок классификации. Коэффициенты δ_j вычисляются на основании тренировочной выборки «Чужие» (один набор коэффициентов может использоваться для нормирования признаков сразу для множества НПБК, принимающих на входы аналогичные признаки).

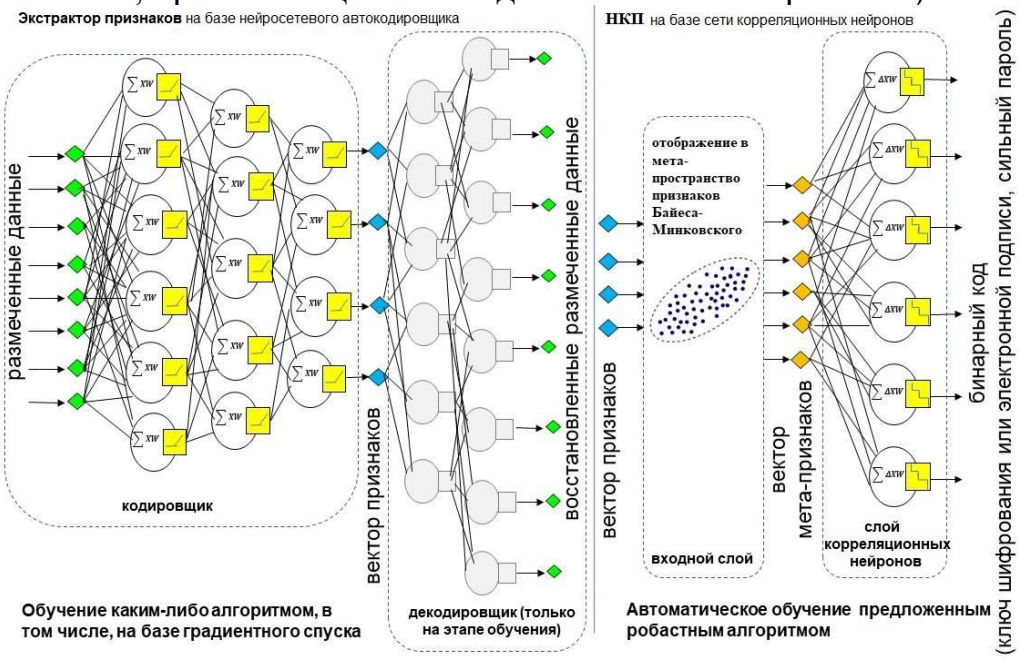

Рисунок 4. – Структурная схема связывания ключа и биометрического образа: слева блок извлечения признаков, справа блок классификации.

Так, для обучения корреляционного нейрона достаточно определить связанные мета-признаки, вычислить веса, пороги и задать хеш-таблицу.

Количество входов η для всех корреляционных нейронов НПБК должно быть равным. При синтезе НПБК для конкретного пользователя необходимо убедиться, что имеется достаточное количество пар признаков с уровнями взаимной корреляции $C_{j,t}<C_-$ и $C_{j,t}>C_+$ (по результатам экспериментов оптимальными являются C_+=0,5 и C_-=-0,5). Следует рассчитать корреляционную матрицу по данным выборки «Свой». Любая пара коррелированных признаков потенциально порождает один мета-признак Байеса-Минковского. Пусть N_- и N_+ – количества нейронов, ориентированных на обработку данных с взаимной корреляцией $C_{j,t}<C_-$ и $C_{j,t}>C_+$. Должно соблюдаться условие $N_- \approx N_+$ (допускается расхождение на 1-3 нейрона). Каждый нейрон должен обрабатывать уникальную комбинацию мета-признаков и генерировать на выходе 2 бита. Нужное количество нейронов определяется, исходя из длины ключа L. Например, при L=1024 бит $N_-=N_+=L/2/2$=256. Тогда если η=4, то для синтеза сети потребуется 2048 пар признаков. Алгоритм синтеза и обучения НПБК можно изложить как последовательность шагов:

1. Расчет корреляционной матрицы признаков.
2. Подсчет пар отрицательно коррелированных признаков ($C_{j,t}<C_-$). Если количество пар менее $\eta \cdot N_-$, то C_- увеличивается на 0,05 и шаг 2 повторяется.
3. Синтез и обучение N_- нейронов для анализа отрицательно коррелированных данных в соответствии с алгоритмом на рисунок 3. Если количество нейронов, удовлетворяющих условиям алгоритма на рисунок 3, оказалось менее N_-, то C_- увеличивается на 0,05 и шаги 2-3 повторяются.
4. Подсчет пар положительно коррелированных признаков ($C_{j,t}>C_+$). Если количество пар менее $\eta \cdot N_+$, то C_+ уменьшается на 0,05 и шаг 4 повторяется.
5. Синтез и обучение N_+ нейронов для анализа положительно коррелированных данных в соответствии с алгоритмом на рисунок 3. Если количество нейронов, удовлетворяющих условиям алгоритма на рисунок 3, оказалось менее N_+, то C_+ уменьшается на 0,05 и шаги 4-5 повторяются.

По мере сужения интервала (C_-; C_+) уже созданные нейроны допустимо не удалять. Алгоритм исполняется, пока не выполнится условие $N_-=N_+=L/4$ либо пока не будет нарушено условие $|C_{-+}| \geq 0{,}3$.

Таблицы весовых коэффициентов w_i и номера хеш-таблиц обученного НПБК представляют собой защищенный эталон пользователя.

В третьей главе показано, что ПФС отражается на динамических биометрических признаках (клавиатурного и рукописного почерка, голоса) спонтанными изменениями. Если ПФС пользователя не совпадает на этапах обучения системы и аутентификации, то вероятность ошибок «ложного отказа» (FRR) и «ложного допуска» (FAR) многократно повышается. Эту проблему можно решить с использованием алгоритмов онлайн обучения моделей ИИ.

Проведен анализ подходов к построению адаптивных моделей ИИ (моделей, способных к онлайн-обучению), включая методы глубокого обучения с подкреплением, эволюционный и иммунный подходы. Методы глубокого обучения с подкреплением и эволюционные нейронные сети работоспособны только при больших объемах обучающей выборки. Искусственные иммунные системы и сети (ИИС) обладают двойной пластичностью, позволяющей легко изменять в процессе функционирования не только собственные параметры, но и структуру (в отличие от нейронных сетей). Проанализированы существующие подходы к построению моделей ИИС (на базе дендритных клеток, негативного отбора, клональной селекции и сетевых алгоритмов), а также методы и алгоритмы биометрической аутентификации на их основе.

Модели, в которых совместно применяются элементы аппаратов ИИС и ИНС, принято называть нейро-иммунными (частный случай нейросетевых моделей). Синтез и обучение нейро-иммунных моделей выполняется с использованием принципов ИИС, однако процесс классификации образов при помощи нейро-иммунной модели после ее обучения схож работе ИНС.

Детектор (аналог искусственного нейрона) – искусственная иммунная клетка, которая обладает способностью обнаруживать чужеродные антигены, анализируя распознаваемый образ и реагируя на него пропорционально тому, насколько этот образ соответствует антигену (шкала реакций [0;1]). Каждый детектор следует рассматривать как бинарный классификатор, являющийся «обверткой» над корреляционным (классическим, квадратичным) нейроном, состоящий из нескольких функций, последовательно применяющихся к $\bar{a}$ (6):

$$u_i = \phi_\chi\left(y' = \varphi\left(y = f_x\left(\bar{\alpha} = R(\bar{a}, \Psi_i), \bar{g}, \Theta_i\right), T_i\right)\right), \quad (6)$$

1.) $\bar{\alpha} = R(\bar{a}, \Psi_i)$ – функция-рецептор извлекает η из n признаков, Ψ_i – множество номеров признаков, которые должен анализировать i-й детектор;

2.) $y = f_x(\bar{\alpha}, \breve{g}, \Theta_i)$ – функция-ядро вычисляет близость образа к эталону класса «Свой»; x – тип ядра (например, на базе корреляционного нейрона (3)); $\breve{g}$ – вектор параметров функционала, которые влияют на характер вычислений (например, степенной коэффициент p для перехода в пространство мета-признаков Байеса-Минковского); $\Theta_i=\{w_1,w_2,\ldots,w_\eta,\delta_1,\delta_2,\ldots,\delta_\eta\}$ – множество параметров обученного нейрона. Элементы из множеств $\breve{g}$ и Θ_i зависят от типа нейрона или меры близости, на базе которых строится детектор. Помимо корреляционных нейронов в основе детектора может лежать квадратичный, классический нейрон, байесовский классификатор и др. Однако большинство существующих метрик компрометируют знания ИИ, в отличие от корреляционного нейрона. Разные ядра образуют различные виды детекторов, которые дают слабо коррелированные решения относительно друг друга, что позволяет объединять такие нейроны (детекторы) в сеть для получения синергетического эффекта. Из любого ядра можно получить разные меры близости за счет изменения параметров $\breve{g}$.

3.) $y' = \varphi(y, T_i)$ – функция нормирования откликов y относительно порога T_i, который вычисляется в процессе настройки i-го детектора. Для корреляционных нейронов функция имеет вид: $\varphi(y, T_i) = y / T_i$, где T_i – это максимальное значение функции-ядра i-го детектора, при поступлении на его вход обучающих образов «Свой».

4.) $u_i = \phi_\chi(y'_i)$ – функция активации, дополнительный нелинейный элемент детектора, который определяет особенности реагирования на антиген. Функция активации также необходима, чтобы привести отклик детектора к области значений [0;1]. В отличие от функции активации в защищенном режиме (5) адаптивная модель ИИ использует сигмоиды.

Одной из теоритических проблем аппарата ИИС является слабая обоснованность используемых мер близости. Согласно теореме «об отсутствии бесплатных завтраков» (No Free Lunch) ни одна мера близости не может быть оптимальной для всего множества задач распознавания образов. Поэтому в настоящей работе каждый детектор определяет близость уникальным способом, а состав детекторов «подстраивается» под задачу в процессе обучения.

Разработанная адаптивная модель ИИ (рисунок 5) также как и модель защищенного исполнения настраивается на верификацию образа конкретного пользователя и обучается на малых выборках образов «Свой» и «Чужие» (можно использовать одну выборку «Чужих» для обучения всех моделей).

Детекторы разделены на две группы: *врожденный (ВИ) и приобретенный иммунитет (ПИ)*, и рассматриваются как *два комитета (ансамбля)* слабых классификаторов, обучаемых при помощи разных алгоритмов. Коллективное решение N детекторов вычисляется как среднее частных решений:

$$\ddot{u} = \Phi(\overline{D}^* = \{D_1^*,\ldots D_N^*\}, \bar{a}) = \frac{1}{N}\cdot\sum_{i=1}^{N}\phi(D_i^*, \bar{a}) = \frac{1}{N}\cdot\sum_{i=1}^{N}u_i \qquad (7)$$

ВИ передается посредством генов. Костный мозг (рисунок 5) является местом пребывания детекторов, параметры и состав которых определяются в процессе итерационного обучения ИИС с учителем с использованием *тренировочной* и *валидационной* выборок, которые являются *непересекающимися* подмножествами *обучающей* выборки. Для этого предложен специальный алгоритм (рисунок 6а), отличающийся от алгоритма (рисунок 3).

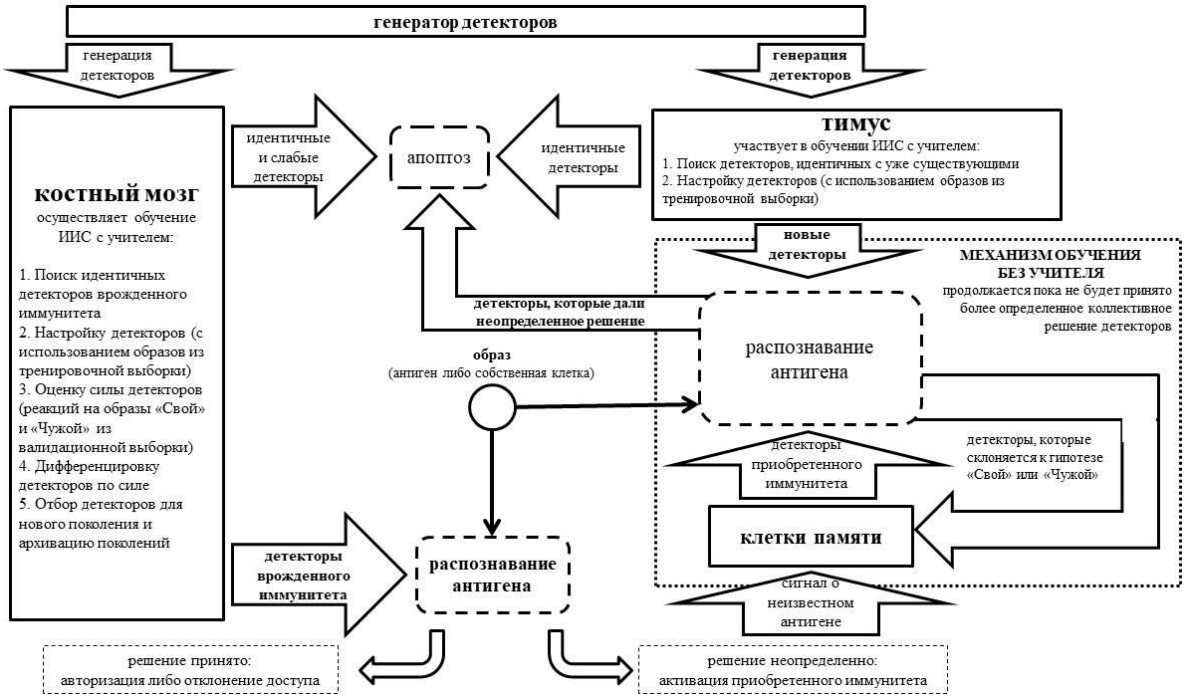

Рисунок 5. – Функциональная схема ИИС

ПИ развивается с течением жизни и определяет способность организма обезвреживать специфические антигены, которые попадали в организм ранее. Адаптивный иммунный ответ приводит к появлению клеток памяти (также представленных детекторами), которые долгое время пребывают в «спящем состоянии» до повторной встречи с антигеном. В разработанной модели ПИ формируется в процессе функционирования ИИС. Если решение об отнесении образа к классам «Свой» или «Чужой» является неоднозначным, могут генерироваться новые иммунокомпетентные детекторы.

Идея объединения классификаторов в комитет основана на теореме Кондорсе, которая утверждает: если мнения экспертов независимы, и вероятность правильного решения каждого из них больше 0,5, то с увеличением количества экспертов вероятность правильного коллективного решения возрастает и стремится к единице. Однако на практике решения классификаторов, играющих роль экспертов, в той или иной мере коррелированы, чем ниже коррелированность решающих правил, тем более ощутим положительный эффект при их комплексировании. Таким образом, имеются следующие *гиперпараметры*, которые влияют на эффективность комитета детекторов:

- RD – матрица коэффициентов корреляции $r(\bar{u}_i, \bar{u}_l)$ между реакциями всех возможных пар детекторов, где $\bar{u}_i$ – вектор реакций i-го детектора на примеры образов «Чужих» из тренировочной или валидационной выборки; N – количество детекторов; Δu – сила детекторов, их способность давать как можно более высокие показатели разницы средних уровней реакции на образы «Свой» $\mu_u^{(C)}$ и «Чужой» $\mu_u^{(Ч)}$ (8):

$$\Delta u = \mu_u^{(Ч)} - \mu_u^{(C)}, \quad \mu_u^{(Ч)} > \mu_u^{(C)} \qquad (8)$$

Апробирована стратегия снижения уровня коррелированности решений детекторов, которая оказалась недостаточно продуктивной. Не наблюдалось сходимости алгоритма: процесс обучения был длительным, и не всегда удавалось найти N детекторов с заданным минимальным уровнем взаимной коррелированности решений. По этой причине в настоящей работе выбрана стратегия повышения си-

лы детекторов при условии, что они не должны быть идентичными. При появлении в ИИС идентичных или слабых детекторов происходит их уничтожение (*апоптоз*, рисунок 5).

Детектор можно описать множеством параметров $D_i = \{\Psi_i, \breve{g}, x, \chi\}$. *Сгенерировать детектор* означает сгенерировать D_i. Необходимо, чтобы решения всех детекторов ВИ и ПИ не являлись полностью коррелированными. Поэтому после генерации детектора осуществляется проверка идентичности его параметров и параметров уже существующих детекторов. При обнаружении «двойника» его следует удалить и сгенерировать детектор снова. При этом значения параметров $\breve{g}$ можно считать равными, когда они отличаются менее чем на 10^{-1}. Чем сильней различия между D_i и D_l, тем менее коррелированы решения i-го и l-го детекторов.

В разработанной ИИС реализуется идея *случайных подпространств признаков*, но в отличие от алгоритма «случайный лес» Ψ_i задается с учетом корреляции между признаками. Этот прием называется *симметризацией корреляционных связей*. Другая идея заключается в объединении разнородных *случайных классификаторов*. Примером подобной техники является нейросетевое обобщения множества различных критериев. *Настройка детектора* связана с вычислением порога T_i и эталонных описаний признаков Θ_i (w_j, δ_j). Настроенный детектор обозначим $D^*_i = \{\Psi_i, \breve{g}, x, \chi, \Theta_i, T_i\}$.

При разработке алгоритмов обучения учтены следующие методы:

1. Бэггинг (bootstrap aggregating) – обучение базовых классификаторов на разных подмножествах обучающей выборки. Бэггинг уменьшает дисперсию голосов базовых классификаторов и помогает избежать переобучения.
2. Бустинг (boosting) – семейство алгоритмов машинного обучения, преобразующих слабые модели к сильным. Бустинг строит ансамбль путём тренировки каждого нового классификатора, уделяя больше внимания обучению на примерах, которые предыдущие модели классифицировали ошибочно.

В разработанном алгоритме обучения с учителем на каждой итерации происходит генерация новой популяции детекторов, которые настраиваются с учетом нескольких случайных тренировочных примеров (бэггинг), и выполняется промежуточная оценка их эффективности (рисунок 6а). Слабые детекторы уничтожаются и появляется новое поколение более эффективных детекторов. Мерой эффективности детекторов можно считать Δu (8). По результатам последней валидации вычисляются оценки $\mu_u^{(С)}$ и $\mu_u^{(Ч)}$ для коллективного решения детекторов ВИ. Эти параметры используются для построения *интервала неопределенности решения (ИНР)* $[\mu_u^{(С)}; \mu_u^{(Ч)}]$. ИНР является частью механизма подкрепления при онлайн-обучении модели в процессе ее функционирования. Этот механизм активируется при формировании ПИ.

На каждой итерации обучения с учителем синтезируются новые образы «Чужих» (рисунок 6а) путем скрещивания тренировочных примеров, которые хуже всего классифицируются детекторами ВИ. Скрещивание образов $\bar{a}_k$ и $\bar{a}_m$ происходит по формуле (9) (в соответствии с ГОСТ Р 52633.2-2010):

$$a_{c,j} = \frac{C_{syn} + 1 - c}{C_{syn} + 1} \cdot a_{k,j} + \frac{c}{C_{syn} + 1} \cdot a_{m,j} \qquad (9)$$

где C_{syn} – количество синтетических примеров, порождаемых парой «сильных Чужих» предыдущего поколения (в настоящей работе $C_{syn} = 1$), c – номер синтетического примера, j – номер признака. Синтетические образы добавляются в тренировочную выборку, что позволяет следующему поколению детекторов эф-

фективнее обучаться классифицировать образы «Чужих», наиболее близких к образам «Свой», с большей эффективностью (один из вариантов бустинга). Таким образом, *ИИС одновременно «учится» создавать образы более сильных «Чужих» и распознавать их.*

На скорость и эффективность алгоритма обучения с учителем влияют следующие основные параметры: $I_{ВИ}$ – количество итераций обучения; $N_{ВИ}$ – количество детекторов ВИ; Q – количество сильных «Чужих» (на каждой итерации генерируется $C_{syn} \cdot Q \cdot (Q-1)/2$ примеров). Тренировочная выборка «Чужие» увеличивается с каждой итерацией обучения при добавлении синтетических примеров, валидационная выборка остается неизменной.

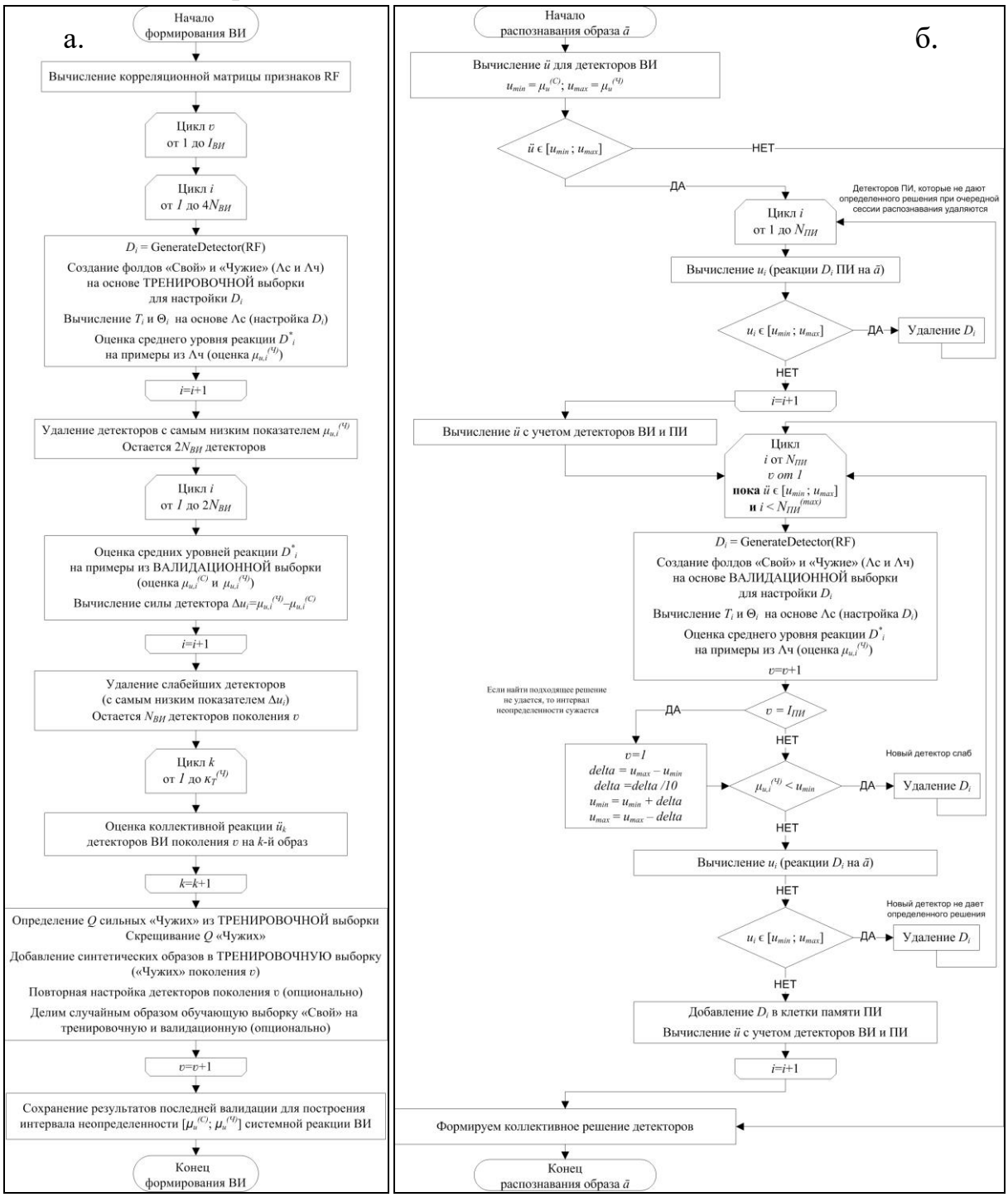

Рисунок 6. – Алгоритмы обучения: а. с учителем (или формирование ВИ); б. онлайн-обучения с подкреплением (или формирование ПИ)

Разработан алгоритм онлайн-обучения с подкреплением (рисунок 6б) для дообучения модели в процессе исполнения. Введем следующее правило, основанное на ИНР: если $u_i > \mu_u^{(Ч)}$ или $u_i < \mu_{u,i}^{(С)}$, то решение детектора D^*_i считается определенным, а если $\mu_u^{(С)} < u_i < \mu_u^{(Ч)}$, то его решение не определено. Если при распознавании образа коллективное (7) решение детекторов ВИ считается неопределенным, то активируется механизм ПИ (рисунок 6б). Тогда генерируются новые детекторы, которые настраиваются на других данных – примерах из валидационной выборки. Для новых детекторов вычисляются реакции u_i, но при формировании коллективного решения учитываются голоса только тех детекторов, которые дают определенный ответ (эти детекторы становятся клетками памяти), детекторы ПИ с неопределенным ответом уничтожаются.

На скорость и эффективность алгоритма онлайн-обучения влияют следующие параметры: $I_{ПИ}$ – количество итераций обучения; $N_{ПИ}^{(max)}$ – максимальное количество детекторов ПИ ($N_{ПИ}$ – их фактическое количество). Введение $I_{ПИ}$ позволяет избежать бесконечного цикла дообучения. Механизм ПИ компенсирует недостаток априорных знаний о классах «Свой» и «Чужой» и снижает вероятность возникновения концептуального дрейфа.

Проведен эксперимент с 2 общедоступными и 1 собственной базами клавиатурного почерка. Опыты проводились при различном объеме обучающей выборки «Свой»: от K_G=20 до K_G=40. Тренировочная $K_G^{(T)}$ и валидационная $K_G^{(V)}$ выборки «Свой» делились в соотношении: $K_G^{(T)} = 2 \cdot K_G^{(V)}$. Тренировочная $K_I^{(T)}$ и валидационная $K_I^{(V)}$ выборки «Чужих» включали по одному примеру от каждого испытуемого. Остальные примеры использовались в качестве тестовой выборки. Тестирование проводилось методом перекрёстного сравнения. Если обучать и тестировать систему на образах испытуемого, которые были записаны в разные дни, то репрезентативность выборки снижается и наблюдается большая разница (более 15%) между оценками коэффициента равной вероятности ошибок (EER) до и после онлайн-обучения.

Процесс обучения адаптивной модели ИИ с учителем оказался достаточно устойчивым (наблюдалась незначительная склонность к переобучению при высоких значениях $I_{ВИ}$). В зависимости от используемого набора данных лучшие показатели ошибок составили: EER=0,079, EER=0,053, EER= 0,026.

Разработанная адаптивная модель и алгоритмы ее обучения удовлетворяют основным принципам построения ИИС: 1. Распределенный характер вычислений и проявление эмерджентности; 2. Достаточно устойчивый процесс обучения; 3. Способность к адаптации; 4. Взаимодействие – ВИ формирует параметры, которые влияют на механизм подкрепления детекторов ПИ; 5. Надежность решений зависит от объема и чувствительности популяции детекторов; 6. Формирование памяти при помощи механизмов ВИ и ПИ.

Четвертая глава посвящена высоконадежной многофакторной аутентификации и безопасному комплексированию независимых биометрических данных. На данный момент не утверждено единого стандарта для безопасного объединения нескольких разнородных биометрических образов при защищенном исполнении процедур аутентификации (проект ГОСТ Р 52633.7 «Защита информации. Техника защиты информации. Высоконадежная мультибиометрическая аутентификация» находится на публичном обсуждении). Существующий стандарт ГОСТ Р 54411-2011 «Мультимодальные и другие мультибиометрические технологии» не рассматривает вопросы защищенного исполнения процедур биометрической

аутентификации, а только варианты объединения классификационных решений и биометрических образов при их последовательном или одновременном представлении. В настоящей работе предлагается иной вариант объединения биометрических образов.

Представлены комплексный метод и алгоритм трехфакторной высоконадежной аутентификации с последовательным предоставлением образов (рисунок 7). Первый фактор аутентификации – это тайный акустический образ уха. Под тайным образом понимается, что он не компрометируется в естественной среде (фотография уха не информативна для создания состязательных примеров). Для защиты данных уха от компрометации при хранении и передачи по каналам связи используется разработанная модель НПБК. Второй и третий факторы могут быть открытыми (рукописная подпись, фиксированная фраза) или тайными (рукописный и голосовой пароли). Для этих типов биометрических данных важно поддерживать актуальность, так как они изменчивы, поэтому для их анализа применяется адаптивная нейро-иммунная модель ИИ. Для защиты знаний адаптивной модели ИИ ее параметры после обучения шифруются на ключе, формируемом НПБК. Небольшое число ошибочных бит в ключе, генерируемом НПБК, может быть скорректировано за счет использования алгоритмов помехоустойчивого кодирования и кодов, исправляющих ошибки. Это позволяет балансировать показатели FRR и FAR на выходе НПБК. Если число ошибочных бит в ключе больше исправляющей способности кода (когда на вход НПБК поступает образ «Чужой»), то параметры адаптивной модели ИИ дешифруются неверно и доступ отклоняется, иначе исполняется алгоритм классификации и онлайн-обучения, который иллюстрируется на рисунок 6б.

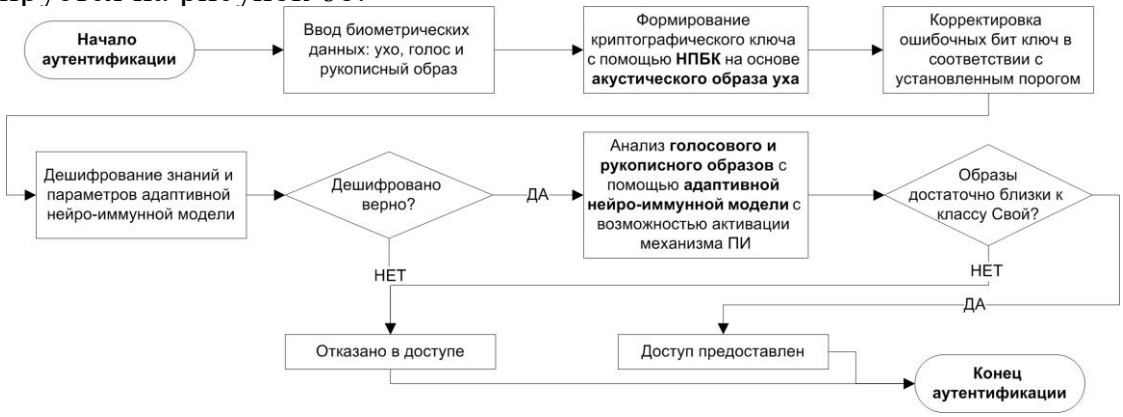

Рисунок 7. – Алгоритм трехфакторной аутентификации в защищенном режиме

Рассмотрим каждый фактор (тип образов) по отдельности.

Строение уха полностью закладывается до 8 лет. Длина, толщина и форма ушного канала различаются у людей. Ушной канал создает резонанс (в среднем 2,5 кГц), чтобы получить информацию о его строении, можно воздействовать на него акустическими волнами, которые отражаются от стенок. Эхо-сигнал имеет отличия, обусловленные индивидуальными особенностями канала. Параметры эхо-сигнала или его передаточной функции можно воспринимать как вектор биометрических признаков.

Создано устройство для регистрации характеристик уха, которое состоит из двух электретных микрофонов, звукоизолирующего корпуса наушников, двух динамиков и звуковой карты CREATIVE. Для сбора биометрических образов привлечено 75 человек (мужчин и женщин в равном соотношении, в возрасте от 18 до 40 лет *без отологических патологий*). Испытуемым предложено прослушать зву-

ковой моно-сигнал возрастающей и убывающей частоты, получаемый путем линейной частотной модуляции. Частота сигнала варьировалась от 1 кГц до 14 кГц, длительность составляла 10 секунд, громкость – 80 дБ. Эхо-сигнал одновременно регистрировался монтированными в корпус наушников микрофонами. Все испытуемые прослушали сигнал через два динамика по 15 раз, каждый раз снимая и надевая наушники (чтобы учесть зависимость эхо-сигнала от монтажа). Регистрируемый эхо-сигнал можно назвать эхограммой или *акустическим образом уха*. Сформирован набор данных в виде совокупности wav-файлов.

Для анализа эхо-сигналов применялось быстрое оконное преобразование Фурье (STFT). Спектрограммы сигналов были преобразованы в усредненный по всем окнам амплитудный спектр $\bar{A}'$ (размер окна W_{size}=65536, шаг W_{step}=16384), из которого удалялись отчеты, соответствующие частотам менее 1 кГц и более 20 кГц (рисунок 8). Применялись следующие оконные функции: Хэмминга, Блэкмана, Барлетта, прямоугольное, Гаусса, Лапласса.

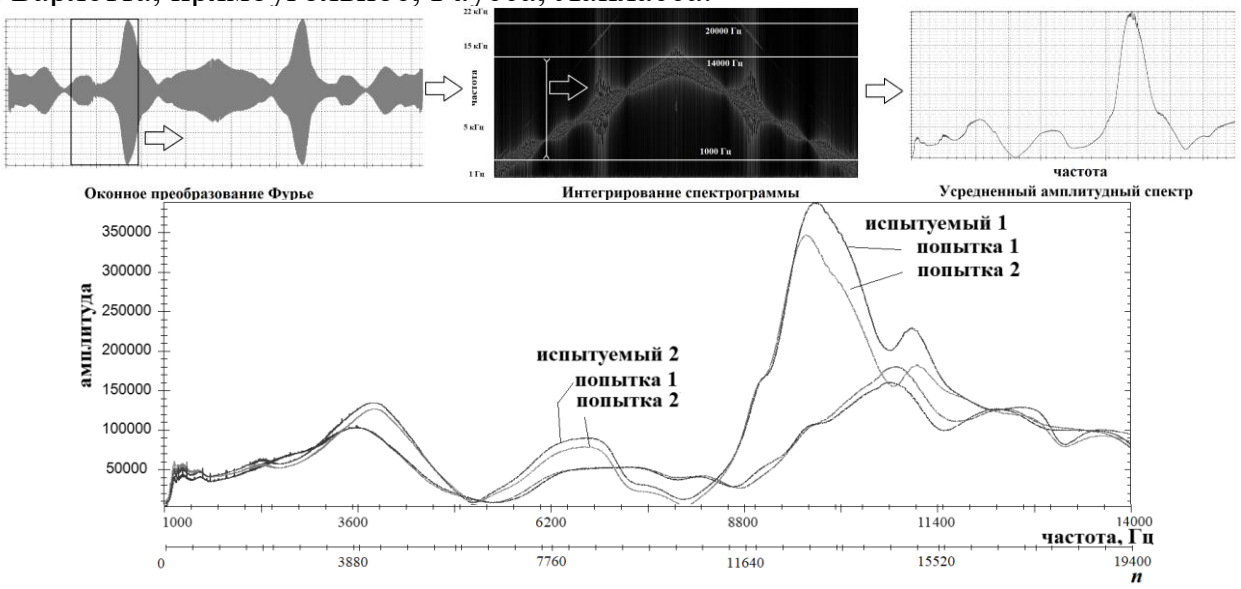

Рисунок 8. – Получение усредненного амплитудного спектра (вверху) и различия спектров у испытуемых (внизу)

Чтобы выявить локальные особенности усредненного спектра, построены кепстрограммы $K_{s,k}$ путем применения STFT к $\bar{A}'_{s,k}$ без операции логарифмирования. Так частотная шкала v спектральной функции $A'_{s,k}(v)$ принималась за временную шкалу, $\bar{A}'_{s,k}$ делился на частотные интервалы Δv в соответствии с размером окна W^*_{size} и шагом W^*_{step}. Далее каждый интервал раскладывался в ряд Фурье, и для интервалов строились *спектры кепстральных коэффициентов*. Комбинируя разные типы окон на этапе вычисления $\bar{A}'_{s,k}$ и $K_{s,k}$ можно получить больше информации. Проведен эксперимент по классификации образов испытуемых на основе спектральных и кепстральных признаков с помощью «наивного» байесовского классификатора и ИНС (8 архитектур, из них 6 сверточные сети). По результатам тестирования байесовский классификатор в сочетании с кепстральными признаками показал меньше ошибок (EER=0,0053 против EER=0,0266 для лучшей ИНС). Установлено, что кепстральные признаки являются более информативными, однако для обучения ИНС на кепстрограммах требуется больший объем обучающей выборки, что нереализуемо на имеющемся наборе данных.

Апробированы две схожие архитектуры нейросетевых автокодировщиков (на базе кодировщиков из 8 и 9 одномерных сверточных слоев и одного полносвязного, а также идентичных декодировщиков из 10 сверточных слоев) для по-

строения на их основе блоков извлечения признаков из усредненных спектров. Коррелированность признаков, извлекаемых разными нейронными сетями, в данном случае является желательным свойством.

Акустический образ уха имеет схожесть с голосовым сигналом, как и их усредненные спектры. Решено использовать следующую схему переноса обучения. Из речевых наборов данных TIMIT и VoxCeleb1 в совокупности было извлечено 71264 голосовых образов дикторов. Эти образы были преобразованы в усредненные спектры (W_{size}=4096 и W_{step}=2048). Для аугментации данных использовались 4 оконные функции: Хэмминга, Блэкмана, Барлетта, прямоугольная, в результате получено 285056 усредненных голосовых спектров, которые использовались для обучения двух автокодировщиков. ИНС обучались (оптимизатор Adam) с возрастающим объемом батчей (далее размер батча/количество эпох): 256/10, 512/10, 1024/3, 2048/3, 4096/2, 8192/2.

Проведен эксперимент по верификации личности испытуемых с помощью двух различных моделей НПБК – на базе корреляционных нейронов (№1) и ГОСТ Р 52633.5 (№2). Все образы были обработаны кодировщиками. Наилучшими результатами аутентификации субъектов являются следующие:

– НПБК №1: EER=0,0238 (FRR=0,093, FAR<0,0001), L=8192, K_G=6, K_I=49.
– НПБК №2: EER=0,03136 (FRR=0,2342, FAR<0,0001), L=716, K_G=8, K_I=49;

Достоверность оценок для вероятностей FRR и FAR составляет 0,99 и 0,96. *Предложенная модель НПБК при меньшей тренировочной выборке дает меньше ошибок и более, чем в 10 раз увеличивает длину ключа, чем НПБК, обучаемый по ГОСТ Р 52633.5-2011.* Метод аутентификации по акустическим параметрам уха на основе предложенной модели НПБК можно применять как в составе алгоритма (рисунок 7), так и отдельно. Для анализа рукописных и голосовых образов сформирован набор данных. Рукописные образы собраны с использованием планшета Wacom (частота опроса 200 Гц, 1024 уровня давления), голосовые – с использованием микрофонов Pioneer, Sony (диапазон частот 70 – 12000 Гц). Также из открытых источников взяты примеры для расширения тестовой выборки «Чужих». Набор данных можно разделить на выборки:

– «Все Свои»: 260 подписантов и 260 дикторов (пол и возраст распределены равномерно от 18 до 35 лет) воспроизвели образ 90 раз, данные собраны в три этапа с интервалом в несколько недель, на каждом этапе испытуемый ввел 30 примеров, на третьем этапе испытуемые находились в сонном ПФС;
– «Неизвестные Чужие»: по одному тестовому примеру других 6500 рукописных и 6500 голосовых образов, воспроизведенных другими субъектами.

Оцифрованный рукописный образ состоит из функций координат *x_coord(t), y_coord(t)* и давления пера на планшет *pressure(t)*, где t – это время в дискретной форме. Рассмотрено 2 набора признаков (таблица 1): с поддержкой давления и без (n=782/n= 521). Для голосовых сигналов также апробировано 2 набора признаков (таблица 2): при частоте дискретизации сигнала 24 кГц и 8 кГц (n=570/n=350) для низко информативных каналов передачи данных.

Метод двухфакторной аутентификации по голосовым и рукописным образам можно также использовать как отдельно, так и в составе алгоритма на рисунок 7. Экспериментальная оценка надежности двухфакторной аутентификации показала следующие результаты: FRR = 0,03 при FAR < 10^{-10}.

Использование трех факторов аутентификации, а также предложенный метод и алгоритм (рисунок 7) многократно повышают надежность и защищенность

биометрической системы от деструктивных воздействий. Экспериментальная оценка надежности трехфакторной аутентификации показала следующие результаты: FRR = 0,12 при FAR < 10^{-14}.

Таблица 1. Краткое описание методик извлечения признаков рукописного образа

Группа признаков рукописного образа	*n*
образ делится на 16 равных по числу точек отрезков, строится матрица расстояний между их краями в 2-х и 3-х мерном пространстве (*pressure*(*t*) – третье измерение)	240 / 120
вычисление коэффициентов корреляции между *x_coord*(*t*), *y_coord*(*t*), *pressure*(*t*), их производных и функцией скорости пера $v_{xy}(t)$, производной от *x_coord*(*t*), *y_coord*(*t*)	21 / 10
вычисление параметров внешнего вида образа – угол наклона, отношение длины к ширине, центр в 2-х (3-х) мерном пространстве, описываемый 2 или 3 координатами	5 / 4
вычисление средних значений фрагментов функций *pressure*(*t*), *x_coord'*(*t*), *y_coord'*(*t*), $v_{xy}(t)$ (образ делится на 5 равных по числу точек отрезков)	20 / 15
вычисление детализирующих коэффициентов вейвлет преобразования Хаара (алгоритм Малла), полученных на 4 нижних уровнях разложения для *x_coord*(*t*), *y_coord*(*t*), *pressure*(*t*), $v_{xy}(t)$ (функции сначала приводились к 128 отчетам (интерполяция))	240 / 180
вычисление усредненного амплитудного спектра с помощью STFT (размер окна – 128 отчетов, шаг – 16 отчетов) для *x_coord*(*t*), *y_coord*(*t*), *pressure*(*t*), $v_{xy}(t)$	256 / 192

Таблица 2. Краткое описание методик извлечения признаков голосовых паролей

Группа признаков голосовых паролей	*n*
вычисление усредненного по всем окнам амплитудного спектра низких частот речевого сигнала, вычисленного с помощью STFT (размер окна – 2048(512) при F=24 (8) кГц, шаг – 16). Предварительно речевой сигнал нормируется по энергии, удаляется тишина	40 / 20
вычисление коэффициентов нижних частот кепстра, который берется от полного усредненного по всем окнам амплитудного спектра, получаемого в соответствии с 2.1	40 / 20
вычисление коэффициентов нижних частот кепстра, который берется от полного логарифмированного усредненного амплитудного спектра	40 / 20
вычисление кепстра второго порядка от полных кепстров 2.1 и 2.2 (кепстры 2.1 и 2.2 повторно подвергаются прямому преобразованию Фурье)	256 / 128
подсчет частоты переходов сигнала через нулевое деление окном (размер окна – 2048(512) при F=24 (8) кГц, шаг – 16), грубо характеризуют ЧОТ (нулевую форманту)	64 / 32
вычисление амплитудного спектра функции автокорреляции речевого сигнала	128
вычисление частоты переходов через «ноль» и экстремумов функции автокорреляции	2

Пользователям рекомендуется периодически (не реже, чем раз в пол года) проходить процедуру аутентификации, чтобы нейро-иммунная модель адаптировалась к дрейфу и точность не снижалась.

В **пятой главе** описываются разработанная технология автоматического синтеза и обучения нейросетевых моделей доверенного ИИ. Обозначены ключевые положения первой редакции разработанного стандарта, основанного на данной технологии, который в значительной степени расширяет уже имеющийся стандарт ГОСТ Р 52633.5-2011 и отличается от него тем, что: он распространяется на нейросетевые преобразователи не только биометрических, но и любых иных образов в код, а также на задачи идентификации образов на открытом множестве. Стандарт позволяет повысить длину пароля и криптографического ключа, связываемого с классом «Свой»; снизить вероятности ошибок 1-го и 2-го рода; улучшить хэширующие (перемешивающие) свойства НПБК; повысить уровень защищенности знаний НПБК от компрометации; сделать невозможным или, по крайней мере, усложнить несанкционированное управление ИИ путем манипуляций с моделями машинного обучения, анализ логики работы ИИ с целью повлиять на его решения, извлечение знаний из памяти ИИ (в том числе путем зондирования

модели), а также их интерпретацию; снизить вероятность успеха состязательных атак, реализуемых злоумышленником путем наложения шумов на исходный образ «Свой» или путем создания синтетических примеров образов «Чужой».

Приведено описание внедрений результатов работы на примере ООО «Открытый код», ООО «Системы информационной безопасности», ООО «КРАФТ ЛАБ», ООО «АИ ЗИОН», МСЧ № 4 Омской области, особое внимание уделяется применению предложенной технологии при разработке линейки программных продуктов AIConstructor (AIC). AIC desktop ориентирован на исследования по машинному обучению и классификации образов в области ИБ. Программный комплекс поддерживает множество форматов данных (xml, txt, bmp, wav, edf, csv и др.), имеет конструктор глубоких нейронных сетей. AIC ModelOps Platform – это система управления жизненным циклом ИИ для цифровой трансформации предприятий. Программный продукт включает следующие модули: управление экспериментом (интеграция с фреймворками, отслеживание экспериментов, инструменты для командной работы исследователей, редактор кода эксперимента); конструктор конвейеров обработки данных; диагностика (мониторинг модели, определение дрейфа). Описана разработанная библиотека машинного обучения на базе корреляционных нейронов, интегрированная в AIC ModelOps Platform.

На базе результатов разработаны образовательные программы и внедрены в учебный процесс ФГАОУ ВО СПбГЭТУ «ЛЭТИ» и ФГАОУ ВО ОмГТУ (по дисциплинам «Распознавание образов», «Машинное обучение в приложениях биометрии», «Биометрия и защита информации», «Этика и правовые проблемы искусственного интеллекта», «Защищенное исполнение искусственного интеллекта», «Доверенный искусственный интеллект»).

В **заключении** сформулированы основные результаты и итоги работы, намечены направления перспективных исследований.

В приложении 1 представлены результаты дополнительных экспериментов по анализу биометрических образов, в приложении 2 представлены копии актов внедрения результатов. Приложения 3 и 4 касаются разработанного стандарта (письмо в ТК 164, список разработчиков, план проспект). Приложения 5 и 6 касаются вступления в состав экспертов от России в ISO/IEC JTC 1/SC 42 «Artificial intelligence». Приложение 7 содержит пример работы в AIC desktop, приложение 8 – реализацию НПБК на базе корреляционных нейронов на C#, приложения 9 и 10 – копии свидетельств о регистрации программ и патента.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ И ВЫВОДЫ

В диссертационной работе на основе выполненного автором исследования решена актуальная научная повышения надежности многофакторной биометрической аутентификации и защищенности биометрических систем от компьютерных атак на основе технологии защищенного исполнения нейросетевых моделей и алгоритмов ИИ, имеющая важное хозяйственное значение с точки зрения обеспечения информационной безопасности компьютерных ресурсов и конфиденциальных данных, а также знаний, моделей и алгоритмов ИИ.

Получены следующие основные результаты:

1. Разработана концепция защищенного исполнения нейросетевых алгоритмов ИИ, которая позволяет сформировать устойчивость модели к извлечению знаний.

Это достигается путем преобразования корреляционных связей между признаками в высокоинформативные мета-признаки Байеса-Минковского, которые сложно фальсифицировать. Установлено, что один мета-признак может содержать в 2-3 раза больше информации, чем содержится суммарно в паре исходных признаков, от которых он порожден. Доказано, что корреляция между признаками увеличивает количество информации о классифицируемом образе. Предложены отображения для перехода в пространства мета-признаков Байеса-Минковского, что не требует хранения какой-либо дифференциальной информации о параметрах классов образов. Свойства пространств мета-признаков исследованы экспериментально.

2. Разработаны модель корреляционных нейронов и модель нейросетевого преобразователя биометрия-код на их основе, анализирующие корреляционные связи между признаками вместо признаков, а также робастной алгоритм их автоматического синтеза и обучения на малых выборках. Это позволило повысить защищенность биометрических данных от компрометации и длину ключа, связываемого с биометрическими образами субъектов, снизить вероятность ошибок биометрической аутентификации в защищенном режиме исполнения и повысить устойчивость ИИ к состязательным атакам.

3. Разработаны адаптивная нейро-иммунная модель и алгоритмы ее пакетного обучения с учителем и онлайн-обучения с подкреплением, позволяющие предупредить или снизить влияние концептуального дрейфа, даже если исходная обучающая выборка недостаточно репрезентативна или незначительна в объеме. Алгоритм обучения с учителем позволяет сформировать врожденный иммунитет модели, позволяющий разделять входные образы на два класса. В процессе функционирования модель адаптируется к изменению данных, используя алгоритм онлайн-обучения с подкреплением, в результате чего формируется приобретенный иммунитет, корректирующий решения модели в спорных случаях.

4. Разработаны методы и алгоритм высоконадежной многофакторной аутентификации на основе рукописных и голосовых образов, а также акустических образов уха с обеспечением защиты биометрических данных от компрометации. Новизна заключается в использовании нового типа биометрических данных – акустических параметров ушного канала, получаемых методом эхолокации, а также учете информативности, стабильности и приоритезации признаков, совместным использованием НПБК и нейро-иммунной модели, способе кепстрального анализа сигналов. Акустические образы уха не компрометируются в естественной среде, так как фотография уха неинформативна для синтеза состязательных примеров. Предложенные методы и алгоритм дают более низкий процент ошибок по сравнению с известными мировыми аналогами: $FRR = 0{,}12$ при $FAR < 10^{-14}$ и $FRR = 0{,}03$ при $FAR < 10^{-10}$.

5. Разработана технология автоматического синтеза и обучения нейросетевых моделей доверенного ИИ для высоконадежной биометрической аутентификации и других ответственных приложений ИИ. На базе технологии разработана линейка программных продуктов AIConstructor и первая редакция национального стандарта ГОСТ Р «Искусственный интеллект. Нейросетевые алгоритмы в защищённом исполнении. Автоматическое обучение нейросетевых моделей на малых выборках в задачах классификации». Это первый стандарт, который регламентирует создание и обучение нейросетевых моделей ИИ, исполняемых в защищенном от исследования режиме. Выполнено 8 внедрений результатов на 7 предприятиях.

Перспективы дальнейшего развития темы. Все более актуальными становятся вопросы создания систем доверенного ИИ, минимизации рисков, связанных с внедрением и поддержкой систем ИИ, защиты от компьютерных атак и обеспечения функциональной безопасности ИИ. Основные перспективы развития темы состоят именно в данном направлении. Защищенное исполнение алгоритмов ИИ необходимо на объектах критической информационной инфраструктуры. Перспективным направлением для развития темы также является управление жизненным циклом ИИ. Разработанный аппарат может применяться для сокращения объемов обучающей выборки, формирования инструментов повышения объяснимости решений, для обнаружения и корректировки дрейфа моделей ИИ.

Кроме того, развитие и применение разработанных концепции, моделей и алгоритмов видится перспективным в области стандартизации ИИ в тех отраслях, в которых ставятся повышенные требования к безопасности.

СПИСОК ОСНОВНЫХ ПУБЛИКАЦИЙ

Статьи в научных изданиях из Перечня рецензируемых научных изданий, рекомендованных ВАК, либо в научных изданиях, индексируемых в базе данных RSCI:

1. Сулавко, А.Е. Метод биометрической аутентификации на основе кепстральных характеристик эхограмм наружного уха и нейросетевого преобразователя биометрия-код / А.Е. Сулавко, А.А. Храмов// Прикладная информатика. Т. 17, № 1. – С. 69–82. 2022.
2. Иванов, А.И. Проект третьего национального стандарта России по быстрому автоматическому обучению больших сетей корреляционных нейронов на малых обучающих выборках биометрических данных / А.И. Иванов, А.Е. Сулавко // Вопросы кибербезопасности. – 2021. – № 3. – С. 84–93. – DOI:10.21681/2311-3456-2021-3-84-93
3. Преобразователь образов голосовых паролей дикторов в криптографический ключ на основе комитета предварительно обученных сверточных нейронных сетей / Сулавко А.Е., Иниватов Д.П., Стадников Д.Г. [и др.] // Вопросы защиты информации. – 2021. - №4. – С. 23-33.
4. Оценка идентификационного потенциала электроэнцефалограмм с использованием статистического подхода и сверточных нейронных сетей / А.Е. Сулавко, П.С. Ложников, А.Г. Чобан [и др.]// Информационно-управляющие системы. – 2020. – № 6. – С. 37–49.
5. Сулавко, А.Е. Высоконадежная аутентификация по рукописным паролям на основе гибридных нейронных сетей с обеспечением защиты биометрических эталонов от компрометации / А.Е. Сулавко // Информационно-управляющие системы. – 2020. – № 4. – С. 61–77. – DOI: 10.31799/1684-8853-2020-4-61-77
6. Оценка ускорения вычислений от перехода к воспроизведению эффектов нейродинамики при анализе числа возможных состояний больших сетей искусственных нейронов / А.И. Иванов, А.И. Газин, А.Е. Сулавко, Д.Г. Стадников // Вопросы защиты информации. – 2020. – № 4. – С. 32–38.
7. Искусственный интеллект в защищенном исполнении на базе иммунных сетевых моделей распознавания образов на примере преобразователей биометрия-код / Е.В. Шалина, Н.В. Малинин, А.Е. Сулавко, Д.Г. Стадников // Вопросы защиты информации. – 2020. – № 2. – С. 31–40.
8. Гарипов, И.М. Методы распознавания личности на основе анализа характеристик наружного уха (Обзор) / И.М. Гарипов, А.Е. Сулавко, И.А. Куприк // Вопросы защиты информации. – 2020. – № 1. – С. 33–41.
9. Сулавко, А.Е. Биометрическая аутентификация пользователей информационных систем по клавиатурному почерку на основе иммунных сетевых алгоритмов / А.Е. Сулавко, Е.В. Шалина // Прикладная информатика. – 2019. – № 3 (81). – С. 39–53. – DOI: 10.24411/1993-8314-2019-10014
10. Иммунные алгоритмы распознавания образов и их применение в биометрических системах (Обзор) // А.Е. Сулавко, Е.В. Шалина, Д.Г. Стадников, А.Г. Чобан // Вопросы защиты информации. – 2019. – № 1. – С. 38–46.

11. Сулавко, А.Е. Тестирование нейронов для распознавания биометрических образов при различной информативности признаков / А.Е. Сулавко // Прикладная информатика. – 2018. – № 1. – С. 128–143.
12. Сулавко, А.Е. Архитектура перспективных нейронов для обработки биометрических данных с высокой взаимной корреляционной зависимостью / А.Е. Сулавко // Вопросы защиты информации. – 2018. – № 1. – С. 35–48.
13. Биометрическая аутентификация по клавиатурному почерку с учетом силы нажатия на клавиши, параметров вибрации и движения рук оператора / А.Е. Сулавко, А.Р. Хамзин, А.А. Лыжин [и др.] // Вопросы защиты информации. – 2018. – № 2. – С. 41–50.
14. Анализ методов распознавания образов человека по особенностям электроэнцефалограмм (обзор) / А.Е. Сулавко, А.И. Куприк, М.А. Старков, Д.Г. Стадников // Вопросы защиты информации. – 2018. – № 4. – С. 36–46.
15. Распознавание психофизиологических состояний пользователей на основе скрытого мониторинга действий в компьютерных системах / Васильев В.И., Сулавко А.Е., Борисов Р.В. [и др.] // Искусственный интеллект и принятие решений. - 2017. - №3. - С. 95-111.
16. Сулавко, А.Е. Влияние психофизиологического состояния подписантов на биометрические параметры рукописных образов и результаты их верификации / А.Е. Сулавко, А.Е. Самотуга // Информационно-управляющие системы. – 2017. – № 6. – С. 29–42. – DOI: 10.15217/issn1684-8853.2017.6.29
17. Аутентификация пользователей компьютера на основе клавиатурного почерка и особенностей лица / П.С. Ложников, А.Е. Сулавко, Е.В. Бурая, В.Ю. Писаренко // Вопросы кибербезопасности. – 2017. – № 3. – С. 24–34. – DOI: 10.21681/2311-3456-2017-3-24-34
18. Комплексирование независимых биометрических признаков при распознавании субъектов на основе сетей квадратичных форм, персептронов и меры ХИ-модуль / А.Е. Сулавко, А.В. Еременко, Е.В. Толкачева, Р.В. Борисов // Информационно-управляющие системы. – 2017. – № 1 (86). – С. 50–62.
19. Распознавание пользователей компьютерных систем по клавиатурному почерку с учетом регистрации дополнительных признаков при помощи специальных датчиков / А.В. Еременко, А.Е. Сулавко, Д.В. Мишин, А.А. Федотов // Датчики и системы. – 2017. – № 3. – С. 9–16.
20. Идентификационный потенциал клавиатурного почерка с учетом параметров вибрации и силы нажатия на клавиши / А.В. Еременко, А.Е. Сулавко, Д.В. Мишин, А.А. Федотов // Прикладная информатика. 2017. – Т. 12, № 1 (67). – С. 79–94.
21. Генерация ключевых последовательностей и верификация субъектов на основе двумерного изображения лица / А.Е. Сулавко, А.В. Еременко, С.С. Жумажанова, Е.В. Бурая // Автоматизация процессов управления. – 2017. – № 1. – С. 58–66.
22. Идентификация психофизиологических состояний подписантов по особенностям воспроизведения автографа / А.Е. Сулавко, А.В. Еременко, Е.А. Левитская, А.Е. Самотуга // Информационно-измерительные и управляющие системы. – 2017. – № 1. – С. 40–48.
23. Оценка информативности характеристик рукописных образов для идентификации психофизиологического состояния человека / А.Е. Сулавко, А.В. Еременко, Е.А. Левитская, А.Е. Самотуга, Е.В. Толкачева // Информационно-измерительные и управляющие системы. – 2017. – №11. – С. 35–46.
24. Комплексная система распознавания водителей транспортных средств и их психофизиологического состояния по динамическим биометрическим признакам / А.Е. Сулавко, С.С. Жумажанова, З.В. Семенова [и др.] // Автоматизация. Современные технологии. – 2017. – № 8. – С. 373–380.
25. Сулавко, А.Е. Влияние функционального состояния оператора на параметры его клавиатурного почерка в системах биометрической аутентификации / А.Е. Сулавко // Датчики и системы. – 2017. – № 11. – С. 19–30.
26. Влияние психофизиологического состояния диктора на параметры его голоса и результаты биометрической аутентификации по речевому паролю / А.Е. Сулавко, А.В. Еременко, Р.В. Борисов, Д.П. Иниватов // Компьютерные инструменты в образовании. – 2017. – № 4. – С. 29–47.

27. Влияние психофизиологического состояния подписанта на результаты его идентификации по рукописному образу естественным и искусственным интеллектами // А.Е. Сулавко, С.С. Жумажанова, А.А. Нигрей, Л.Н. Закутнева // Безопасность информационных технологий. – 2017. – Т. 24, № 4. – С. 87–97.
28. Экспериментальная оценка надежности верификации подписи сетями квадратичных форм, нечеткими экстракторами и персептронами / П.С. Ложников, А.Е. Сулавко, А.В. Еременко, Д.А. Волков // Информационно-управляющие системы. – 2016. – № 5. – С. 73–85.
29. Нечеткий экстрактор для генерации ключей шифрования на основе параметров клавиатурного почерка / А.Е. Сулавко, А.В. Еременко, Е.В. Толкачева, С.С. Жумажанова // Информационные технологии и вычислительные системы. – 2016. – № 4. – С. 69–79.
30. Еременко, А.В. Современное состояние и пути модернизации преобразователей биометрия-код / А.В. Еременко, А.Е. Сулавко, Д.А. Волков // Информационные технологии. – 2016. – № 3. – С. 203–210.
31. Метод защиты текстовых документов на электронных и бумажных носителях на основе скрытого биометрического идентификатора субъекта, получаемого из подписи / А.В. Еременко, А.Е. Сулавко, Е.В. Толкачева, Е.А. Левитская // Информационные технологии. – 2016. – Т. 22, № 8. – С. 628–634.
32. Сулавко, А.Е. Генерация криптографических ключей на основе голосовых сообщений / А.Е. Сулавко, А.В. Еременко, Р.В. Борисов // Прикладная информатика. – 2016. – № 5. – С. 76–89.
33. Еременко, А.В. Двухфакторная аутентификация пользователей компьютерных систем на удаленном сервере по клавиатурному почерку / А.В. Еременко, А.Е. Сулавко // Прикладная информатика. – 2015. – № 6. – С. 48–59.
34. Сулавко, А.Е. Разграничение доступа к информации на основе скрытого мониторинга действий пользователей в информационных системах: портрет нелояльного сотрудника / А.Е. Сулавко, А.В. Еременко, Е.А. Левитская // Известия Транссиба. – 2015. – № 1 (21). – С. 80–89.
35. Сулавко, А.Е. Метод сжатия собственных областей классов образов в пространстве малоинформативных признаков / А.Е. Сулавко, А.В. Еременко // Искусственный интеллект и принятие решений. – 2014. – № 2. – С. 102–109.
36. Сулавко, А.Е. Исключение искаженных биометрических данных из эталона субъекта в системах идентификации / А.Е. Сулавко, А.В. Еременко, А.Е. Самотуга // Информационные технологии и вычислительные системы. – 2013. – № 3. – С. 96–101.
37. Еременко, А.В. Исследование алгоритма генерации криптографических ключей из биометрической информации пользователей компьютерных систем / А.В. Еременко, А.Е. Сулавко // Информационные технологии. – 2013. – № 11. – С. 47–51.
38. Епифанцев, Б.Н. Альтернативные сценарии авторизации при идентификации пользователей по динамике подсознательных движений / Б.Н. Епифанцев, П.С. Ложников, А.Е. Сулавко // Вопросы защиты информации. – 2013. – № 2. – С. 28–35.

Публикации в отечественных журналах из перечня изданий ВАК, включенных в международные базы Web of Science, Scopus:

39. Сулавко, А.Е. Абстрактная модель искусственной иммунной сети на основе комитета классификаторов и ее использование для распознавания образов клавиатурного почерка / А.Е. Сулавко // Компьютерная оптика. – 2020. – Т. 44, № 5. – С. 830–842. – DOI: 10.18287/2412-6179-CO-717
40. Сулавко, А.Е. Высоконадежная двухфакторная биометрическая аутентификация по рукописным и голосовым паролям на основе гибких нейронных сетей / А.Е. Сулавко // Компьютерная оптика. – 2020. – Т. 44, № 1. – С. 82–91. – DOI: 10.18287/2412-6179-CO-567
41. Identification of the Psychophysiological State of the User Based on Hidden Monitoring in Computer Systems / V.I. Vasilyev, A.E. Sulavko, S.S. Zhumazhanova, R.V. Borisov // Scientific and Technical Information Processing. – 2018. – Vol. 45, № 6. – P. 398–410. DOI: 10.3103/S0147688218060096
42. Иванов, А.И. Оценка надежности верификации автографа на основе искусственных нейронных сетей, сетей многомерных функционалов Байеса и сетей квадра-

тичных форм / А.И. Иванов, П.С. Ложников, А.Е. Сулавко // Компьютерная оптика. – 2017. – Т. 41, № 5. – С. 765–774. – DOI: 10.18287/2412-6179-2017-41-5-765-774

43. Identification Potential of Online Handwritten Signature Verification / B.N. Epifantsev, P.S. Lozhnikov, A.E. Sulavko, S.S. Zhumazhanova // Optoelectronics, Instrumentation and Data Processing. – 2016. – № 3 (52). – P. 238–244. – DOI: 10.3103/S8756699016030043

Публикации в изданиях, включенных в международные базы Web of Science, Scopus:

44. Sulavko, A.E. Biometric-Based Key Generation and User Authentication Using Acoustic Characteristics of the Outer Ear and a Network of Correlation Neurons / A.E. Sulavko // Sensors. – 2022. – Vol. 22. – P. 9551. – DOI: 10.3390/s22239551
45. Sulavko, A.E. Personal Identification Based on Acoustic Characteristics of the Outer Ear Using Cepstral Analysis, Bayesian Classifier and Artificial Neural Networks / A.E. Sulavko, A.E. Samotuga, I.A. Kuprik // IET Biometrics. – 2021. – Vol.10. – №6. – P. 692–705. – DOI: 10.1049/bme2.12037
46. Sulavko, A.E. Bayes-Minkowski measure and building on its basis immune machine learning algorithms for biometric facial identification / A.E. Sulavko // Journal of Physics: Conf. Series. – 2020. – Vol. 1546. – P. 012103-1–012103-7. – DOI:10.1088/1742-6596/1546/1/012103
47. Statistical approach for subject's state identification by face and neck thermograms with small training sample / S.S. Zliumazhanova, A.E. Sulavko, D.B. Ponomarev, V.A. Pasenchuk // IFAC-PapersOnLine. – 2019. – Vol. 52, № 25. – P. 46–51, DOI: 10.1016/j.ifacol.2019.12.444.
48. Biometric authentication on the basis of lectroencephalograms parameters / A.E. Sulavko, A.E. Samotuga, D.G. Stadnikov, V.A. Pasenchuk, S.S. Zhumazhanova // Journal of Physics: Conference Series. – 2019. – Vol. 1260, №2. – P. 022011. – DOI:10.1088/1742-6596/1260/2/022011
49. Flexible fast learning neural networks and their application for building highly reliable biometric cryptosystems based on dynamic features / V.I. Vasilyev, P.S. Lozhnikov, A.E. Sulavko [et al] // IFAC-PapersOnLine. – 2018. – Vol. 51, № 30. – P. 527–532. – DOI: 10.1016/j.ifacol.2018.11.272
50. Subjects Authentication Based on Secret Biometric Patterns Using Wavelet Analysis and Flexible Neural Networks / A.E. Sulavko, D.A. Volkov, S.S. Zhumazhanova, R.V. Borisov // XIV International Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering (APEIE). – Novosibirsk, 2018. – P. 218–227, DOI: 10.1109/APEIE.2018.8545676
51. Sulavko, A.E. Biometric pattern recognition using wide networks of gravity proximity measures / A.E. Sulavko, S.S. Zhumazhanova // Journal of Physics: Conf. Series. – 2018. – Vol. 1050: Mechanical Science and Technology Update. – P. 012082-1–012082-13. – DOI: 10.1088/1742-6596/1050/1/012082
52. Sulavko, A.E. Perspective Neural Network Algorithms for Dynamic Biometric Pattern Recognition in the Space of Interdependent Features / A.E. Sulavko, S.S. Zhumazhanova, G.A. Fofanov // Dynamics of Systems, Mechanisms and Machines: conference proceeding, 13–15 November 2018 / Omsk State Technical University. – Omsk, 2018. – P. 1–12, DOI: 10.1109/Dynamics.2018.8601440
53. Sulavko, A.E. Users' identification through keystroke dynamics based on vibration parameters and keyboard pressure / A.E. Sulavko, A.A. Fedotov, A.V. Eremenko // Dynamics of Systems, Mechanisms and Machines: conference proceeding, 14–16 November 2017 / Omsk State Technical University. – Omsk, 2017. – P. 1–7, DOI: 10.1109/Dynamics.2017.8239514
54. Sulavko, A.E. Human psychophysiological state recognition based on analysis of thermograms of face and neck regions / A.E. Sulavko, S.S. Zhumazhanova // Dynamics of Systems, Mechanisms and Machines: conference proceedings (Omsk, 14-16 November 2017) / Omsk State Technical University. – Omsk, 2017. – DOI: 10.1109/Dynamics.2017.8239515
55. Methods of Generating Key Sequences based on Parameters of Handwritten Passwords and Signatures / P.S. Lozhnikov, A.E. Sulavko, A.V. Eremenko, D.A. Volkov // Information. MDPI. – 2016. – №7(4). – P.59.DOI: 10.3390/info7040059

56. Lozhnikov, P.S. Usage of fuzzy extractors in a handwritten-signature based technology of protecting a hybrid document management system / P.S. Lozhnikov, A.E. Sulavko, D.A. Volkov // 10th International Conference on Application of Information and Communication Technologies (AICT), 12–14 October 2016. – Baku, 2016. – P. 395–400, DOI: 10.1109/ICAICT.2016.7991728
57. Identification potential capacity of typical hardware for the purpose of hidden recognition of computer network users / V.I. Vasilyev, A.E. Sulavko, A.V. Eremenko, S.S. Zhumazhanova // Dynamics of Systems, Mechanisms and Machines: conference proceeding, 15–17 November 2016 / Omsk State Technical University. – Omsk, 2016. – P. 1–5, DOI: 10.1109/Dynamics.2016.7819106
58. Lozhnikov, P.S. Personal Identification and the Assessment of the Psychophysiological State While Writing a Signature / P.S. Lozhnikov, A.E. Sulavko, A.E. Samotuga // Information. – 2015. – №6. – P. 454–466. – DOI: 10.3390/info6030454
59. Lozhnikov, P.S. Application of noise tolerant code to biometric data to verify the authenticity of transmitting information / P.S. Lozhnikov, A.E. Sulavko, D.A. Volkov / SIBCON.2015.7147126 // 2015 International Siberian Conference on Control and Communications (SIBCON). – IEEE, 2015. – P. 1–3. – DOI: 10.1109/SIBCON.2015.7147126

Другие публикации по теме диссертации:

60. Иванов, А.И. Использование сетей корреляционных нейронов с многоуровневым квантованием: защита от извлечения знаний из параметров решающего правила: препринт / А.И. Иванов, А.Е. Сулавко. – Пенза: Изд-во ПГУ, 2020. – 48 с.(
61. Сулавко, А.Е. Искусственный интеллект в защищенном исполнении / А.Е. Сулавко // Информационная безопасность: современная теория и практика: сб. науч. тр. студентов, аспирантов и преподавателей по материалам III Межвуз. науч.-практ. конф. (Омск, 24 нояб. 2020 г.) / Сиб. гос. автомобильно-дорож. ун-т (СибАДИ). – Омск: Изд-во СибАДИ, 2020. – С. 112–114.
62. AIConstructor – облачная среда разработки искусственного интеллекта для цифровой трансформации предприятий без написания кода / А.Е. Сулавко, Д.П. Иниватов, А.В. Еременко, Е.В. Шалина // Информационная безопасность: современная теория и практика: сб. науч. тр. студентов, аспирантов и преподавателей по материалам III Межвуз. науч.-практ. конф. (Омск, 24 нояб. 2020 г.) / Сиб. гос. автомобильно-дорож. ун-т (СибАДИ). – Омск: Изд-во СибАДИ, 2020. – С. 115–120.
63. Сулавко, А.Е. Разностные нейроны Байеса с множеством квантователей для высоконадежной аутентификации и защищенного исполнения искусственного интеллекта / А.Е. Сулавко // Безопасность информационных технологий: сб. науч. ст. по материалам II Всерос. науч.-техн. конф. – Пенза: Изд-во ПГУ, 2020. – С. 103–111.
64. Сулавко, А.Е. Модель защищенного нейро-иммунного контейнера для задач биометрической аутентификации / А.Е. Сулавко, А.А. Лыжин // Фундаментальные и прикладные исследования молодых учёных: сб. материалов IV Междунар. науч.-практ. конф. студентов, аспирантов и молодых учёных (Омск, 6–7 февр. 2020 г.). – Омск: Изд-во СибАДИ, 2020. – С. 378–382.
65. Сулавко, А.Е. Идентификация образов электроэнцефалограмм пользователей компьютерных систем при наборе парольных фраз на клавиатуре / А.Е Сулавко, С.С. Жумажанова, Д.Г. Стадников // Искусственный интеллект и принятие решений. – 2019. – № 2. – С. 15–27.
66. Идентификация личности по особенностям лица с использованием искусственной иммунной системы и формулы гипотез Байеса / А.Е. Сулавко, Е.В. Шалина // Интеллектуальный анализ сигналов, данных и знаний: методы и средства: сб. ст. II Всерос. науч.-практ. конф. с междунар. участием им. В. В. Губарева (Новосибирск, 11–13 дек. 2018 г.) / Новосиб. гос. техн. ун-т. - Новосибирск: Изд-во НГТУ, 2018. – С. 303–307.
67. Сулавко, А.Е. Биометрическая аутентификация по клавиатурному почерку на основе иммунного алгоритма распознавания образов / А.Е. Сулавко, Е.В. Шалина, Д.Г. Стадников // Интеллектуальный анализ сигналов, данных и знаний: методы и средства: сб. ст. II Всерос. науч.-практ. конф. с междунар. участием им. В. В. Губарева (Новосибирск, 11–13 дек. 2018 г.) / Новосиб. гос. техн. ун-т. - Новосибирск: Изд-во НГТУ, 2018. – С. 307–315.

68. Сулавко А.Е., Жумажанова С.С., Стадников Д.Г., Пасенчук В.А., Приз И.Л., Нигрей А.А. Идентификация человека с высокой точностью по особенностям работы головного мозга на основе визуальной стимуляции // Биомедицинская радиоэлектроника. – 2018. – №12. – С. 22-35. DOI:10.18127/j15604136-201812-03
69. Sulavko, A.E. Comparison of functionals based on statistic tests for generating fast learning wide neural networks / A.E. Sulavko // Инфографика и информационный дизайн: визуализация данных в науке: материалы Междунар. науч.-практ. конф. (Омск, 17–18 нояб. 2017 г.) / ОмГТУ. – Омск: Изд-во ОмГТУ, 2017. – С. 210–223.
70. Об оценке возможностей человека по распознаванию рукописных образов в процессе их воспроизведения на экране монитора / В.И. Васильев, А.Е. Сулавко, С.С. Жумажанова, А.А. Нигрей // Омский научный вестник. – 2017. – № 5. – С. 175–180.

Рецензируемая монография:

71. Идентификационный потенциал пользователей компьютерных систем в процессе их профессиональной деятельности: монография / Б.Н. Епифанцев, А.Е. Сулавко, А.С. Ковальчук [и др.]. – Омск: СибАДИ, 2017. – 1 DVD-R. ISBN: 978-5-00113-046-8

Патенты и свидетельства о регистрации программ для ЭВМ:

72. Свидетельство о государственной регистрации программы для ЭВМ № 2021660512 Российская Федерация. AIC desktop: № 2021617236: заявл. 17.05.2021: опубл. (зарег.) 28.06.2021 / А.Е. Сулавко, Д.Г. Стадников, А.Г. Чобан, Д.П. Иниватов; заявитель Ом. гос. техн. ун-т. – 1 с.
73. Свидетельство о государственной регистрации программы для ЭВМ № 2019663412 Российская Федерация. Программный модуль для цифрового подписания PDF-документов «PdfDigiSign»: № 2019662174: заявл. 07.10.2019: опубл. (зарег.) 16.10.2019 / П.С. Ложников, М.А. Семиколенов, А.Е. Сулавко; заявитель Ом. гос. техн. ун-т. – 1 с.
74. Свидетельство о регистрации электронного ресурса № 23578 от 26.04.2018. Разрез файлов формата WAV / Д.П. Иниватов, А.Е. Сулавко; Ом. гос. техн. ун-т. – Москва: ОФЭРНиО. – 1 с.
75. Свидетельство о государственной регистрации программы для ЭВМ № 2017616888 Российская Федерация. Среда для имитационного моделирования экспериментов и проверки гипотез по распознаванию образов «SHV-kernel»: № 2017614035: заявл. 24.04.2017: опубл. (зарег.) 19.06.2017 / А.Е. Сулавко; заявитель Ом. гос. техн. ун-т. – 1 с.
76. Патент № 2543927 Российская Федерация, МПК G06K 9/00. Способ идентификации личности по особенностям динамики написания пароля: № 2014116281/08; заявл. 22.04.2014; опубл. 10.03.2015 / Б.Н. Епифанцев, П.С. Ложников, А.Е. Самотуга, А.Е. Сулавко.
77. Свидетельство о государственной регистрации программы для ЭВМ № 2012612981 Российская Федерация. Распределенная система управления доступом к ресурсам компьютера на основе регистрируемых событий: №: заявл.: опубл. (зарег.) 15.06.2012 / А.Е. Сулавко, А.Л. Богдан; заявитель Ом. гос. техн. ун-т. – 1 с.
78. Свидетельство о государственной регистрации программы для ЭВМ № 2011619263 Российская Федерация. Мультифакторная система аутентификации «TEOFRAST-M»: № 2011619263: заявл. 05.12.2011: опубл. (зарег.) 01.02.2012 / П.С. Ложников, В.А. Перевальский, А.Е. Сулавко; заявитель Ом. гос. техн. ун-т. – 1 с.
79. Свидетельство о государственной регистрации программы для ЭВМ № 2011611363 Российская Федерация. Система безопасности компьютера на основе регистрируемых событий в компьютерных сетях: № заявл.: опубл. (зарег.) 28.04.2011/ А.Е. Сулавко, С.А. Голованов; заявитель Ом. гос. техн. ун-т. – 1 с.
80. Свидетельство о государственной регистрации программы для ЭВМ № 2010610473 Российская Федерация. Программный модуль для обеспечения безопасности бухгалтерских информационных систем «TEOFRAST-B»: № 2009616252: заявл.10.11.2010: опубл. (зарег.)11.01.2010 / П.С.Ложников, А.В. Еременко, В.А. Перевальский, А.Е.Сулавко; заявитель Ом. гос.техн.ун-т.– 1 с.

Диссертант Сулавко А.Е.