На правах рукописи

Панфилова Ирина Евгеньевна

МОДЕЛИ И АЛГОРИТМЫ НЕЙРОСЕТЕВОЙ БИОМЕТРИЧЕСКОЙ АУТЕНТИФИКАЦИИ В ЗАЩИЩЕННОМ РЕЖИМЕ ИСПОЛНЕНИЯ

Специальность 2.3.6. Методы и системы защиты информации, информационная безопасность

АВТОРЕФЕРАТ

диссертации на соискание ученой степени кандидата технических наук

Самара – 2024

Работа выполнена в федеральном государственном бюджетном образовательном учреждении высшего образования «Самарский государственный технический университет».

Научный руководитель: доктор технических наук, доцент, проректор по научной и инновационной деятельности федерального государственного автономного образовательного учреждения высшего образования «Омский государственный технический университет»,
г. Омск
Ложников Павел Сергеевич

Официальные оппоненты:
Котенко Игорь Витальевич, доктор технических наук, профессор, главный научный сотрудник лаборатории проблем компьютерной безопасности федерального государственного бюджетного учреждения науки «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук»
Исмагилова Альбина Сабирьяновна, доктор физико-математических наук, доцент, заведующий кафедрой управления информационной безопасностью федерального государственного бюджетного образовательного учреждения высшего образования «Уфимский университет науки и технологий»

Ведущая организация: федеральное государственное бюджетное образовательное учреждение высшего образования «Пензенский государственный университет», г. Пенза

Защита диссертации состоится 13.12.2024 г. в 10 ч. 00 мин. на заседании диссертационного совета 24.2.479.07 на базе федерального государственного бюджетного образовательного учреждения высшего образования «Уфимский университет науки и технологий», по адресу: 450008, г. Уфа, ул. К. Маркса, д. 12.

С диссертацией можно ознакомиться в библиотеке федерального государственного бюджетного образовательного учреждения высшего образования «Уфимский университет науки и технологий» и на сайте https://uust.ru/

Автореферат разослан «___»_____2024 года.

Ученый секретарь
диссертационного совета,
д.т.н., профессор Вульфин Алексей Михайлович

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Активное развитие искусственного интеллекта (ИИ) обуславливает появление принципиально новых исследовательских задач, направленных на обеспечение его безопасности. Так, по данным MarketsandMarkets, глобальный рынок кибербезопасности для ИИ достигнет 38,2 млрд. долларов к 2026 году, при среднегодовом темпе роста в 26,3% (при 9,3 млрд. долларов в 2020 г.). Теоретические аспекты обеспечения безопасности ИИ отражены в концепции доверенного искусственного интеллекта (ДИИ), предполагающей высокий уровень надёжности, прозрачности и конфиденциальности технологий ИИ. Перечисленные требования особенно актуальны для систем и алгоритмов, принимающих ответственные решения или оперирующих критически важными данными. Таковым в полной мере может считаться искусственный интеллект, лежащий в основе биометрических систем аутентификации. Варианты защищённого исполнения таких систем (при которых невозможны раскрытие логики работы ИИ, извлечение его знаний и управление ими) представлены в рамках исследований *высоконадежной биометрической аутентификации*, в отечественной практике отраженных в серии стандартов ГОСТ Р 52633. Базовыми характеристиками высоконадежной нейросетевой биометрической аутентификации можно считать:

− обеспечение сокрытия решающих правил и конфиденциальности знаний;
− защиту биометрических данных от утечки;
− обеспечение устойчивости к внешним воздействиям (атакам).

Перечисленным критериям, на сегодняшний день, в большей степени соответствуют нейросетевые преобразователи «биометрия-код» (НПБК) − разновидность биометрических криптосистем (БКС), основанных на принципах работы искусственных нейронных сетей. Основной целью НПБК является связывание биометрического образа человека с криптографическим ключом. Чем длиннее криптографический ключ, продуцируемый НПБК, тем ниже возможность компрометации исходного биометрического образа. Разнообразие модификаций нейросетевых преобразователей, представленное как зарубежной, так и отечественной литературой, наглядно демонстрируют актуальность разработки процедур биометрической аутентификации в защищенном режиме исполнения – *защищенной биометрической аутентификации (ЗБА)*, одновременно с этим, демонстрируя наличие ряда нерешенных задач. Одной из таких задач является необходимость поддержания высокого уровня защищенности процедуры аутентификации от деструктивных воздействий при заданном уровне точности распознавания биометрических образов (аутентификации). Под деструктивными воздействиями понимаются:

− атаки извлечения знаний НПБК;
− компрометация открытых биометрических образов;
− атаки на биометрическое предъявление (спуфинг атаки), направленные на получение несанкционированного доступа к объектам защиты системы ЗБА.

Особенности реализации указанной задачи зависят от конкретной биометрической модальности: отпечатки пальцев, рукописный почерк, голос, лицо и т.д. В этой связи, работа с биометрическими образами, обладающими особенностями сбора и представления, повышает требования к структуре и функциональным возможностям нейросетевых преобразователей «биометрия-код», лежащих в основе ЗБА. Среди таких биометрических модальностей особенно выделяется лицо человека: системы защищенной биометрической аутентификации по лицу в полной мере подвержены всем перечисленным выше деструктивным воздействиям.

С учетом перечисленных особенностей, можно сформулировать общую **научную задачу**, заключающуюся в необходимости изменения логики функционирования и концептуального исполнения НПБК с целью повышения его защищенности по отношению к деструктивным воздействиям при работе с биометрическими образами лиц. Проведенные в ходе работы исследования показали, что предложенная система защищенной биометрической аутентификации является полноценным решением поставленной научной задачи.

Степень проработки темы исследования. Вопросам защищенной биометрической аутентификации посвящены работы таких отечественных и зарубежных исследователей, как: Ахметов Б. С., Безяев А.В., Васильев В.И., Волчихин В.И., Иванов А.И., Малыгина Е.А., Сулавко А.Е., Derakhshani R., Dong X., Liu W., Rathgeb C., Rattani A., Talreja V. и др. Анализ современного состояния области демонстрирует значительные преимущества нейросетевых преобразователей «биометрия-код» (НПБК) перед альтернативными решениями на основе глубоких нейронных сетей и/или нечетких экстракторов. Однако существующие реализации НПБК обладают значительными недостатками либо с точки зрения длины продуцируемого ключа (только 128 бит для классического НПБК), либо с точки зрения точности работы с лицом человека (корреляционные нейроны работают только с сильно коррелированными признаками, что не характерно для образов лица). Кроме того, ЗБА на основе существующих реализаций НПБК оказывается уязвимой по отношению к спуфинг атакам, что способно нивелировать преимущества ее защищенного исполнения.

Перечисленные недостатки демонстрируют необходимость изменения не только логики работы НПБК, но и концептуального исполнения защищенной биометрической аутентификации с его участием. Результатом вносимых изменений должна стать система защищенной биометрической аутентификации по лицу на основе нейросетевого преобразователя «биометрия-код», устойчивая к атакам извлечения знаний НПБК и компрометации открытых биометрических образов лиц, а также к атакам на биометрическое предъявление (спуфинг атак).

Цель диссертационной работы: повысить защищенность процедуры биометрической аутентификации личности на основе нейросетевого преобразователя «биометрия-код», использующего открытые биометрические образы лица человека.

Для достижения поставленной цели необходимо решить **следующие задачи:**

1. Разработать концепцию защищенной биометрической аутентификации по лицу на основе НПБК, устойчивой к внешним воздействиям в виде атак на биометрическое предъявление (спуфинг атак).
2. Разработать модель нейрона и основанную на ней модель нейросетевого преобразователя «биометрия-код», осуществляющих процедуру биометрической аутентификации по лицу с обеспечением защиты знаний и биометрических образов лиц от компрометации.
3. Разработать алгоритмы обучения нейросетевого преобразователя биометрических образов лица в код на малых выборках.
4. Разработать систему биометрической аутентификации по лицу, устойчивую к атакам извлечения знаний НПБК и компрометации открытых биометрических образов лиц, а также к атакам на биометрическое предъявление (спуфинг атак).

Объектом исследования являются биометрические системы аутентификации человека на основе нейросетевых алгоритмов.

Предметом исследования являются нейросетевые модели преобразователей биометрических образов лица в сильный пароль или криптографический ключ.

Методы исследования. Применялись методы классификации и идентификации образов, биометрической аутентификации, глубокого обучения, теории вероятностей и математической статистики, распознавания образов, компьютерного моделирования, кодирования информации, аппарат искусственных нейронных сетей (ИНС), тригонометрические вычисления.

Достоверность и обоснованность работы подтверждается корректной постановкой задач и выбором известных методов, успешно применяемых в других областях, практическим применением системы, построенной в соответствии с разработанными моделями и алгоритмами, а также апробацией на научных конференциях, публикацией результатов в научных изданиях, в том числе из Перечня ВАК, актами о внедрении результатов работы в образовательную и производственную сферы.

Научная новизна состоит из предложенных в работе:

1. Концепции защищенной биометрической аутентификации по лицу, *отличающейся* применением механизма защищенного нейросетевого контейнера (ЗНК) для безопасного взаимодействия блока аутентификации на основе пользовательского НПБК и блока обнаружения спуфинг атак на основе НПБК, представленного в виде классификатора реальных и поддельных изображений лиц, что позволяет обеспечивать устойчивость процедуры аутентификации к атакам на биометрическое предъявление (спуфинг атакам), а также дополнительную защиту таблиц нейросетевых функционалов пользовательского НПБК.
2. Модели тригонометрического нейрона, а также основанной на ней модели НПБК, *отличающихся* применением новой тригонометрической меры оценки расстояния между образами субъектов в подпространстве пар признаков вместо исходных признаков, что обеспечивает защиту образов лиц от компрометации путем продуцирования длинного криптографического ключа при высокой точности классификации. Предложенные модели не используют параметры распределений и/или характеристики образов легитимных пользователей, что обеспечивает защиту знаний НПБК от компрометации.
3. Алгоритма калибровки нейросетевых преобразователей «биометрия-код» и алгоритма автоматического обучения НПБК на основе тригонометрических нейронов, *отличающихся* использованием дополнительной информации, полученной путем оценки не участвующего в обучении набора биометрических образов лиц, что дает возможность быстрого и робастного обучения пользовательских НПБК на малых выборках образов лиц.
4. Структуры системы защищенной биометрической аутентификации по лицу, *отличающейся* наличием независимых блоков извлечения признаков, обучения нейросетевых преобразователей и аутентификации, а также применением варианта исполнения ЗНК, при котором в режиме обучения ключом НПБК для обнаружения спуфинг атак осуществляется защита структуры пользовательского НПБК, а обратный описанному процесс происходит при аутентификации. Разработанная структура позволяет повысить защищенность процедуры биометрической аутентификации личности по лицу на основе НПБК в отношении спуфинг атак, а также атак извлечения знаний НПБК и компрометации биометрических образов лиц.

Теоретическая значимость диссертационной работы заключается в новом математическом аппарате для построения защищенной биометрической аутентификации, работающей со слабо коррелированными признаками лица человека. Предложенная математическая модель нейросетевого преобразователя «биометрия-код» повышает устойчивость биометрической аутентификации по лицу к деструктивным воздействиям в виде атак извлечения знаний и компрометации биометрических образов путем продуцирования длинного криптографического ключа (2048 бит), значительно превышающего длину ключа НПБК, обученного в соответствии с ГОСТ Р 52633.5 (128 бит), а также дополняют функционал нейросетевых преобразователей на основе корреляционных нейронов и позволяют работать со слабо колерованными признаками лица. Особенности работы нейронов НПБК позволяют использовать малое число примеров биометрических образов для обучения преобразователя и осуществлять процедуру обучения автоматически, не раскрывая параметров легитимных пользователей, что может быть актуально для иных приложений ИИ, исполняемых в защищенном режиме.

Практическая значимость работы заключается в разработке системы защищенной биометрической аутентификации по лицу и ее программной реализации. Система основана на предложенных в работе концепции, моделях, алгоритмах и структуре. Коэффициент равной вероятности ошибок при работе системы составил EER=2,5%, что говорит о сравнительно низком уровне ошибок распознавания образов при высоком уровне защищенности процедуры биометрической аутентификации по лицу от атак на биометрическое предъявление, а также атак компрометации знаний НПБК и биометрических данных.

Положения, выносимые на защиту:

1. Концепция защищенной биометрической аутентификации по лицу, обеспечивающей противодействие атакам на биометрическое предъявление.
2. Модель тригонометрического нейрона и основанная на ней модель нейросетевого преобразователя «биометрия-код», осуществляющие процедуру защищенной биометрической аутентификации по лицу.
3. Алгоритм калибровки нейросетевых преобразователей «биометрия-код» и алгоритм автоматического обучения НПБК на основе тригонометрических нейронов, позволяющие производить быстрое и робастное обучение НПБК на малых выборках образов лиц.
4. Система защищенной биометрической аутентификации по лицу и ее программная реализация, обеспечивающая защищенность процедуры биометрической аутентификации личности по лицу на основе НПБК в отношении спуфинг атак и атак компрометации знаний НПБК и открытых биометрических образов лиц.

Личный вклад. Все результаты, изложенные в диссертации, включая программные реализации предложенных в работе алгоритмов, получены автором самостоятельно. Проработка цели и задач, способов их решения и вариантов представления результатов осуществлены автором совместно с научным руководителем.

Апробация результатов работы. Полученные результаты работы докладывались на следующих конференциях: Всероссийская молодежная научно-практическая конференция «Нанотехнологии. Информация. Радиотехника» (г. Омск); V Всероссийская научно-техническая конференция «Безопасность информационных технологий» (г. Омск)**;** IEEE Conference on the Intelligent Methods, Systems, and Applications (Giza, Egypt); III Всероссийская научная школа-семинар «Современные тенденции развития методов и технологий защиты информации» (г. Москва).

Работа выполнена в рамках государственного задания Минобрнауки России на 2023-2025 годы № FSGF-2023-0004. Часть работы по теме диссертации проводилась в рамках гранта ИБ МТУСИ № 40469-18/23-К. Грант выполнялся автором единолично.

Соответствие паспорту специальности. Тема и содержание диссертации соответствуют паспорту специальности 2.3.6. Методы и системы защиты информации, информационная безопасность, пункту 12: «Технологии идентификации и аутентификации пользователей и субъектов информационных процессов. Системы разграничения доступа», а также пункту 15: «Принципы и решения (технические, математические, организационные и др.) по созданию новых и совершенствованию существующих средств защиты информации и обеспечения информационной безопасности».

Публикации. По теме диссертации лично и в соавторстве опубликовано 10 печатных работ, 6 из которых изданы в журналах, рекомендованных ВАК; 1 научная публикация индексирована в международной информационно-аналитической системе научного цитирования Scopus. Получено 2 свидетельства о государственной регистрации программы для ЭВМ.

Объем и структура работы. Диссертация состоит из введения, 4 глав, заключения, списка литературы (195 наименований) и 2 приложений. Общий объем диссертации составляет 166 страниц, включающих в себя 15 таблиц и 28 рисунков.

ОСНОВНЫЕ ПОЛОЖЕНИЯ РАБОТЫ

Во введении представлены актуальность темы диссертации, сформулированы цель, задачи, объект и предмет исследования, а также научная новизна, теоретическая и практическая значимость полученных в ходе работы результатов.

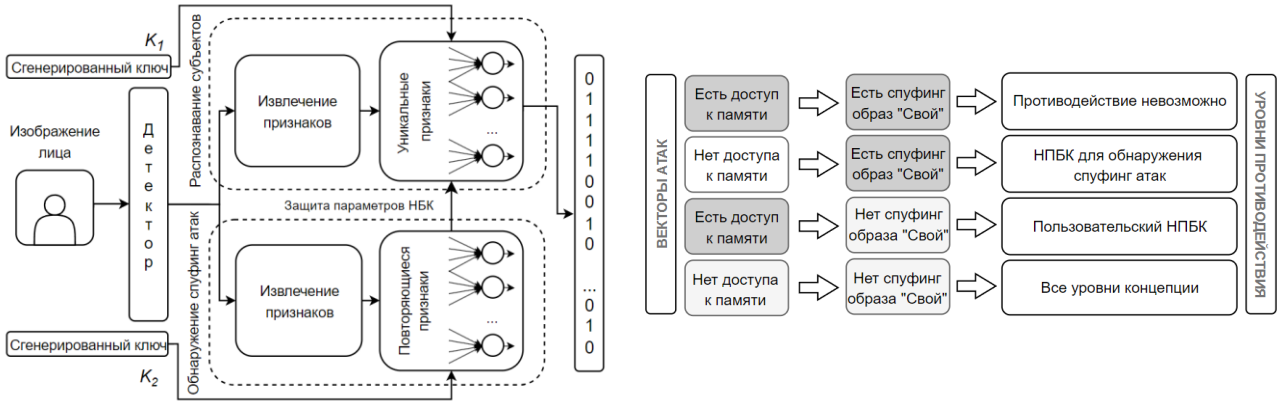
Первая глава посвящена анализу современного состояния исследований в области защищенного исполнения искусственного интеллекта в задачах биометрической аутентификации субъектов. Рассмотрены основные проблемы функциональной безопасности искусственного интеллекта, в том числе оперирующего биометрическими данными. Отдельно определена проблема биометрической аутентификации на основе ИИ,

заключающаяся в уязвимости к атакам на биометрическое предъявление (спуфинг атакам). В результате анализа спуфинг атак и атак, ориентированных на компрометацию знаний ИИ и биометрических образов, подчеркнута актуальность разработки систем защищенной биометрической аутентификации.

Защищенная биометрическая аутентификация строится на методах защиты биометрических шаблонов (ЗБШ), среди которых выделены 3 направления: биокриптографические системы (БКС), отменяемая биометрия и гомоморфное шифрование. На фоне принципиальных недостатков отменяемой биометрии и гомоморфного шифрования, отмечены преимущества БКС, обеспечивающих надежное связывание биометрического образа с криптографическим ключом, заменяющим длинные и ненадежные пароли. В качестве ключевых направлений развития БКС определены нечеткие экстракторы и нейросетевые преобразователи «биометрия-код» (НПБК). Концепция НПБК относится к высоконадежной биометрической аутентификации, критерии которой определены в серии стандартов ГОСТ Р 52633. Анализ обозначенных вариантов исполнения БКС продемонстрировал однозначное преимущество нейросетевых преобразователей перед нечеткими экстракторами, а также позволил выделить ряд недостатков, свойственных существующим реализациям НПБК: классическому НПБК и НПБК на базе корреляционных нейронов. Отмечена необходимость нивелирования рассмотренных недостатков в отношении лицевой биометрии.

В заключении главы сформулированы цели и задачи исследования в соответствии с поставленной научной задачей.

Во второй главе предложена концепция защищенной биометрической аутентификации по лицу на основе нейросетевого преобразователя «биометрия-код», устойчивой к атакам на биометрическое предъявление (рис. 1 (а)). Предположенная концепция не только устойчива к спуфинг атакам, но также обеспечивает защиту параметров пользовательского НПБК с помощью ключа, продуцируемого НПБК для обнаружения атак на биометрическое предъявление.



а) б)

Рисунок 1. Концепция защищенной биометрической аутентификации по лицу, устойчивой к атакам на биометрическое предъявление: а) структурная схема предложенной концепции; б) векторы атак в отношении предложенной концепции

Так как процедуры обнаружения спуфинг атак и распознавания субъектов (аутентификации) не могут быть исполнены в рамках одного алгоритма, предложенная концепция подразумевает наличие двух параллельно функционирующих блоков. Нейросетевой преобразователь, отвечающий за распознавание, должен обеспечивать связывание биометрического образа лица человека с длинным криптографическим ключом K_1, отвечающим требованиям высоконадежной биометрической аутентификации. В свою очередь, блок обнаружения спуфинг атак предлагается исполнить в режиме, при котором будет минимизировано влияние ошибок классификации спуфинг атак на результирующую процедуру аутентификации. Организация взаимодействия двух функциональных блоков концепции осуществляется с учетом четырех возможных векторов атак (рис. 1 (б)). Наиболее

негативный сценарий, при котором злоумышленник имеет доступ к памяти и обладает спуфинг-образом «Свой», является единственным, при котором невозможно противодействие ни на одном из уровней концепции.

Нейросетевые параметры НПБК, осуществляющего аутентификацию, хранятся в *нейросетевых биометрических контейнерах* (НБК) и представляют собой таблицы связей и/или весов НПБК. Предлагается защищать нейросетевые контейнеры путем размножения ошибок образа «Чужой» с применением обратимых и необратимых преобразований. Для защиты таблиц нейросетевых функционалов применяется *механизм защищенного нейросетевого контейнера* (ЗНК): после обучения таблицы каждого нейрона шифруются наложением гаммы, представляющей ключ НПБК для обнаружения спуфинг-атак. В режиме ЗНК энтропия выходов пользовательского НПБК при поступлении образа «Чужой» или спуфинг-образа повышается. Это приводит к тому, что нейроны с неверно восстановленными таблицами связей дают почти случайный отклик. Таким образом, ЗНК препятствует осуществлению направленного перебора образов «Чужой» для несанкционированного восстановления ключа (или его отдельных бит).

Глубокие нейросетевые архитектуры, решающие задачу распознавания атак на биометрическое предъявление, обладают слабой обобщающей способностью по отношению к разнообразию реализации таких угроз и часто «переучиваются» на специально подобранных наборах данных. Нивелировать указанные недостатки можно путем разделения блока векторного представления образов и блока принятия решения. В рамках экспериментальной реализации концепции в качестве блока векторного представления была обучена глубокая нейронная сеть, основанная на архитектуре FeatherNet. Обучение осуществлялось на открытом наборе данных CelebA-Spoof. На тестовой выборке максимальное значение точности составило 93%. Для дальнейшего использования обученной сети в качестве экстрактора признаков (от англ. extraction - извлечение), слои классификатора «замораживаются», а работа осуществляется с 96-мерным вектором признаков на выходе предшествующего классификатору слоя сети.

В качестве классификатора реальных и поддельных входных образов лиц реализован нейросетевой преобразователь биометрия-код, процедура обучения которого представлена в ГОСТ Р 52633.5-2011. Адаптация логики работы НПБК под задачу бинарной классификации осуществляется путем построения преобразователя для реальных изображений лиц (класс C_1). В таком случае, поддельные изображения (класс C_2) расцениваются НПБК как «Чужие», а случайный код на выходе свидетельствует о решении в пользу класса C_2. Принадлежность классу C_i, i = 1, 2, оценивается исходя из получаемого на выходе НПБК бинарного кода (ключ K_2 в случае класса C_1). Особенностью построения НПБК в качестве классификатора является допущение о возможности дублировании входов нейрона в связи с отсутствием необходимости сокрытия структуры преобразователя. Для оценки качества осуществляемой бинарной классификации i-ого входного образа применяется следующее правило:

$$\begin{cases} (\bar{a}_i \in C_1 \wedge h_i < threshold) \rightarrow TP \\ (\bar{a}_i \in C_1 \wedge h_i > threshold) \rightarrow FN \\ (\bar{a}_i \in C_2 \wedge h_i < threshold) \rightarrow FP \\ (\bar{a}_i \in C_2 \wedge h_i > threshold) \rightarrow TN \end{cases}$$

где h_i – i-ое значение расстояния Хэмминга между ожидаемым кодом и выходом НПБК, $threshold$ – порог, определяющий допустимое количество ошибок в коде i-ого входного образа для корректного отнесения его к одной из групп классифицированных образов: TP – True Positive, FN – False Negative, FP – False Positive или TN – True Negative. Полученный классификатор не требует итерационного обучения (осуществляется автоматически) и большого числа обучающих примеров. На основе полученных значений указанных метрик (TP, FN, FP, TN) высчитывается точность классификации. Лучший результат на тестовых выборках составил 97,2% точности (ACER=2,9%) при значении $threshold$ = 3. Несмотря на сравнительно невысокие показатели (табл. 1) точности работы НПБК в качестве классификатора, сохраняется два ключевых преимущества предложенного решения:

безопасная реализация концепции защищенной биометрической аутентификации по лицу, при которой модуль обнаружения спуфинг атак не становится «слабой» точкой потенциальных уязвимостей, и возможность быстрого (автоматического) переобучения анти-спуфинг модуля на малых выборках новых данных.

Таблица 1. – Сравнительные результаты работы предложенного решения с альтернативными моделями глубокого обучения для обнаружения спуфинг атак

№	Модель	Датасет	Accuracy	ACER
1	DeepPixBiS	OULU-NPU (p.1)	-	5.97%
2	CDCN++	OULU-NPU (p.1)	-	1.3%
3	AENet	CelebA-Spoof Dataset	99,6%	3.09%
4	MCCNN (BCE+OCCL)-GMM	SiW-M	-	14.9 ± 7.8%
5	FasTCo	SiW-M	-	10.1 ± 5.6%
6	MobileNetv3	CelebA-Spoof Dataset	99,8%	3.8%
7	**FeatherNet + НПБК**	**CelebA-Spoof Dataset**	**97,2%**	**2,9%**

В третьей главе рассматривается процедура построения и обучения нейросетевого преобразователя «биометрия-код», обеспечивающего процедуру биометрической аутентификации по лицу с обеспечением защиты от компрометации биометрических данных субъектов и знаний НПБК. Особенность построения НПБК заключается в применении нового математического аппарата построения нейронов на основе тригонометрической меры оценки расстояния между образами «Своего» и образами «Чужих» в подпространствах пар признаков.

В задачах биометрической идентификации и верификации, поступающий на вход системы вектор признаков $\bar{a} = (a_1, a_2, \dots, a_n)$, описывающий биометрический образ субъекта, сравнивается с некоторым эталонным вектором $\bar{a}'$, однозначно характеризующим этого пользователя. Однако анализируя взаимосвязи между признаками, в том числе, корреляционные связи, можно извлечь дополнительную информацию о различии или сходстве сравниваемых образов. В таком случае, любая пара признаков $(a_i,\ a_j)$ может быть изучена на предмет функциональной зависимости. Пространство, образуемое в ходе проведения такого анализа, будем называть *подпространством пар признаков*. Количество возможных подпространств пар признаков определяется по формуле:

$$n' = 0.5(n(n-1))$$

где n – исходное количество признаков. Так как сами по себе подпространства пар признаков могут нести только информацию о корреляционной зависимости, то для получения дополнительной информации необходимо совершить его отображение в некоторое новое пространство. Такое отображение можно описать следующим образом:

$$a'_l = f(a_i, a_j)$$

где i и j – номера признаков исходного вектора признаков ($i \neq j$), a'_l – *мета-признак*, т.е. признак, полученный путем синтеза двух или более исходных признаков с помощью функционального преобразования f, l – номер мета-признака. Полученное пространство, образуемое мета-признаками, есть *пространство мета-признаков*. Описанное функциональное преобразование должно отвечать, как минимум, двум критериям:

1. Не содержать в себе характеристик, компрометирующих образ легитимного пользователя («Своего»). Этот критерий является основополагающим для построения защищенной системы биометрической аутентификации.
2. Позволять с достаточно высокой точностью описывать расположение образа в пространстве мета-признаков с учетом высокой вариативности биометрических образов.

В качестве функционального преобразования, отвечающего указанным критериям, в работе предложено выражение (1):

$$a'_l = \sqrt{(a_i - m_i)^2 + (a_j - m_j)^2} * sin(\widehat{\bar{d}, \bar{v}}), \quad (1)$$

где m_i – среднее арифметическое значений i-ого признака, m_j – среднее арифметическое значений j-ого признака, $\bar{d}$ – вектор, характеризующий расстояние от точки $O = (m_i,\ m_j)$

(координаты центра собственной области класса «Чужие») до точки $M = (a_i,\ a_j)$, характеризующей биометрической образ, $\bar{v}$ – вектор, равный вектору $\bar{d}$, но расположенный в начале тригонометрической окружности, $sin(\widehat{\bar{d}, \bar{v}})$ – синус угла между векторами $\bar{d}$ и $\bar{v}$.

Мера (1) основана на так называемом тригонометрическом нивелировании расстояния. Несмотря на то, что метрика не является полным аналогом тригонометрического нивелирования, она оказывает корректирующее воздействие на исходное расстояние и дает более информативное представление о расположении образов относительно друг друга. Альтернативой полученной метрики может выступать преобразование, позволяющее избегать дополнительных вычислительных процедур. Такая метрика характеризуется исключительно синусом угла между векторами $\bar{d}$ и $\bar{v}$:

$$a'_l = sin(\widehat{\bar{d}, \bar{v}}) \quad (2)$$

Простейший тригонометрический нейрон строится на метрике (3) и принимает на вход мета-признаки, построенные с помощью отображений, учитывающих любой вариант тригонометрического нивелирования. Представленный нейрон суммирует значения мета-признаков, однако возможны и более сложные схемы интегрирования входных значений, а также иные варианты тригонометрического нивелирования расстояний.

$$y = \sum_{l=1}^{k} a'_l, \quad (3)$$

где k – количество синапсов (входов) нейрона. Будем называть такой нейрон *тригонометрическим*. Принятие решения по сумме входных значений мета-признаков нейрон осуществляет согласно двухуровневой пороговой функции активации $\varphi(y)$:

$$\varphi(y) = \begin{cases} 1,\ y \geq T_2 \\ 0,\ T_1 < y < T_2 \\ -1,\ y \leq T_1 \end{cases} \quad (4)$$

где T_1 и T_2 – пороги принятия решения в пользу одного из трех значений функции. Пороговая функция необходима для квантования результатов преобразований, так как нейроны должны генерировать на выходах бинарный код.

Чтобы НПБК осуществлял связывание ключа пользователя с биометрическим образом, значения на выходе функции активации необходимо преобразовать в двоичные состояния типа {«10», «00», «01»}. Тогда каждый нейрон будет продуцировать 2 бита информации. Для осуществления данной процедуры нужно выбрать номер хеш-преобразования для нейрона из таблицы 24 вариантов хеширующих преобразований отклика нейрона в двоичный код.

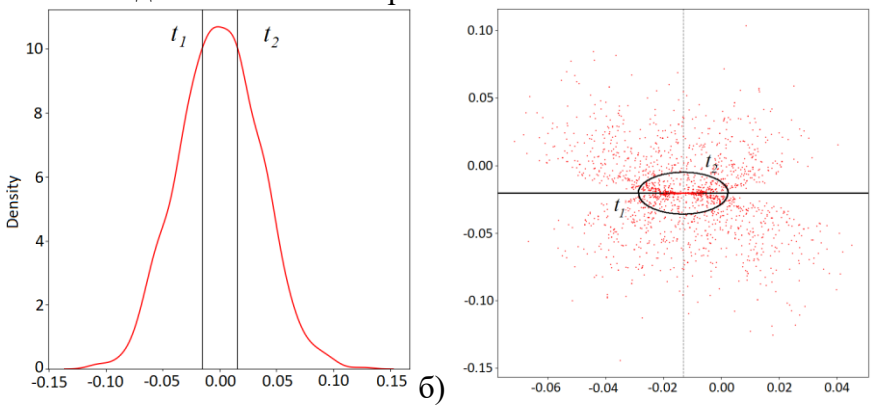
Тригонометрический нейрон является частично связным. Для его настройки требуется подобрать несколько пар признаков и определить для каждой пары пороги t_1 и t_2. Пороги должны разделять соответствующие подпространства на три сектора ([-∞; t_1], (t_1; t_2), [t_2; ∞]) таким образом, чтобы в каждый сектор попадало примерно равное количество обучающих примеров «Чужие». Пары признаков следует выбирать так, чтобы все обучающие примеры «Свой» для каждой пары попадали в один определенный сектор (рис. 2).

Все пороги тригонометрического нейрона зависят только от обучающей выборки образов «Чужие», не являющихся секретными. Единственным секретным элементом нейрона является номер сектора, который был выбран при обучении и который связан с двумя битами ключа пользователя. Злоумышленник не знает, на какие биты настроен нейрон, и не знает, какой сектор является верным, что определяет безопасность биометрического шаблона.

Для корректной работы НПБК необходимо правильно настроить пороги для каждого подпространства пар признаков. Проще всего разделение на секторы можно продемонстрировать на графике плотности вероятности мета-признаков (рис. 2а). Для определения двух порогов и обхода необходимости построения эмпирической функции плотности вероятности (в случае, если закон распределения мета-признаков неизвестен или не соответствует нормальному распределению), предлагается рассчитывать пороги согласно следующему алгоритму (рис. 3):

1. Подпространство пары признаков переводится в пространство мета-признаков.
2. Полученные значения мета-признаков ранжируются по возрастанию.

3. Полученный массив делится на 3 сектора.



а) б)

Рисунок 2. Визуализация порогов принятия решения тригонометрическим нейроном на основе метрики (1): а) пороги, расположенные на графике оценки плотности вероятности распределения мета-признаков, б) пороги, расположенные в пространстве мета-признаков

Для каждой пары признаков значение порогов сохраняются и используются при синтезе и обучении НПБК. После завершения работы алгоритма, нет необходимости хранить данные «Чужих», использовавшиеся в ходе его работы, если пороги не нуждаются в уточнении (донастройке).

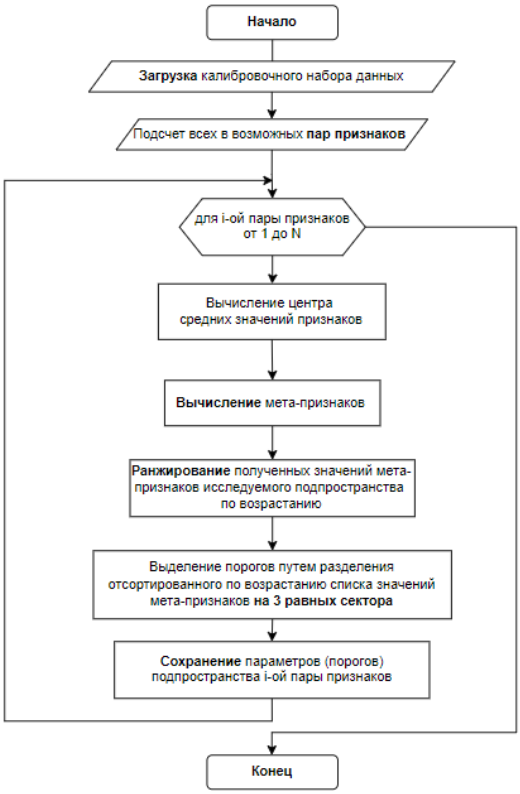


Рисунок 3. Блок-схема алгоритма калибровки НПБК

НПБК строится отдельно для каждого субъекта, то есть для каждого класса обучающего набора биометрических данных. Процедура синтеза и обучения НПБК осуществляется автоматически без применения итерационного обучения на основе метода обратного распространения ошибки. Алгоритм обучения НПБК приведен на рисунке 4.

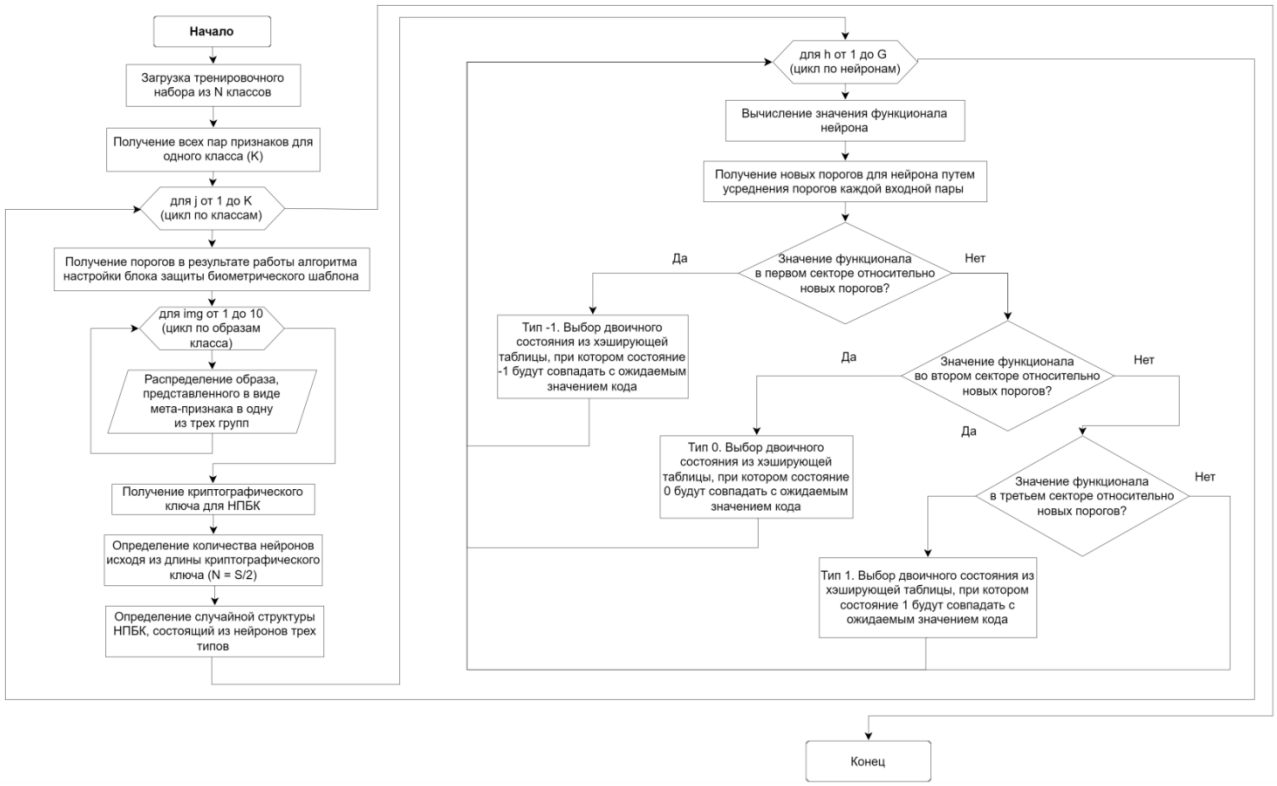


Рисунок 4. Блок-схема алгоритма обучения НПБК на основе тригонометрического нейрона

Первым этапом построения НПБК на базе тригонометрического нейрона будет отбор таких пар признаков обучающего набора данных, в подпространстве которых биометрические образы субъекта будут располагаться строго в одном из трех секторов. Отметим, что «строгость» расположения образов в секторах можно варьировать, исходя из объема обучающей выборки «Свой» (чем больше обучающих примеров, тем больше отклонений от заданного сектора допустимо). При этом следует учитывать, чем больше обучающих примеров «Свой» лежит вне выбранного сектора, тем выше будет вероятность ошибок «ложного отказа». Слишком строгие правила при больших объемах обучающей выборки «Свой» могут привести к невозможности синтеза НПБК из-за отсутствия достаточного количества подходящих пар признаков.

Формируется три непересекающиеся группы пар признаков (отдельная группа на каждый сектор). Для каждого нейрона случайным образом выбираются пары признаков из определенной (одной) группы. Количество пар признаков равно числу синапсов нейрона. Входы каждого нейрона должны быть уникальны и ни одна пара не может повторно использоваться в другом нейроне. На базе каждой группы пар признаков формируется примерно равное количество нейронов, чтобы избежать статистических смещений и снижения информационной энтропии кодов на выходе НПБК при поступлении на его входы образов «Чужой» (допускается незначительное расхождение в 2-3 нейрона). При случайном перемешивании нейронов в структуре НПБК становится невозможно осуществить направленный перебор входных образов для определения ключа пользователя (или его частей).

Пороговые значения для функции активации (4) нейрона вычисляются по формулам:

$$T_1 = \frac{1}{k}\sum_{z=1}^{k} t_{1z} \qquad \text{и} \qquad T_2 = \frac{1}{k}\sum_{z=1}^{k} t_{2z},$$

где k – количество синапсов (входов) нейрона; z – номер синапса (входа) нейрона; t_{1z} и t_{2z} – пороги для пары признаков, поступающей на вход нейрона, полученные в результате работы алгоритма калибровки.

Структура НПБК должна строиться из расчета $N = S/q$, где N – количество нейронов, S – желаемая длина криптографического ключа, q – количество бит, продуцируемое одним нейроном и вычисляемое по формуле:

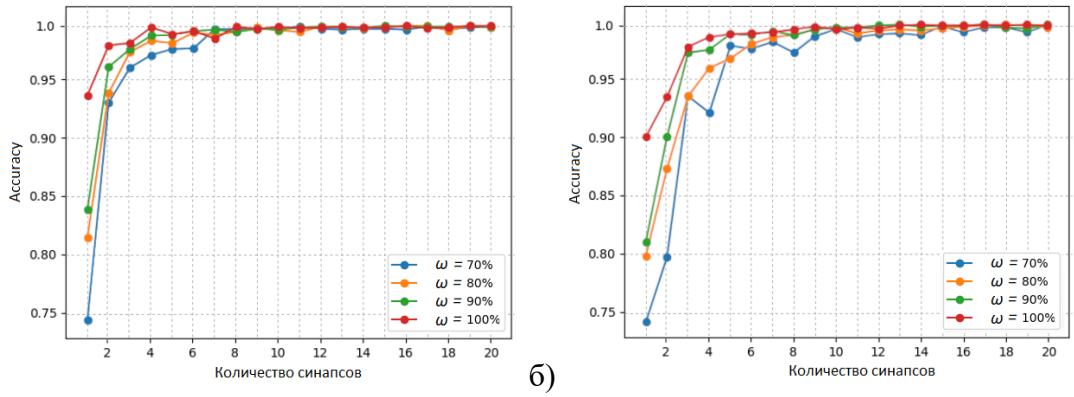
$$q = [\, log_2 h \,],$$

где h – количество порогов, делящее функцию плотности вероятности мета-признаков на равные секторы ($h = m\text{-}1$), [.] – операция округления до большего значения.

После обучения нейронов распределение хеш-преобразований является равномерным, поэтому достигается равновероятное появление состояний «0» и «1» на выходе НПБК при поступлении на входы образов «Чужих», и как следствие высокая энтропия.

Перед первичным синтезом и обучением НПБК экспериментально были определены оптимальные параметры конфигурации нейросетевого преобразователя (рис. 5). Указанными параметрами являются количество входов одного нейрона (количество синапсов) и процент образов «Свой» обучающего набора, попадающих в один из трех секторов, при котором принимается решение отнесения пары признаков к одной из трех групп. Следует подбирать такие конфигурации нейрона, при которых возможно получение максимальной точности при наименьших значениях k и ω. Это позволит получить ключ наивысшей длины.

С помощью графиков, представленных на рисунке 5, были отобраны и протестированы по несколько оптимальных конфигурации для каждого из функционалов НПБК. Результаты эксперимента, проведенного на специально сформированном наборе данных (SFDv1), включающем в себя видеозаписи лиц 75 испытуемых, представлены в таблице 2.



а) б)

Рисунок 5. Зависимость точности работы НПБК от количества синапсов нейрона при разных процентных значениях образов «Свой», попадающих в один из трех секторов:
а) функционал (1); б) функционал (2)

Обозначенным требованиям соответствуют только одна конфигурация НПБК на базе меры (2). Тем не менее, такой НПБК способен продуцировать 2048 битный ключ, используя всего 70% образов попадания в один из секторов. В случае нейронов на базе меры (2) подходящими являются сразу две конфигурации, позволяющие для 100% пользователей успешно продуцировать длинные криптографические ключи (2048 бит). Однако согласно эксперименту, представленному на рисунке 5, наиболее высокой точностью распознавания отдельного нейрона обладает конфигурация №1 (99,93% против 99,81% для №2). В связи с этим, указанная конфигурация была выбрана в качестве основной для НПБК на базе меры (2).

Таблица 2. –Тестирование оптимальных конфигураций НПБК для двух функционалов

№	«Строгость» попадания в сектор, ω	Количество синапсов, k	Количество нейронов	Длина ключа, бит	Успешный синтез
Нейрон на базе меры (2)					
1	70%	7	256	512	100%
			512	1024	100%
			1024	2048	100%
Нейрон на базе меры (3)					
1	70%	10	256	512	100%
			512	1024	100%

			1024	2048	100%
2	80%	9	256	512	100%
			512	1024	100%
			1024	2048	100%

Итоговая оценка эффективности выбранного варианта исполнения НПБК на базе меры (2) для разных наборов данных представлена на рисунке 6 (представлены средние значения). Для обучения НПБК потребовалось всего 70% образов «Свой» от 10 представленных для каждого пользователя в обучающем наборе данных (7 образов). Малое количество обучающих примеров обеспечивается за счет работы алгоритмов предварительной настройки и обучения НПБК. Как видно из рисунков, лучшего показателя удалось достичь при проведении экспериментов на основе набора данных Faces94 (EER ≈ 0.008). При этом длина ключа достигает значения в 2048 бит, что является высоким показателем.

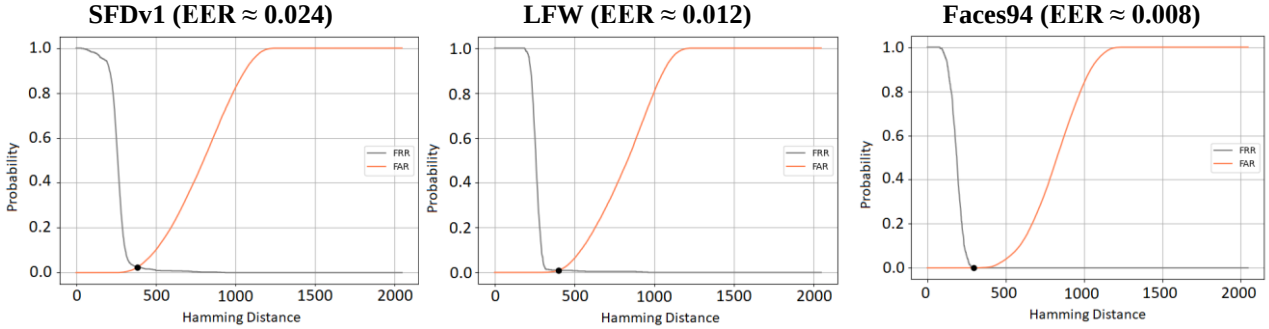


Рисунок 6. Тестирование НПБК на базе тригонометрических нейронов на разных наборах данных

В четвертой главе представлена структура системы защищенной биометрической аутентификации на основе предложенных в работе концепции, моделей и алгоритмов (рис. 7). В основе структуры лежит способ построения ЗБА, при котором ключом НПБК для обнаружения спуфинг атак осуществляется гаммирование структуры пользовательского НПБК. Разработанная структура состоит из трех функциональных блоков:

1. *Блок детекции лиц и извлечения признаков*. Глубокие нейронные сети блока не требуют повторного переобучения при изменении контекста функционирования системы или компрометации биометрических образов и, при необходимости, могут быть заменены альтернативными архитектурами.

2. *Блок обучения нейросетевых преобразователей* для задач распознавания лиц и обнаружения спуфинг атак. НПБК для задачи обнаружения спуфинг атак обучается один раз в соответствии с процедурой, описанной в главе 2 и универсален для всех пользователей. Для пользовательского НПБК определена процедура шифрования параметров НБК ключом НПБК для обнаружения спуфинг атак. Для этого предполагается, что НБК можно представить в виде матрицы (таблицы):

$$F = \|f_{ij}\|, \quad i = \overline{1,n} \quad j = \overline{1,m}$$

где F – матрица номеров признаков НПБК, n – количество синапсов одного нейрона, m – общее число нейронов НПБК. Каждое значение f_{ij} матриц может быть представлено в двоичной системе счисления и записано в соответствующий двоичный вектор:

$$\bar{f} = (0,1,1,0 \ldots 0)$$

где $\bar{f}$ – двоичный вектор, описывающий структуру (последовательность номеров признаков) НПБК. Полученные в результате рассмотренных преобразований векторы подвергаются гаммированию внешнем ключом НПБК для обнаружения спуфинг атак. Для этого методом скользящего окна осуществляется сложение по модулю 2 N-битного ключа K_2 и равного ему участка соответствующего двоичного вектора. Смещение окна производится на значение длины ключа K_2, а в случае если ключ K_2 «не умещается» в $|\bar{f}|$, остаток вектора складывается с его началом.

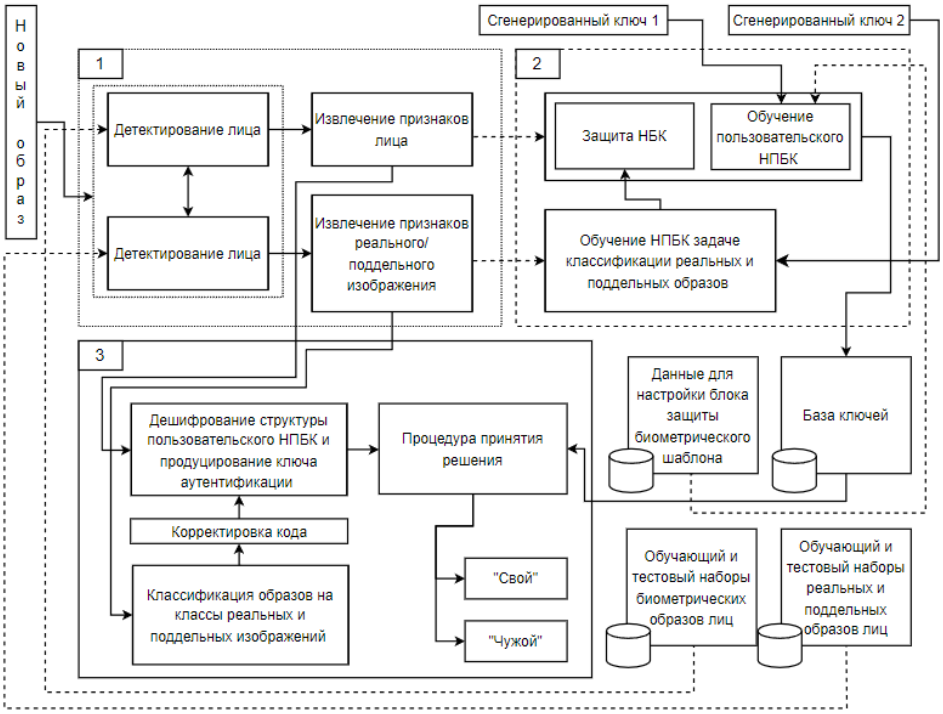


Рисунок 7. Структура системы защищенной биометрической аутентификации по лицу

3. Блок аутентификации, осуществляющий распределение новых входных образов по классам «Свой»/«Чужой». Аутентификация субъекта осуществляется в два этапа: 1) полученный на вход образ определяется как поддельное или реальное изображение; 2) в случае реального изображения ключ НПБК для обнаружения спуфинг атак успешно дешифровывает параметры НБК, после чего пользовательский НПБК принимает решение об аутентификации. В силу того, что сложение по модулю является обратной по отношению к себе операцией, дешифрование структуры НПБК осуществляется аналогичным шифрованию образом: $|\bar{f}|' \oplus K_2$, где $|\bar{f}|'$ – зашифрованный ранее двоичный вектор. Для нивелирования эффекта повышения энтропии выходов пользовательского НПБК при поступлении образа «Чужой» или спуфинг-образа при режиме работы ЗНК необходимо осуществлять предварительную корректировку кода на выходе НПБК для обнаружения спуфинг атак. Описанную корректировку предлагается реализовывать с помощью кодов Безяева, позволяющих избегать применения избыточных самокорректирующихся кодов.

Система на основе предложенной структуры реализована в качестве программного комплекса на языке Python, а также протестирована на базе разработанной в ходе исследования среды управления жизненным циклом ИИ «AIC ModelOps Platform». Итоговое тестирование системы проводилось на специально сформированном наборе данных из 120 субъектов. Коэффициент равной вероятности ошибок при работе системы составил EER=0,025, что говорит о сравнительно низком уровне ошибок распознавания образов при высоком уровне защищенности процедуры биометрической аутентификации по лицу от атак на биометрическое предъявление, а также атак компрометации знаний НПБК и биометрических данных.

В заключении приводятся основные результаты и выводы, полученные автором.

В приложениях приведены акты внедрения и свидетельства о регистрации программ для ЭВМ и электронных ресурсов.

ОСНОВНЫЕ ВЫВОДЫ И РЕЗУЛЬТАТЫ РАБОТЫ

1. Разработана концепция защищенной биометрической аутентификации по лицу, обеспечивающая устойчивость к атакам на биометрическое предъявление (спуфинг атак). За счет применения дополнительного нейросетевого преобразователя, обученного задаче классификации реальных и поддельных изображений, решается задача противодействия спуфинг атакам, а также обеспечивается дополнительная защита параметров

пользовательского НПБК. Защита осуществляется путем применения механизма защиты нейросетевого контейнера (ЗНК). Точность классификации нейросетевого преобразователя, адаптированного для задачи обнаружения спуфинг атак, составляет 97,2% на тестовых выборках, что сравнимо с точностью распознавания спуфинг атак с помощью аналогичных архитектур, обучаемых на основе Cross-Entropy Loss.

2. Разработаны модель тригонометрического нейрона и основанная не ней модель нейросетевого преобразователя биометрия-код (НПБК), позволяющие работать со слабо коррелированными признаками лица человека и продуцировать длинный криптографический ключ на выходе НПБК при высокой точности классификации биометрических образов. Предложенные модели не используют параметры распределений и/или характеристики образов «Свой», что обеспечивает защиту биометрических данных и знаний НПБК от компрометации. В рамках проведенных экспериментов продемонстрированы высокие показатели эффективности итогового исполнения НПБК, продуцирующего ключ длиной в 2048 бит по сравнению со 128 битами классического НПБК, обученного в соответствии с ГОСТ Р 52633.5. При этом достигнуты сравнительно высокие показатели точности распознавания EER≈0.024.

3. Разработаны алгоритмы предварительной настройки параметров нейросетевого преобразователя биометрия-код и алгоритм автоматического обучения НПБК на основе тригонометрических нейронов на малых выборках. Совместная работа алгоритмов позволяет предварительно оценить особенности распределений биометрических образов лиц, не компрометирующих легитимных пользователей, с целью последующей сборки и быстрого обучения НПБК, не требующего большого числа примеров «Свой». Проведенные эксперименты продемонстрировали возможность использования всего 7 образов.

4. Разработана структура системы защищенной биометрической аутентификации по лицу, в которой за счет работы независимых блоков извлечения признаков, обучения нейросетевых преобразователей и аутентификации, а также применения варианта исполнения ЗНК, при котором шифруется структура (порядок расположения синапсов в нейронах) пользовательского НПБК обеспечивается защищенность процедуры биометрической аутентификации личности по лицу на основе НПБК в отношении спуфинг атак и атак компрометации знаний НПБК и открытых биометрических образов лиц. Коэффициент равной вероятности ошибок при работе системы составил EER=0,025, что говорит о сравнительно низком уровне ошибок распознавания образов при высоком уровне защищенности процедуры биометрической аутентификации по лицу от атак на биометрическое предъявление, а также атак компрометации знаний НПБК и биометрических данных.

Результаты диссертационного исследования приняты к внедрению в производственные и бизнес-процессы компаний ООО «Открытый код» г. Самары и ООО «АИ ЗИОН» г. Омска, а также в учебные процессы ФГБОУ ВО «Самарский государственный технический университет» и ФГАОУ ВО «Омский государственный технический университет».

Перспективы дальнейшего развития темы. В рамках дальнейших исследований планируется разработка защищенной биометрической аутентификации, устойчивой к новому типу атак – дипфейкам.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи, опубликованные в журналах, входящих в перечень рецензируемых научных журналов и изданий, рекомендованных ВАК:

1. Панфилова И. Е., Сулавко А. Е. Методы определения живого присутствия пользователя перед видеокамерой в задачах биометрической аутентификации по лицу / Вопросы защиты информации: Науч.-практ. журн. / ФГУП «НТЦ оборонного комплекса «Компас», 2023. Вып. 2 (141). С. 17–26.

2. Васильев В.И., Панфилова И.Е., Сулавко А.Е., Серикова А.Е. Система верификации личности по изображению лица в защищенном режиме на основе искусственных нейронных сетей / Прикладная информатика, 2023. Т. 18. № 5. С. 33–47.

3. Панфилова И. Е. Глубокие нейронные сети в задачах идентификации и верификации лиц / Вопросы защиты информации: Науч.-практ. журн. / ФГУП «НТЦ оборонного комплекса «Компас», 2024. Вып. 2 (145). С. 33–41.

4. Жумажанова С. С., Панфилова И. Е., Ложников П. С., Сулавко А. Е., Серикова А. Е. Биометрическая аутентификация по тепловым изображениям лица на основе преобразователей "биометрия-код" / Вопросы защиты информации: Науч.-практ. журн. / ФГУП «НТЦ оборонного комплекса «Компас», 2023. Вып. 1 (140). С. 8–19.

5. Панфилова И.Е., Ложников П.С. Исследование применимости нейросетевых преобразователей «биометрия-код» для задачи обнаружения атак на биометрическое предъявление / Вестник УрФО № 2(52), 2024. С. 106–121

6. Панфилова И. Е., Сулавко А. Е., Ложников П. С. Повышение защищенности процедуры биометрической аутентификации по лицу на основе нейросетевых преобразователей «биометрия-код» / Вопросы защиты информации: Науч.-практ. журн. / ФГУП «НТЦ оборонного комплекса «Компас», 2024. Вып. 3 (146). С. 3-11

Статьи, индексируемые в международной базе Scopus:

7. A. Sulavko, I. Panfilova, A. Samotuga and S. Zhumazanova, "Biometric Authentication Using Face Thermal Images Based on Neural Fuzzy Extractor," 2023 Intelligent Methods, Systems, and Applications (IMSA), Giza, Egypt, 2023, pp. 80-85, doi: 10.1109/IMSA58542.2023.10217752.

Прочие публикации:

8. Сулавко А.Е., Панфилова И.Е. Верификация личности субъектов по лицу на основе методов глубокого обучения и нейросетевых преобразователей «биометрия-код» / Нанотехнологии. Информация. Радиотехника (НИР-23) : материалы Всерос. молодеж. науч.-практ. конф. (Омск, 18 апр. 2023 г.) / Минобрнауки России, Ом. гос. техн. ун-т, Радиотехн. фак., Каф. «Физика» ; редкол.: Н. О. Голубятникова [и др.]. – Омск : Изд-во ОмГТУ, 2023. С. 336–339

9. Панфилова И.Е., Иниватов Д.П. Обзор методов защиты данных биометрических шаблонов / Безопасность информационных технологий: сб. науч. ст. по материалам V Всерос. науч.-техн. конф., посвящ. 70-летнему юбилею АО «НПП "Рубин"» (г. Пенза, 27 сентября 2023 г.) : в 2 т. – Пенза : Изд-во ПГУ, 2023. Т. 1. С. 135–145.

10. Панфилова И.Е. Методы и алгоритмы повышения надежности нейросетевой биометрической аутентификации / Сборник трудов III Всероссийской научной школы-семинара «Современные тенденции развития методов и технологий защиты информации». Москва, МТУСИ, 25-237 октября 2023 г. М., 2023. С. 289–295.

Свидетельства о государственной регистрации программы для ЭВМ

11. Свидетельство о государственной регистрации программы для ЭВМ: Реализация моделей и алгоритмов обучения нейросетевых преобразователей биометрия-код на основе тригонометрических нейронов, позволяющих симметризовать классы образов относительно пространства признаков: Свидетельство о государственной регистрации программы для ЭВМ № 2023684211: / И. Е. Панфилова, А. Е. Сулавко, А. Е. Серикова, Ю. Дорогов. – заявка № 2023682402 заявл. 27.10.2023: опубл. 14.11.2023

12. Свидетельство о государственной регистрации программ для ЭВМ: AIC ModelOps Platform: Свидетельство о государственной регистрации программы для ЭВМ № 2022680686/ А.Е. Сулавко, Д.Г. Стадников, А.Г. Чобан, А.Е. Самотуга, И.Е. Панфилова. – заявка № 2022668418; заявл. 10.10.2022; опубл. 3.11.2022.

Диссертант  Панфилова И.Е.