На правах рукописи

Иниватов Даниил Павлович

НЕЙРОСЕТЕВАЯ АУТЕНТИФИКАЦИЯ ПО ГОЛОСОВЫМ ПАРОЛЯМ С ЗАЩИТОЙ БИОМЕТРИЧЕСКИХ ЭТАЛОНОВ И УЧЁТОМ ФУНКЦИОНАЛЬНОГО СОСТОЯНИЯ ПОЛЬЗОВАТЕЛЯ

Специальность 2.3.6. Методы и системы защиты информации, информационная безопасность

АВТОРЕФЕРАТ

диссертации на соискание ученой степени кандидата технических наук

Омск – 2025

Работа выполнена на кафедре комплексной защиты информации ФГАОУ ВО «Омский государственный технический университет».

Научный руководитель: доктор технических наук,
Сулавко Алексей Евгеньевич

Официальные оппоненты:
Катасёв Алексей Сергеевич, доктор технических наук, профессор, федеральное государственное бюджетное образовательное учреждение высшего образования «Казанский национальный исследовательский технический университет им. А.Н. Туполева-КАИ», профессор кафедры систем информационной безопасности
Исмагилова Альбина Сабирьяновна, доктор физико-математических наук, доцент, федеральное государственное бюджетное образовательное учреждение высшего образования «Уфимский университет науки и технологий», заведующий кафедрой управления информационной безопасностью

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

Защита диссертации состоится 19.01.2026 г. в 10 ч. 00 мин. на заседании диссертационного совета 24.2.479.07 на базе ФГБОУ ВО «Уфимский университет науки и технологий», по адресу: 450008, г. Уфа, ул. К. Маркса, д. 12.

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

Автореферат разослан «___ » _____ 2025 года.

Ученый секретарь
диссертационного совета,
доктор технических наук  Вульфин Алексей Михайлович

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Современные биометрические технологии становятся неотъемлемой частью систем информационной безопасности, находя применение в повседневных задачах, таких как аутентификация в смартфонах, получение доступа к банковским системам или подтверждение личности перед началом рабочего дня в некоторых сервисах. Быстрый рост вычислительных мощностей, развитие множества типов датчиков и алгоритмов обработки данных создали условия для повсеместного внедрения биометрических методов идентификации личности. Голосовая биометрия сочетает удобство бесконтактной аутентификации и способность оценивать функциональное состояние (ФС) пользователя, однако её широкое применение ограничено уязвимостями к атакам воспроизведения, синтеза речи и компрометации шаблонов, а также более высоким уровнем ошибок по сравнению со статическими методами (лицо, отпечатки пальцев, радужная оболочка). При этом мировой рынок голосовой биометрии в 2023 году оценивался в 2,5 млрд долларов США. Согласно прогнозам, к 2032 году его объем достигнет 11,5 млрд, демонстрируя среднегодовой темп роста на уровне 18,7%, что подчёркивает стратегическое значение технологий голосового распознавания.

В связи с принятием ФЗ № 572 «Об осуществлении идентификации и (или) аутентификации физических лиц с использованием биометрических персональных данных, о внесении изменений в отдельные законодательные акты Российской Федерации и признании утратившими силу отдельных положений законодательных актов Российской Федерации» регламентирующего порядок использования изображений лица и голосовых характеристик человека, современная биометрическая система должна строиться на основе доверенного искусственного интеллекта (ИИ), устойчивого к деструктивным факторам (атаки на биометрические системы, дрейф биометрических данных, ФС пользователя) и обладающего поддержкой защищенного режима исполнения.

Степень разработанности темы исследования. 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, применении шифрования. В России действует серия национальных стандартов ГОСТ Р 52633, не имеющая 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.

Вопросам биометрической аутентификации, оценки изменчивости биометрических параметров, обеспечения защищённости биометрических данных посвятили множество исследований такие учёные как Ахметов Б.С., Брюхомицкий Ю.А., Епифанцев Б.Н., Жумажанова С.С., Иванов А.И., Катасёв А.С., Ложников П.С., Трошков А.М., Dodis Y., Jain A.K., Khan A., Monrose F., Schiel F., Snyder D. и др. Однако, несмотря на высокие показатели точности голосовой аутентификации, разрабатываемые системы демонстрируют низкую устойчивость к дрейфу биометрических данных вследствие изменённого ФС пользователя. Также текущие законодательные акты и стандарты в области биометрии не гарантируют защиту речевого образа от фальсификации в связи с возможностью кражи голосовых параметров пользователя. При наличии у злоумышленника речевых записей легального пользователя системы он сможет ими воспользоваться для проникновения. В условиях уязвимости речевых образов к атакам воспроизведения и компрометации становится актуальной научная задача разработки биометрической системы аутентификации по голосу, обеспечивающей защиту биометрических шаблонов и устойчивость к изменениям ФС пользователя с использованием тайного пароля или контрольной фразы. Для её решения необходимо построить модель изменения голосовых параметров в зависимости от ФС пользователя, разработать метод извлечения робастных признаков из голосовых

сигналов, алгоритма настройки модели нейросетевого преобразователя из голосовых данных в код, обеспечивающего одновременно защищенность биометрического шаблона и устойчивость к изменению состояния и методику аутентификации пользователя по голосу с учетом его ФС.

Объект исследования. Системы биометрической аутентификации по голосу.

Предмет исследования. Нейросетевые модели, методы и алгоритмы машинного обучения для защищенного исполнения процедур биометрической аутентификации.

Цель диссертационной работы: повысить точность и устойчивость к дрейфу биометрических данных процедуры аутентификации пользователя по голосу на основе нейросетевого преобразователя «биометрия-код» с обеспечением защиты биометрических шаблонов от компрометации.

Поставлены и решены следующие основные задачи:

1. Разработка статистической модели изменения характеристик голосового сигнала в зависимости от ФС пользователя, отражающей влияние состояния пользователя на параметры его голоса.
2. Разработка и исследование метода извлечения информативных признаков из голосового сигнала для повышения точности аутентификации.
3. Разработка алгоритма настройки модели нейросетевого преобразователя «биометрия-код», устойчивого к дрейфу голосовых признаков.
4. Разработка методики аутентификации пользователя по голосу с учетом его ФС и с обеспечением защиты биометрического шаблона от компрометации.
5. Разработка программного модуля, реализующего алгоритм настройки гибридной модели нейросетевого преобразователя «биометрия-код», устойчивого к дрейфу данных.

Методы исследования. В ходе диссертационного исследования применялись методы распознавания образов, машинного обучения (МО), защиты данных от компрометации, аппарат искусственных нейронных сетей (ИНС), ансамблевые методы, методы теории вероятностей и математической статистики, спектрального и корреляционного анализа, идентификации и аутентификации.

Достоверность результатов исследования подтверждается корректной постановкой задач, применением апробированных методов, успешно используемых в других прикладных областях, практической реализацией системы на основе разработанных моделей и алгоритмов, а также представлением результатов на научных конференциях и публикацией в научных изданиях, в том числе в изданиях из Перечня ВАК, актами о внедрении и использовании результатов работы в образовательную и производственную сферы.

Научная новизна:

1. Предложена статистическая модель голосовых образов пользователя, основанная на экспериментальных данных об изменении голосовых параметров у испытуемых обоих полов от 18 до 50 лет в различных ФС, которая, в отличие от существующих ранее моделей, учитывает изменения голосовых биометрических признаков в различных ФС пользователя и позволяет предсказывать характер дрейфа биометрических данных.
2. Предложен метод извлечения признаков из голосовых образов, основанный на ансамбле автокодировщиков, который, в отличие от традиционных методов, осуществляет расширение набора биометрических данных на этапе распознавания пользователя, что позволяет повысить точность аутентификации при дрейфе данных за счёт учёта спектральных особенностей, извлечённых из различных репрезентаций голосового сигнала.
3. Разработан алгоритм настройки НПБК с обеспечением защиты биометрических шаблонов от компрометации, основанный на использовании двух типов нейронов, каждый из которых осуществляет обработку голосовых данных с различным уровнем взаимной корреляции, что, в отличие от ранее известных алгоритмов обучения НПБК, обеспечивает более высокую точность аутентификации пользователя в условиях дрейфа биометрических данных.

4. Разработана методика аутентификации по голосу с распознаванием ФС пользователя, основанная на ансамблевых методах МО, что, в отличие от существующих ранее методик, позволяет проводить не только верификацию личности, но и выполнять классификацию ФС пользователя, регистрируя отклонения его состояния от нормы, что служит решающим фактором для ограничения прав доступа в процессе авторизации.

Теоретическая значимость диссертационной работы заключается в предложенной статистической модели изменения характеристик голосового сигнала в зависимости от состояния пользователя, основанной на экспериментальных данных об амплитудно-частотных характеристиках речи, которая учитывает неоднородность реакций пользователей в различных состояниях. Эта модель позволяет прогнозировать области спектра, подверженные дрейфу голосовых признаков, вызванному изменениями ФС. Разработаны новый алгоритм настройки НПБК, учитывающего влияние ФС пользователя на голосовые характеристики, метод извлечения признаков из голосового образа и методика аутентификации по голосу с распознаванием ФС пользователя. Разработанный метод извлечения признаков на основе ансамбля автокодировщиков, осуществляющего расширение набора биометрических данных на этапе распознавания образа увеличивает точность и надёжность аутентификации. Особый вклад вносит анализ корреляционных связей голосовых параметров, на основе которого происходит выбор типа нейронов, используемых в НПБК, что повышает точность аутентификации и устойчивость к дрейфу этих параметров, в частности при изменении функционального состояния.

Практическая значимость заключается в повышении надёжности биометрической аутентификации субъекта по голосу и защищённости компьютерных ресурсов от неавторизованного доступа при помощи разработанного программного комплекса. Коэффициент равной вероятности ошибок в 2,1% для задачи аутентификации пользователя подтверждает высокий уровень точности и даёт возможность настроить систему на $FAR<10^{-4}$ при FRR = 0,129, что обеспечивает высокую степень защищённости процедуры аутентификации по голосу.

Положения, выносимые на защиту:

1. Статистическая модель, учитывающая особенности изменения голосовых образов пользователя в различных состояниях, позволяющая прогнозировать дрейф данных и выявлять признаки, чувствительные к изменению функционального состояния пользователя.
2. Метод извлечения признаков при помощи ансамблирования автокодировщиков на базе свёрточных слоёв, обрабатывающих различные представления одного и того же голосового образа, что позволяет извлекать более полную информацию, необходимую для биометрической аутентификации, благодаря учёту спектральных особенностей различных репрезентаций голосового сигнала.
3. Алгоритм настройки гибридной модели нейросетевого преобразователя биометрия-код на основе двух типов нейронов, чувствительных к корреляции между признаками с обеспечением защиты биометрических шаблонов от компрометации.
4. Методика биометрической аутентификации по голосу с распознаванием ФС, которая позволяет адаптировать уровень прав пользователя в зависимости от выявленного состояния.
5. Программный модуль, реализующий алгоритм настройки гибридной модели нейросетевого преобразователя биометрия-код, устойчивого к дрейфу данных и обеспечивающего защиту биометрических шаблонов.

Соответствие научной специальности. Работа соответствует научной специальности 2.3.6 Методы и системы защиты информации, информационная безопасность, п. 12: Технологии идентификации и аутентификации пользователей и субъектов информационных процессов. Системы разграничения доступа, и п. 15: Принципы и решения (технические, математические, организационные и др.) по созданию новых и совершенствованию существующих средств защиты информации и обеспечения

информационной безопасности.

Апробация работы. Основные положения и результаты работы докладывались на следующих конференциях: 14th International Scientific-Technical Conference on Actual Problems of Electronic Instrument Engineering" (APEIE 2018), Новосибирск; II Межвузовская научно-практическая конференция «Информационная безопасность: современная теория и практика» (г. Омск, 2019 год); Девятый международный круглый стол «Современная мировая экономика: проблемы и перспективы в эпоху развития цифровых технологий и биотехнологии» (г. Москва, 2019 год); Всероссийская научно-практическая конференция с международным участием «Россия и мировые тенденции развития» (г. Омск, 2021 год); II Международная научно-практическая конференция «Цифровизация и кибербезопасность: современная теория и практика» (г. Омск, 2022 год); VII Международная научно-практическая конференция «Фундаментальные и прикладные исследования молодых учёных» (г. Омск, 2023 год); XVII Международная научно-техническая конференция «Динамика систем, механизмов и машин» (г. Омск, 2023 год).

Связь с научными программами. Исследование выполнено при финансовой поддержке Минцифры России (грант ИБ МТУСИ), проект № 40469-15/2022-к в 2022-2023 годах, в котором аспирант был руководителем проекта. Также данная работа была частично профинансирована в рамках государственного задания Минобрнауки России в период 2023–2025 гг. № FSGF-2023-0004.

Публикации. По материалам диссертации опубликовано 17 работ, 4 из которых изданы в журналах, рекомендованных перечнем ВАК; 5 научных публикаций индексированы в международной информационно-аналитической системе научного цитирования Scopus; 8 научных работ опубликованы в других изданиях. Получено 4 свидетельства о регистрации программ для ЭВМ.

Личный вклад. Все представленные в диссертации результаты (метод, модели, алгоритмы, методика и программный модуль) и их анализ принадлежат лично автору. Формулирование цели и задач, способов их решения и вариантов представления осуществлены автором совместно с научным руководителем.

Объём и структура работы. Диссертационная работа состоит из введения, четырёх глав, заключения, списка литературы (119 наименований) и 2 приложений. Общий объём работы составляет 144 страницы включая 37 рисунков и 16 таблиц.

ОСНОВНЫЕ ПОЛОЖЕНИЯ РАБОТЫ

Во введении представлено обоснование актуальности темы диссертации, указаны цели и задачи исследования, а также дается общая характеристика работы. Подчеркивается важность разработки систем биометрической аутентификации в защищённом исполнении с учетом изменчивости ФС пользователя, что является актуальной задачей в условиях роста использования биометрических систем и увеличения угроз информационной безопасности.

В первой главе представлен обзор современных достижений и методов в области аутентификации пользователя по голосу. Рассматриваются ключевые задачи, над которыми работают исследователи в этой области, такие как борьба с шумом, дрейфом голосовых характеристик, спуфингом (подделка образа), состязательными атаками и др.

Приводятся таблицы сравнения результатов исследований, проведенных в этой области, отмечаются перспективные направления. Известно, что биометрическая аутентификация стала критически важным компонентом обеспечения безопасного доступа к системам. Она требует устойчивости к попыткам извлечения из неё знаний. Область защиты биометрических шаблонов (ЗБШ) включает в себя: нечёткие экстракторы, гомоморфное шифрование, позволяющее осуществлять обработку данных в зашифрованном виде, отменяемую биометрию и нейросетевой преобразователь «биометрия-код» (НПБК).

Гомоморфное шифрование позволяет производить обработку данных в зашифрованном виде, но оно требует большой вычислительной мощности, что затрудняет его использование в режиме реального времени. Нечёткий экстрактор извлекает биометрические признаки из специально обработанных данных, предоставляя

дополнительный уровень защиты. Ключевым подходом к ЗБШ служит НПБК, утверждённый стандартом ГОСТ Р 52633. Данная биометрическая криптосистема переходит от открытого шаблона к защищенному ключу, связывая первый со вторым. Она поддерживает схемы шифрования, которые включают подобное преобразование. В работе делается вывод на основе обзора, что НПБК обеспечивает более высокий уровень безопасности по сравнению с нечетким экстрактором, чему во многом способствует возможность генерации более длинных ключей и меньшую вероятность ошибок.

В результате обзора экспериментальных исследований выявлена хронология ключевых работ и их основных направлений. Делается вывод, что современные научные методы способны с высокой точностью справляться с задачами распознавания пользователя по голосу, борьбы с шумом, а также с атаками подделки. Также обзор подчёркивает необходимость дальнейших исследований в области идентификации пользователя по коротким высказываниям, учёта его ФС и защиты биометрических шаблонов.

Во второй главе диссертации исследуется влияние ФС на голосовые характеристики и их воздействие на точность распознавания пользователя, а также описана статистическая модель, отражающая изменения амплитудно-частотных характеристик (АЧХ) речи, а также разброс этих изменений для совокупности пользователей в различных ФС. Рассмотрены различные базы голосовых образов, среди которых можно выделить NIST SRE, VoxCeleb2, TIMIT, RedDots и предлагаемую AIC-spkr-130. Наиболее подходящими базами для нашего исследования является RedDots и AIC-spkr-130. RedDots состоит из записей 45 испытуемых, произносивших 22 фразы на еженедельных сессиях (всего около 98 тысяч аудиофайлов).

AIC-spkr-130 ориентирован на задачу аутентификации по голосовому паролю с защитой биометрического шаблона. Предлагаемый архив состоит из записей 130 испытуемых (78 мужчин и 52 женщин) возраста от 18 до 50 лет в формате wav и частотой дискретизации 8 кГц. Для сбора данных были привлечены русскоговорящие испытуемые без выраженных заболеваний или неврологических нарушений. Эксперимент проводился в начале рабочего дня после полноценного отдыха. Испытуемые последовательно «вводились» в разные ФС, в каждом из которых не менее 60 раз произносили разные парольные фразы на русском языке длительностью в 1-2 секунды. Использовались следующие ФС:

1. Исходное состояние, в котором испытуемый не подвергался воздействиям.

2.1. Состояния алкогольного опьянения с разным уровнем содержания алкоголя в крови: ИФС 1 – около 0.4‰, ИФС 2 – около 0.8‰, ИФС 3 – около 1.6‰.

3. Сонное состояние.

Для построения статистической модели, отражающей изменения АЧХ речи в различных ФС предложен метод, основанный на интегрировании амплитудного спектра в окрестности его экстремумов. Спектр разбивается на отрезки, для каждого из которых вычисляется отношение средней амплитуды в изменённом ФС к средней амплитуде в нормальном ФС. Это позволяет оценить изменения АЧХ голоса при изменении ФС (рис. 1 и 2).

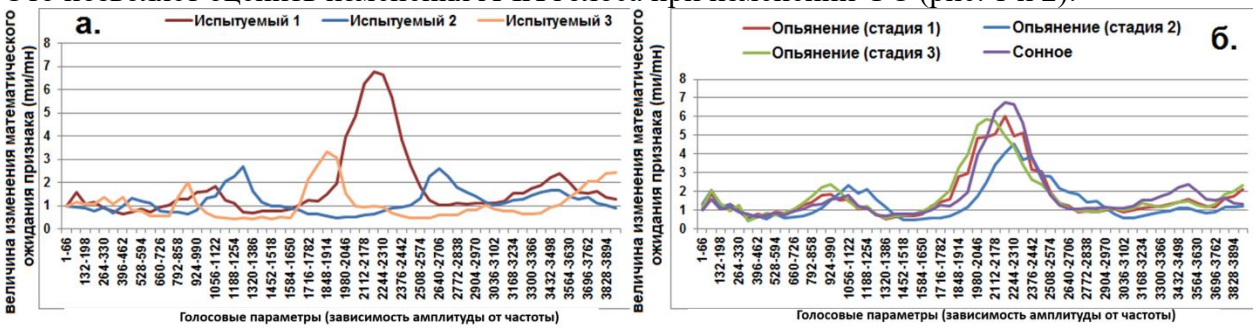

Рисунок 1 – Пример изменения средних значений голосовых параметров: а. для 3-х испытуемых в сонном состоянии, б. для одного пользователя в зависимости от состояния.

Почти всех испытуемых (порядка 90%) по характеру этих изменений можно разделить на 3 условных категории (рис. 2 а): субъекты с ярко выраженным пиком изменений в окрестности частот 1700-2700 Гц и двумя слабо выраженными (700-1300 Гц и 3200-3700

Гц), субъекты с двумя выраженными пиками в интервалах частот 1000-1400 Гц и 2500-2900 Гц, субъекты с тремя пиками – слабым (750-1000 Гц), сильным (1650-2100 Гц) и пиком в области частот 3650-4000 Гц. Во всех измененных состояниях для большинства испытуемых и всех произнесенных ими паролей признаки $\bar{a}_1$ меняются схожим образом (рис. 1 б). Для математического описания этих изменений предложена модель преобразования голосового сигнала. Пусть E представляет вектор амплитуд гармоник (1). Тогда он определяется как:

$$E = [a_1, a_2, a_3, \dots, a_n], \quad (1)$$

где n – количество гармоник, a_j - амплитуда j-й гармоники (j = 1,2, …, n). $K_и$ - вектор коэффициентов для изменённого ФС задаётся как (2):

$$K_и = \left[\frac{M(a_{и1})}{M(a_{н1})}, \frac{M(a_{и2})}{M(a_{н2})}, \frac{M(a_{и3})}{M(a_{н3})}, \dots \frac{M(a_{ип})}{M(a_{нп})}\right], \quad (2)$$

Тогда $E_и$ представляет вектор амплитуд гармоник (3) в изменённом ФС и вычисляется как произведение исходного вектора $E_н$ на коэффициент $K_и$:

$$E_и = E_н \cdot K_и = [a_{н1}, a_{н2}, a_{н3}, \dots, a_{нп}] \times \begin{bmatrix} \frac{M(a_{и1})}{M(a_{н1})} \\ \dots \\ \frac{M(a_{ип})}{M(a_{нп})} \end{bmatrix}, \quad (3)$$

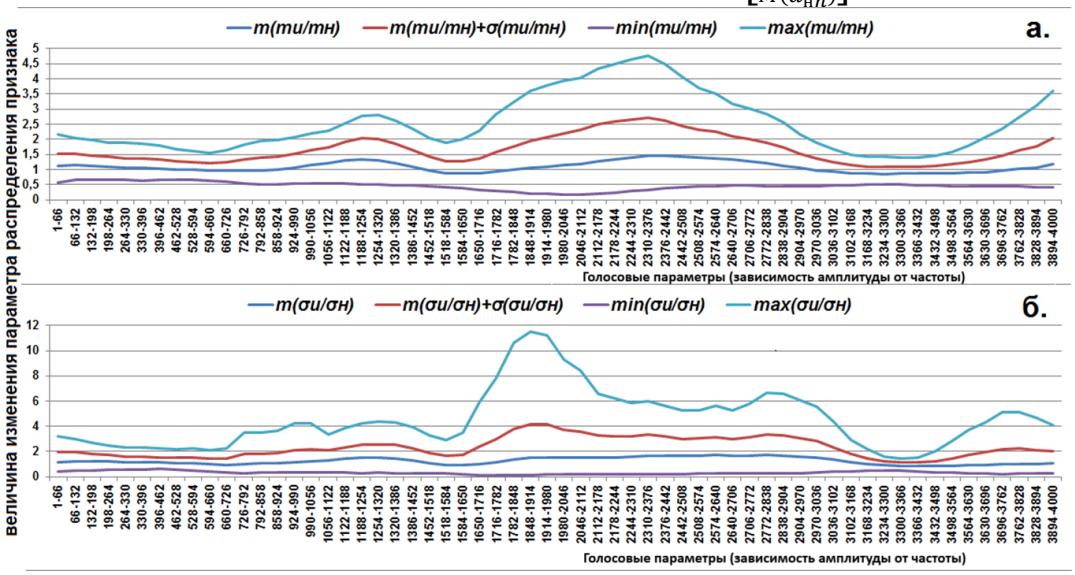

Рисунок 2 – Трансформация параметров распределения голосовых параметров при смене ФС пользователя с нормального (*н*) на измененное (*и*): а. математических ожиданий *m*(), б. среднеквадратичных отклонений *σ*() (*max* и *min* – функции максимума и минимума).

В результате проведённого исследования была описана статистическая модель, отражающая изменения АЧХ речи в различных ФС. Анализ дрейфа признаков позволил выявить области спектра, наиболее чувствительные к изменениям состояния. Также подтверждено, что изменённое ФС оказывает значительное влияние на голосовые характеристики, что, в свою очередь, влияет на точность распознавания пользователя. Доказана необходимость учета изменчивости ФС при проектировании и обучении биометрических систем.

В третьей главе диссертации описывается разработанный метод извлечения признаков из голосового сигнала. Для анализа особенностей голосовых сигналов часто используется быстрое оконное преобразование Фурье. Отсчёты Фурье-спектрограмм для различных частот могут быть поданы на входы глубоким свёрточным нейронным сетям (СНС) для извлечения биометрических признаков.

Высокая размерность вектора входных данных требует больших объемов выборки для обучения многослойной НС, поэтому вместо отсчётов спектрограмм нами был использован усредненный по всем окнам амплитудный спектр (Рис. 3), который интегрирует информацию о локальных характеристиках голосовой записи и сглаживает случайные

выбросы. Получаемый в результате преобразования спектр зависит как от речевых характеристик испытуемого, так и от самого сообщения, поэтому если пользователь будет произносить одну и ту же фразу, то их спектры будут схожи. В случае замены голосовой фразы или испытуемого, существенные изменения будут заметны и на графике. В настоящем исследовании применяются репрезентации речевых образов, основанные на различных оконных функциях: прямоугольное, Гаусса, Блэкмана, Барлетта и Хемминга.

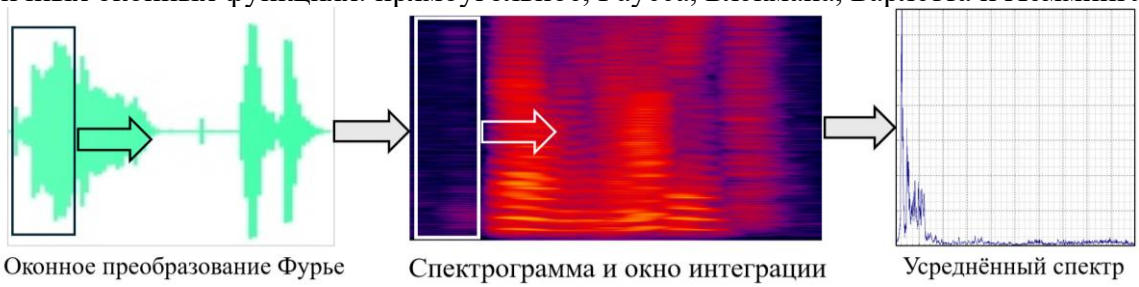

Рисунок 3 – Построение усреднённого амплитудного спектра голосового сигнала.

Под репрезентацией подразумевается промежуточное представление сигнала голоса, полученное после предобработки, но ещё не являющееся окончательным вектором признаков. Разные репрезентации могут использоваться для обучения различных классификаторов, которые можно объединить в единый ансамбль. Данная идея была основана на теореме Кондорсе. Согласно ей, если решения классификаторов независимы и вероятность правильного решения каждого >0,5, то с увеличением числа классификаторов вероятность принятия правильного решения комитета будет стремиться к единице.

В диссертационной работе предлагается комбинировать многослойные СНС и классический НПБК по принципу стекинга. Данный подход позволяет улучшить качество классификации, объединив сильные стороны обеих архитектур. СНС используются для выработки признаков, которые затем подаются на вход НПБК. Автокодировщик – это глубокая СНС, которая позволяет извлекать информативные признаки.

Он состоит из кодировщика, который сжимает входные данные в более компактный вид, и декодировщика, который обучается для их восстановления. Размерность входного слоя должна соответствовать выходному, а скрытые слои - иметь меньшую размерность. Данное свойство позволяет выявлять наиболее информативные признаки речевого образа и снижать размерность пространства признаков. Во время обучения на вход и выход подается усредненный спектр голосового пароля, что позволяет сети обучиться выделять ключевые особенности голосовых данных (признаки, которые будут поданы на вход НПБК).

В качестве основы использована архитектура VGG16, которая была модифицирована для обработки усреднённых спектров. В её состав вошли одномерные свёртки, слои с пакетной нормализацией, а также полносвязный слой с линейной функцией активации (рис. 4 и 5).

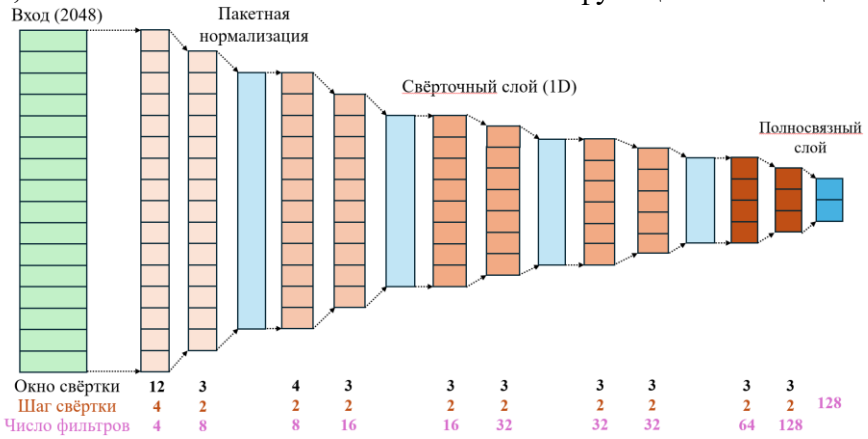

Рисунок 4 – Схема кодировщика для извлечения признаков

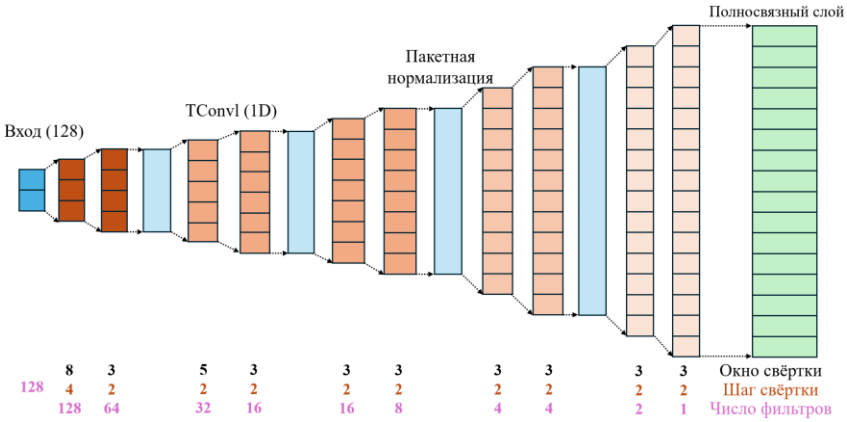

Рисунок 5 – Схема декодировщика

В качестве обучающей выборки был использован набор голосовых данных AIC-spkr-130, алгоритма оптимизации – Adam, функции ошибки – бинарная кросс-энтропия.

Для создания ансамбля было сформировано и обучено на 500 эпохах пять аналогичных топологий автокодировщиков. Их обучение проходило при идентичных условиях: одинаковая выборка, оптимизатор и функция ошибки. Но каждой СНС на вход были поданы разные репрезентации образов – усреднённые спектры, вычисленные с использованием разных оконных функций: прямоугольная, Гаусса, Блэкмана, Барлетта и Хемминга.

В процессе аутентификации речевой сигнал преобразовывается в 5 усреднённых спектров, каждый из которых подаётся на вход соответствующему кодировщику, на выходе которого вычисляется вектор признаков. Данные вектора комбинируются и подаются на вход НПБК, который на выходе производит личный ключ или пароль пользователя (рис. 6). Выявлено, что EER снижается при использовании комитета СНС (EER ≈ 2,1%) в сравнении с самой эффективной одиночной СНС (2,21%), что соответствует снижению EER на 5,07%.

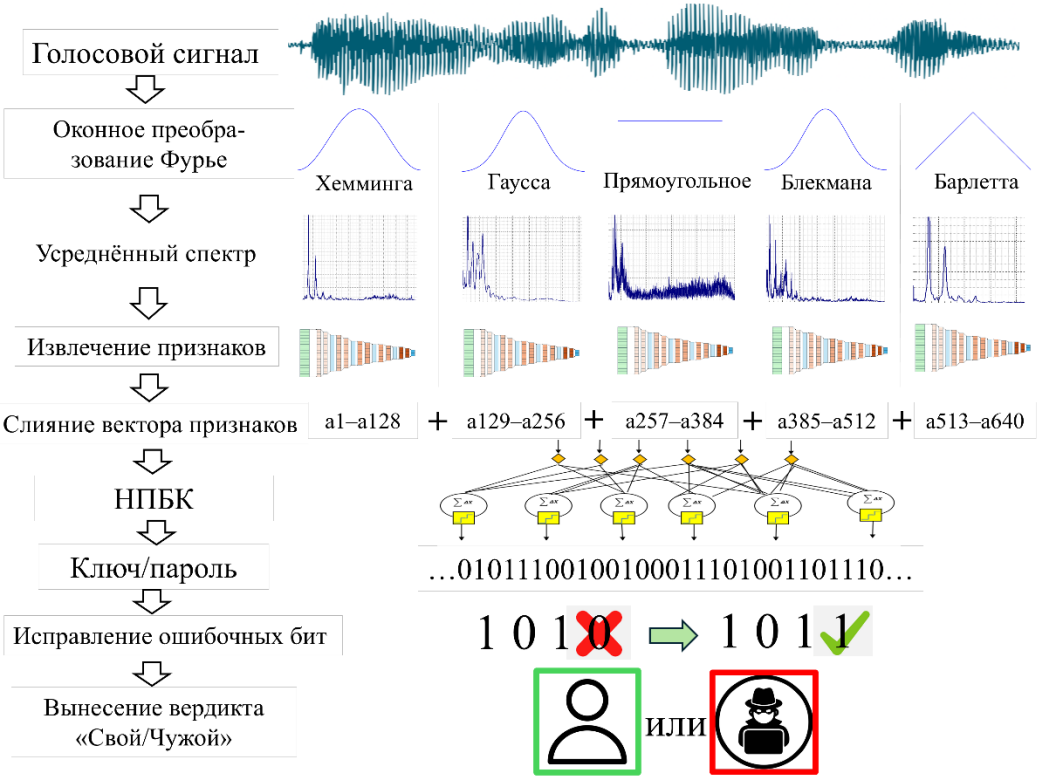

Рисунок 6 – Процесс аутентификации на базе СНС и НПБК

Представлено описание алгоритма настройки НПБК, процесса синтеза тригонометрического корреляционного нейрона, а также результаты проведённого эксперимента. В качестве входных значений для нейрона выступают пары признаков

вместо отдельных значений исходного вектора. Оперирование *мета-признаками*, учитывающими корреляционные связи в паре, позволяет НПБК надежно разделять образы легитимных («Свои») и нелегитимных пользователей («Чужие»). Получить мета-признаки, а также их альтернативные исполнения, можно путем следующего отображения:

$$a'_l = f(a_i, a_j), \tag{4}$$

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 (метрика Байеса-Минковского в случае корреляционного нейрона), l – номер мета-признака. Каждая пара признаков, участвующая в преобразовании, порождает свое *подпространство*, а пространство, порождаемое мета-признаками, называется *мета-пространством*.

В настоящей работе в качестве отображения (4) предлагается использовать аналог косинусного расстояния, работающего на уровне отдельных пар признаков вектора:

$$a'_l = cos(\widehat{\bar{d}, \bar{v}}), \tag{5}$$

Косинусное расстояние определяется между двумя векторами: вектор $\bar{d}$ соединяет «центр масс» (пересечение средних значений двух признаков) с точкой, представляющей искомый образ, а вектор $\bar{v}$ совпадает с первым, но располагается в начале тригонометрической окружности, имеющей тот же центр. Использование метрики (5) позволяет избегать применения евклидова расстояния, которое может демонстрировать одинаковые значения от «центра масс» как до образа «Свой», так и до образа «Чужой».

Однако в подпространствах, в которых признаки сильно коррелированы, векторы $\bar{d}$ и $\bar{v}$ могут быть почти коллинеарными. Косинусное расстояние измеряет только угол между векторами, но не учитывает их длину. Кроме того, если векторы лежат почти на одной прямой, косинусное расстояние может не сильно варьироваться, даже если угол между ними достаточно велик. Поэтому необходима дополнительная метрика, учитывающая корреляционные связи. Такой метрикой может выступать следующее преобразование:

$$a'_l = ctg(\widehat{\bar{d}, \bar{v}}), \tag{6}$$

Котангенс угла между векторами $\bar{d}$ и $\bar{v}$ предоставляет детализированное представление углового соотношения между векторами (рис. 7). Совместное использование косинуса и котангенса угла в основе нейронов позволяет сформировать комплексный подход к анализу векторов признаков, особенно в контексте высоко коррелированных данных, что приводит к уменьшению ошибок аутентификации. Как можно видеть, собственная область класса образов в пространстве коррелированных признаков «вытягивается», а метрика (6) позволяет лучше сепарировать образы коррелированных классов по секторам. Исследования показали, что нейроны на базе косинусного (синусного) преобразования лучше обрабатывают слабо коррелированные признаки, а на базе котангенсного – сильно коррелированные.

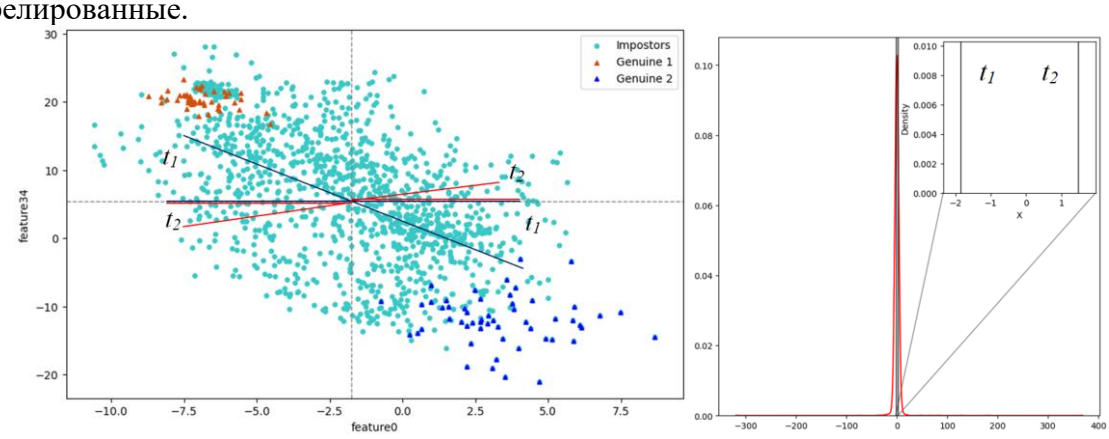

Рисунок 7 – Отображение порогов t_1 и t_2 (-) коррелированном подпространстве пар признаков голосового пароля и плотности вероятности мета-признаков, полученных путем отображения (4) и использования метрики (6).

Тогда *тригонометрический корреляционный нейрон* будет строиться на базе одной из двух предложенных метрик (косинусной или котангенсной), суммируя входные мета-признаки. Принятие решения по сумме входных значений мета-признаков нейрон осуществляет согласно двухуровневой пороговой функции активации $\varphi(y)$:

$$\varphi(y) = \begin{cases} 1, & y \geq T_2 \\ 0, & T_1 < y < T_2, \\ -1, & y \leq T_1 \end{cases} \quad (7)$$

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.

На рис. 8 представлен алгоритм настройки НПБК для расчёта порогов для каждого подпространства пар признаков:

1. Вычисляется коэффициент корреляции пары признаков;
2. Определяется тип мета-признака в соответствии с корреляцией: положительная корреляция – метрика (6) (+), отрицательная – метрика (6) (–), слабая – метрика (5).
3. Строится эмпирическая функция плотности вероятности $f(.)$ мета-признаков, вычисляемых по соответствующей формуле;
4. Функция $f(.)$ интегрируется с целью нахождения функции распределения $F(.)$;
5. Интервал [0, 1] делится на *m* равных секторов ([0,$\frac{1}{m}$,$\frac{2}{m}$,…,1]);

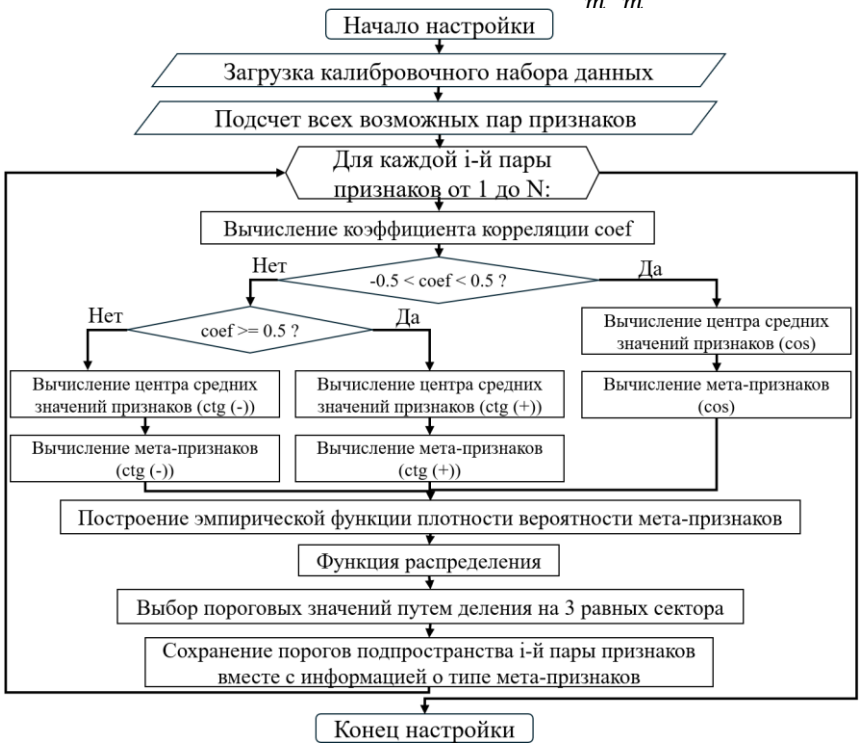

Рисунок 8 – Алгоритм настройки параметров НПБК

Для каждого значения $z \in [0, \text{m-1}]$ из полученного набора высчитывается соответствующее значение функции распределения t_z, для которого $F(t_z) = z$. Полученные значения t_z и являются искомыми порогами. Для каждой пары признаков, рассчитанной с помощью указанного алгоритма, пороги сохраняются, а затем применяются для вычисления порогов функции активации нейрона.

Обучение НПБК осуществляется автоматически без применения метода обратного распространения ошибки. Структура НПБК должна проектироваться из расчета $N = S/q$, где N – количество нейронов, S – желаемая длина криптографического ключа, q – количество бит, продуцируемое одним нейроном и вычисляемое по формуле: $q = [\, log_2 z\,]$, где z – количество порогов, делящее функцию плотности вероятности мета-признаков на равные секторы ($z = m-1$), [.] – операция округления до большего значения.

Для оценки работы НПБК с тригонометрическими нейронами двух типов, был произведен синтез и обучение НПБК для нескольких пользователей. Набор данных был разделён на группы «Чужой», которые использовались для калибровки и «Свой» для обучения НПБК. Для каждого из преобразователей оценивается FRR (вероятность ложного отклонения легального пользователя) и FAR (вероятность ложного пропуска постороннего).

Каждый НПБК обучался при следующих параметрах:

1. Длина ключа 1024 бит;
2. Количество входов нейронов – 4;
3. Количество образов, попадающих в сектор ([-∞; t_1], (t_1; t_2) или [t_2; ∞] на рис. 7) при сборке – 100% (15 образов).

С указанными параметрами производились эксперименты по построению НПБК для косинусных нейронов, котангенсных нейронов, а также в случае их совместного применения. Лучший из полученных результатов составляет EER = 2.1% для AIC-spkr-130 и EER = 3,2% для RedDots, где EER - равная ошибка классификации, определяемая как точка пересечения кривых FAR и FRR. При построении НПБК только на косинусных или котангенсных нейронах значение EER равно 5,5% и 4,4% соответственно.

В четвёртой главе исследуется влияние ФС пользователя на точность биометрической аутентификации. Представленная в 3 главе модель НПБК, которая была настроена по заявленному алгоритму, учитывает дрейф голосовых характеристик под воздействием изменённого состояния для минимизации ошибок в системах распознавания. Эксперименты показывают, что НПБК с тригонометрическими корреляционными нейронами сохраняет стабильность бинарного кода так как при дрейфе положительно коррелированных признаков они изменяются в одном направлении (синхронный сдвиг), как показано на рисунке 9.

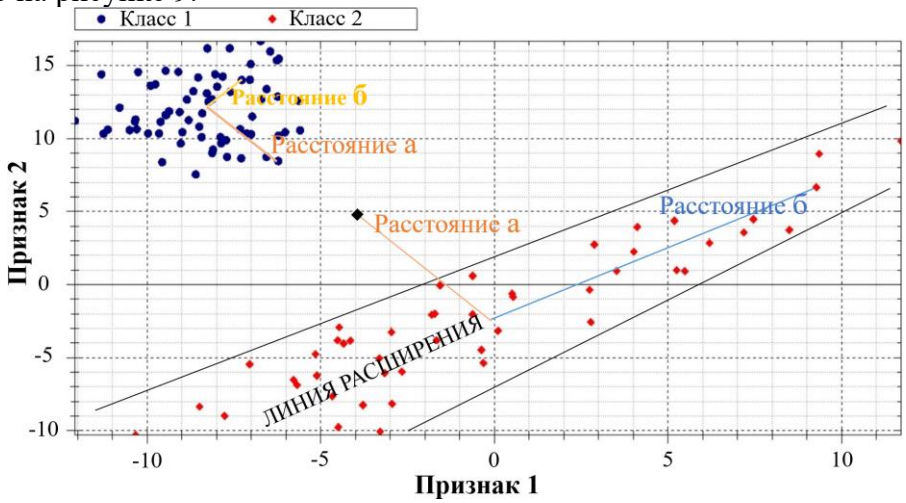

Рисунок 9 – Пространство признаков является «плоским» относительно класса 1 (признаки независимы) и искривленным относительно класса 2 (признаки положительно коррелированы), черная точка – это образ «Чужого»

Для оценки устойчивости НПБК к дрейфу данных проведены тесты с голосовыми сигналами в изменённых состояниях. При переходе к изменённому ФС EER для предложенной модели вырос с 2,1% до 2,71%, что соответствует увеличению ошибок на 29%, в то время как для GMM-UBM speaker verification system, рост достиг 52% (с 8,1 до 12,3%). Такие результаты подчёркивают преимущество тригонометрических корреляционных нейронов, которые фокусируются не только на абсолютных значениях признаков, но и на их взаимосвязях, снижая влияние дрейфа.

Также в данной главе представлена методика аутентификации пользователя по голосу с обеспечением защиты биометрических эталонов и с учётом ФС, включающая уже представленный алгоритм настройки НПБК и разработанный ансамблевый алгоритм МО

для распознавания ФС, классифицирующий пять состояний: нормальное, три градации опьянения и сонное. Также в рамках данной методики были предложены два варианта комплексирования данных алгоритмов для задач аутентификации и генерации ключа электронной цифровой подписи.

Ансамблевый алгоритм на основе CatBoost (метод МО, использующий деревья решений для задачи классификации) достиг наилучшей точности со значением метрики макро-F1 = 0,954 (где F1-score – это среднее между долей правильных положительных ответов и долей истинно положительных случаев, найденных моделью, а макро-F1 – среднее F1-score по всем классам без учёта их сбалансированности) при использовании комбинированных признаков (мел-кепстральных коэффициентов и векторных представлений от предобученных моделей). Таблица 1 содержит данные 5-блочной кросс-валидации, отражающие показатели точности, сбалансированной точности и макро-F1 для моделей MLP, SVM и CatBoost по каждой части, а также средние значения данных метрик.

Таблица 1 – Результаты 5-блочной кросс-валидации и средние значения точности, сбалансированной точности и F1

№ кросс-валидации	MLP accuracy	MLP balanced acc	MLP macro F1	SVM accuracy	SVM balanced acc	SVM macro F1	CatBoost accuracy	CatBoost balanced acc	CatBoost macro F1
1	0.834	0.810	0.810	0.723	0.702	0.704	0.957	0.953	0.950
2	0.844	0.837	0.837	0.714	0.701	0.698	0.954	0.953	0.951
3	0.801	0.781	0.778	0.714	0.694	0.694	0.961	0.957	0.957
4	0.844	0.822	0.824	0.714	0.702	0.698	0.956	0.952	0.954
5	0.835	0.820	0.821	0.717	0.706	0.706	0.951	0.947	0.947
Среднее	0.832	0.814	0.814	0.716	0.700	0.700	**0.956**	**0.952**	**0.954**

Согласно представленной матрице ошибок (рис. 10) для классификатора можно выявить, что состояние «ИФС 2» может быть ошибочно классифицировано как «ИФС 1» или «ИФС 3». А состояние «ИФС 3» может быть ошибочно классифицировано только как «ИФС 1/2», из чего можно сделать вывод, что обученная модель не допустила ошибок в решениях классификатора между такими состояниями как нормальное и «ИФС 2/3».

Рисунок 10 – Матрица ошибок модели CatBoost при классификации ФС

Рассматривается несколько вариантов комплексирования НПБК с моделью классификации ФС, которая распознаёт состояние пользователя. В рамках подобного объединения создаётся система, в которой доступ предоставляется с адаптацией к текущему состоянию пользователя. Данная методика позволяет не только верифицировать пользователя, но и определять уровень его прав в зависимости от выявленного состояния.

На рис. 11 и 12 предложены алгоритмы регулирования уровня доступа в зависимости от состояния пользователя и функционального назначения НПБК (генерация бинарного кода для прохождения аутентификации или закрытого ключа электронной подписи).

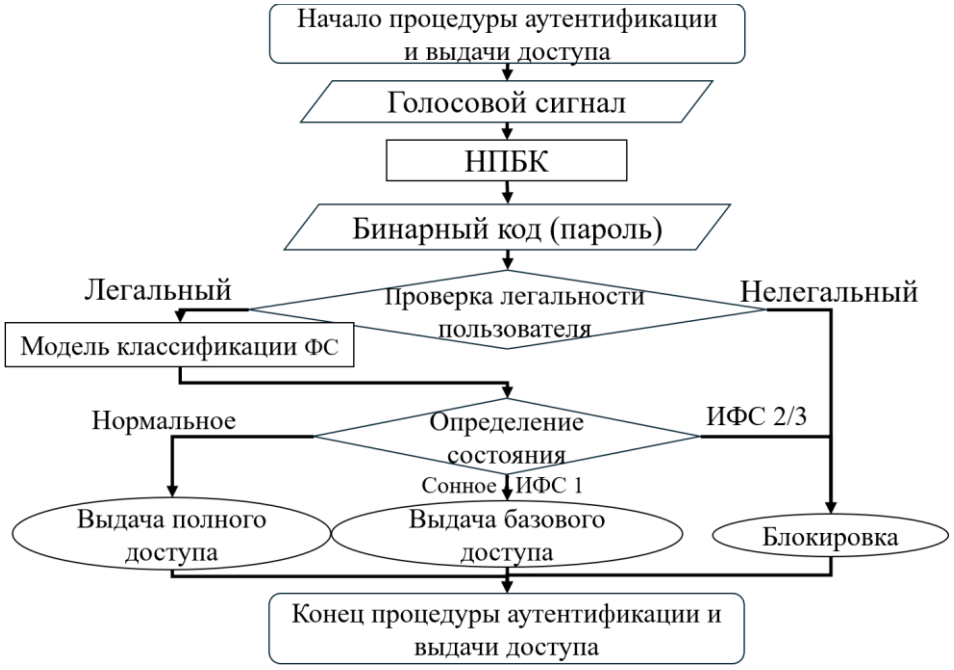

Рисунок 11 – Схема алгоритма комплексирования решения НПБК и модели классификации ФС, обученных для генерации бинарного кода для аутентификации

Рисунок 12 – Схема алгоритма комплексирования решения НПБК и модели классификации ФС, обученных для генерации закрытого ключа электронной подписи

В ходе диссертационной работы был модифицирован программный комплекс AIC Desktop и создана полноценная учебно-исследовательская среда для проектирования, обучения и тестирования нейросетевых моделей, а также для обработки и анализа биометрических данных с сохранением результатов в удобном для дальнейших исследований формате. На основе разработанных методов создан комплекс лабораторных работ, охватывающий задачи машинного обучения, биометрической идентификации, цифровой обработки сигналов, краткое содержание которых представлено в данной главе. Комплекс внедрён в образовательный процесс ФГАОУ ВО «Омский государственный технический университет», что позволило студентам освоить практические навыки работы с биометрическими системами и нейросетевыми алгоритмами.

В заключении подводятся итоги работы, подчеркивается их теоретическая и практическая значимость, приводятся направления дальнейших исследований.

Основные результаты диссертационной работы и выводы:

В рамках диссертационной работы на основе выполненных исследований решена

актуальная научная задача создания биометрической системы аутентификации по голосовым паролям, устойчивой к изменению ФС пользователя, на основе нейросетевых алгоритмов доверенного искусственного интеллекта, имеющая значение для развития методов и систем защиты информации. Разработанные методы обеспечивают адаптивное управление доступом с учётом его ФС, что повышает безопасность и практическую применимость. К основным результатам диссертационной работы относятся следующие теоретические и практические достижения:

1. Разработана статистическая модель, описывающая изменения амплитудно-частотных характеристик речи в различных функциональных состояниях, включая нормальное, сонное и три градации алкогольного опьянения. Модель основана на экспериментальных данных, собранных в ходе экспериментов с голосовыми образами, и учитывает индивидуальные особенности пользователей, проявляющиеся в неоднородности реакций на ФС. Анализ дрейфа признаков позволил выявить области спектра, наиболее чувствительные к изменениям состояния, такие как сдвиги в частотных диапазонах.

2. Предложен метод извлечения признаков из голосовых образов на основе ансамбля автокодировщиков, который осуществляет расширение набора биометрических данных на этапе распознавания. Метод заключается в параллельной обработке различных представлений одного голосового образа, что позволяет формировать более полное и устойчивое описание биометрического шаблона. Ансамблирование СНС увеличивает информативность признаков, снижая чувствительность к шуму и вариациям ФС, что приводит к снижению EER на 5,07%.

3. Разработан алгоритм настройки НПБК, устойчивого к изменению ФС пользователя, за счёт использования тригонометрических корреляционных нейронов, которые оценивают не только значения признаков, но и их взаимосвязи. Это позволяет сохранять стабильность генерируемого бинарного кода при синхронном дрейфе коррелированных характеристик голоса. Эксперименты показали, что в изменённых состояниях (сонное, ИФС 1–3) EER для предложенной модели растёт всего на 29% (с 2,1% до 2,71%), в то время как для GMM-UBM speaker verification system, рост составляет 52% (с 8,1 до 12,3%). Это минимизирует ложные отказы в доступе, повышая надежность аутентификации, в основе которой лежит формируемый алгоритмом НПБК. Кроме того, бинарный код на выходе НПБК может служить не только паролем для аутентификации, но и ключом для электронной подписи, что расширяет применение в защищённых системах.

4. Разработана методика аутентификации по голосу с распознаванием ФС пользователя, включающая предложенный НПБК и ансамблевый алгоритм МО CatBoost, классифицирующий пять состояний (нормальное, ИФС 1, ИФС 2, ИФС 3, сонное) с использованием комбинированных признаков (мел-кепстральные коэффициенты и векторные представления от предобученных моделей). Алгоритм на основе CatBoost достиг макро-F1 0,954, что даёт возможность с высокой точностью определять состояние пользователя. Эксперименты на кросс-валидации подтвердили устойчивость модели и полное отсутствие ошибок в решениях классификатора между такими состояниями как нормальное и ИФС 2/3. Указанная методика предусматривает динамическое управление доступом, уровень которого может меняться в зависимости от ФС пользователя. Объединение данных классификаторов позволяет не только верифицировать пользователя, но и ввести дополнительный контроль, предотвращая пользовательские ошибки в критически важных модулях при выявлении изменённого состояния.

5. Реализован программный модуль в рамках платформы AIC Desktop, реализующий алгоритм настройки гибридной модели НПБК с использованием двух типов тригонометрических корреляционных нейронов. Модуль обеспечивает связку биометрического шаблона с бинарным кодом, и поддерживает защищённый режим исполнения.

Таким образом, в диссертации решены все поставленные задачи, а ее цель, заключающаяся в повышении точности и устойчивости к дрейфу биометрических данных

процедуры аутентификации пользователя по голосу на основе НПБК с обеспечением защиты биометрических шаблонов от компрометации, достигнута.

Результаты работы приняты к внедрению в проектную деятельность ООО «ОСМИ-ИТ» г. Москва, а также применены в учебные процессы ФГАОУ ВО «Омский государственный технический университет».

Перспективы дальнейшей разработки темы включают интеграцию моделей с другими биометрическими модальностями (например, лицо или отпечатки пальцев) для создания мультимодальных систем, устойчивых к комбинированным атакам в том числе распознаванию синтетических образов. Планируются также дополнительные исследования влияния внешних факторов на точность аутентификации и разработка платформы для использования уже полученных наработок в режиме реального времени.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

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

1. 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.

2. Иниватов, Д. П. Аналитическое исследование проблемы биометрической идентификации и аутентификации субъектов по голосу / Д. П. Иниватов // Вопросы защиты информации. – 2023. – № 3(142). – С. 28-38. – DOI 10.52190/2073-2600_2023_3_28.

3. Аутентификация по голосовым паролям с обеспечением конфиденциальности биометрических данных на основе корреляционных нейронов / А. Е. Сулавко, Д. П. Иниватов, В. И. Васильев [и др.] // Информационно-управляющие системы. – 2024. – № 2(129). – С. 21-38. – DOI 10.31799/1684-8853-2024-2-21-38.

4. Иниватов Д.П. Аутентификация по голосовому паролю с учётом функционального состояния пользователя // Вестник УрФО. – 2025. – № 3(57). – С. 61–72.

Статьи, индексируемые в международной базе Scopus:

5. Vasilyev, V.I., Sulavko, A.E., Fofanov, G.A., Inivatov, D.P. Applicability of Classical and Hybrid Neural Network Algorithms in Problems of Recognition of Biometric Patterns // 2018 14th International Scientific-Technical Conference on Actual Problems of Electronic Instrument Engineering, APEIE 2018 - Proceedings. Novosibirsk. p. 563 – 568. Ноябрь 2018.

6. Ivanov A. I., Bannykh A. G., Lozhnikov P. S., Sulavko A.E., and Inivatov D. P. Possibility of Decrease in a Level of Data Correlation During Processing Small Samples Using Neural Networks by Generating New Statistic Tests // Journal of Physics: Conference Series. 2020. Vol. 1546. 012080. DOI: 10.1088/1742-6596/1546/1/012080.

7. Nigrey A. A., Sulavko A. E., Samotuga A. E., Inivatov D. P. On the person and psychophysiological state identification using electroencephalogram parameters // Journal of Physics: Conference Series. 2020. Vol. 1546. 012092. DOI:10.1088/1742-6596/1546/1/012092.

8. Inivatov D. P., Sulavko A. E., Samotuga A. E. The Study of an Impact of Alcohol Intoxication on Voice Parameters in Biometric Authentication Systems //2023 Dynamics of Systems, Mechanisms and Machines (Dynamics). – IEEE, 2023. – С. 1-6.

9. Sulavko A. et al. Biometric-based key generation and user authentication using voice password images and neural fuzzy extractor //Applied System Innovation. 2025.Т. 8. №. 1. С. 13.

Прочие публикации:

10. 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 г., с. 29-47.

11. Особенности построения нейросетевых алгоритмов в задачах распознавания образов / Чобан А.Г., Стадников Д.Г., Иниватов Д.П. [и др.] // Информационная безопасность: современная теория и практика: сборник научных трудов студентов, аспирантов и

преподавателей по материалам II Межвузовской научно-практической конференции. СиБАДИ. г. Омск, 13 сентября 2019 г., С. 127 – 133.

12. Иниватов Д.П., Стадников Д.Г., Чобан А.Г. Определение субъекта по биометрическим параметрам // Современная мировая экономика: проблемы и перспективы в эпоху развития цифровых технологий и биотехнологии // Сборник научных статей по итогам работы девятого международного круглого стола. 15-16 декабря 2019 г. Часть 1 - Москва: ООО «Конверт», 2019. С. 75 – 78.

13. Иниватов, Д. П. Аналитический обзор достигнутых результатов в области идентификации личности с обеспечением защиты биометрических эталонов от компрометации / Д. П. Иниватов, А. Е. Сулавко // Прикладная математика и фундаментальная информатика. – 2021. – Т. 8. – № 2. – С. 29-37.

14. Иниватов, Д. П. «Комитет нейронных сетей» в свете «теоремы жюри Кондорсе» / Д. П. Иниватов // Россия и мировые тенденции развития: Материалы Всероссийской научно-практической конференции с международным участием, Омск, 13–15 мая 2021 года. – Омск: Омский государственный технический университет, 2021. – С. 332-339.

15. Иниватов Д.П. Распознавание субъекта на основе комитета искусственных нейронных сетей // Цифровизация и кибербезопасность: современная теория и практика : сборник научных трудов по материалам II Международной научно-практической конференции, Омск, 20 – 21 октября 2022 года. – Омск: Сибирский государственный автомобильно-дорожный университет (СибАДИ), 2022. – С. 137-140.

16. Иниватов Д.П. Защищенное исполнение процедур биометрической аутентификации дикторов в цифровой среде // Фундаментальные и прикладные исследования молодых учёных : сборник материалов VII Международной научно-практической конференции студентов, аспирантов и молодых учёных, приуроченной к 110-летию со дня рождения Т.В. Алексеевой, Омск, 20–21 апреля 2023 года. – Омск: Сибирский государственный автомобильно-дорожный университет (СибАДИ), 2023. – С. 571-575.

17. Панфилова, И. Е., Иниватов Д.П. Обзор методов защиты данных биометрических шаблонов / Безопасность информационных технологий: Сборник научных статей по материалам V Всероссийской научно-технической конференции, посвященной 70-летнему юбилею АО "НПП "Рубин". В 2-х томах, Пенза, 27 сентября 2023 года. – Пенза: Пензенский государственный университет, 2023. – С. 135-146.

Свидетельства о государственной регистрации программ для ЭВМ

18. Свидетельство о государственной регистрации программы для ЭВМ № 2021660512 Российская Федерация. AIC desktop : № 2021617236 : заявл. 17.05.2021: опубл. 28.06.2021 / А. Е. Сулавко, Д. Г. Стадников, А. Г. Чобан, Д. П. Иниватов.

19. Свидетельство о государственной регистрации программы для ЭВМ № 2021617987. Программный комплекс, обрабатывающий звуковые файлы для последующего анализа нейронными сетями особенностей голоса диктора: № 2021617379: заявл. 21.05.2021: опубл. 21.05.2021 / Д. П. Иниватов.

20. Свидетельство о государственной регистрации программы для ЭВМ № 2022661737. Программный комплекс, осуществляющий преобразования базы звуковых образов Reddots: № 2022660860: заявл. 10.06.2022: опубл. 24.06.2022 / Д. П. Иниватов.

21. Свидетельство о государственной регистрации программы для ЭВМ № 2023661175. Программный комплекс, осуществляющий обезличивание биометрических образов: № 2023660561: заявл. 25.05.2023: опубл. 29.05.2023 / Д. П. Иниватов.

Дисссертант Иниватов Д.П.