

АУТЕНТИФИКАЦИЯ ПО ГОЛОСОВОМУ ПАРОЛЮ С УЧЁТОМ ФУНКЦИОНАЛЬНОГО СОСТОЯНИЯ ПОЛЬЗОВАТЕЛЯ¹

В статье представлена методика биометрической аутентификации по голосу, учитывающая функциональное состояние (ФС) пользователя, такое как нормальное состояние, различные стадии алкогольной интоксикации и сонливость. Предложенный подход объединяет нейросетевой преобразователь «биометрия-код» (НПБК) с моделью классификации ФС на основе ансамблевого алгоритма CatBoost, обеспечивая высокую точность (макро-F1 0.954) в определении состояния пользователя. Методика позволяет не только подтвердить личность, но и ограничить доступ в системе в случае выявления неприемлемого ФС при наиболее важных действиях.

Ключевые слова: голосовая аутентификация, функциональное состояние, защита данных, НПБК, CatBoost, биометрическая аутентификация, алкогольная интоксикация.

Inivatov D. P.

VOICE PASSWORD AUTHENTICATION WITH CONSIDERATION OF USER'S FUNCTIONAL STATE

The article presents a voice-based biometric authentication method that accounts for the user's functional state, including normal state, various stages of alcohol intoxication, and drowsiness. The proposed approach integrates a neural fuzzy extractor with a functional state classification model based on the CatBoost ensemble algorithm, achieving high accuracy (macro-F1 0.954) in state detection. The method not only verifies user identity but also dynamically restricts access based on functional state, minimizing risks of errors in critical systems.

Keywords: voice authentication, functional state, neural fuzzy extractor, CatBoost, biometric authentication, alcohol intoxication, data security.

Введение

Голосовая биометрия используется повсеместно: в системах аутентификации банков, в кол-центрах, умных устройствах. Современные биометрические системы, основанные на голосе, уже достигают значений точности в 98-99% [1] для задач аутентификации пользователя, предполагая, что речевой образ остаётся стабильным. Данный прогресс позволил переориентировать исследования с повышения базовой точности на решение задач, связанных с вариативностью биометрических данных, среди которых функциональное состояние (ФС) пользователя является одним из основных факторов.

ФС, включая сонливость, стресс, утомление и алкогольную интоксикацию, вызывают значительные изменения в амплитудно-частотных характеристиках голоса [2], что не только снижает точность аутентификации, но и создает побочные угрозы. Выдача доступа пользователю в неприемлемом ФС, таком как сильное опьянение или крайняя усталость, может привести к компрометации данных. В банковской системе сотрудник в подобном состоянии может инициировать ошибочные транзакции, повредив финансовые записи, или стать мишенью для фишинговой атаки, открыв доступ к защищаемой информации. Не все отклонения от нормального ФС представляют угрозу, но состояния, влияющие на когнитивные и моторные функции, требуют строгого контроля.

Для минимизации таких рисков традиционная аутентификация, подтверждающая только легальность пользователя, недостаточна. Необходимо внедрять методы, оценивающие ФС. Подобные решения позволяют не только верифицировать личность, но и предотвратить угрозу целостности данных от действий в нежелательных состояниях.

Обзор существующих методов определения изменённого функционального состояния пользователя

Методы машинного обучения (МО) находят широкое применение в системах голосовой аутентификации, позволяя анализировать такие характеристики речевых сигналов, как тембр, высота тона, ритмические особенности, паузы и др. Такие подходы как машина опорных векторов (SVM), нейронные сети (НС), модели Гауссовой смеси и ансамблевые методы, способны определять схожие при-

знаки и проводить классификацию пользователей по соответствующим категориям [3].

Для анализа голосовых данных в задачах классификации речевой сигнал сначала проходит этап предобработки, направленный на устранение шумов и нормализацию, после чего из него выделяются признаки, такие как мел-кепстральные коэффициенты или спектральные характеристики, которые затем обрабатываются алгоритмами МО [4]. Задача оценки ФС может быть сформулирована как бинарная классификация с двумя категориями – нормальным и изменённым функциональным состоянием, либо как многоклассовая, в которой каждый класс отражает своё состояние. Для оценки эффективности предлагаемых решений исследователи активно применяют метрики UAR (невзвешенная средняя точность), которая учитывает сбалансированность классов, и ассигасу (доля пра работу системы проверки состояния водителей для предупреждения неприемлемых ФС за рулём. В исследовании используется сразу 2 базы данных: ALC для классификации состояния алкогольной интоксикации и BESD для распознавания эмоций (нейтральное, гнев, радость, грусть, страх, отвращение). Признаки извлекаются с помощью мел-частотных кепстральных коэффициентов (MFCC), линейных предсказательных коэффициентов и частоты пересечения нуля для ALC, а также высоты тона, энергии и просодических характеристик для определения эмоций. Полученные признаки объединяются, их обработка осуществляется с помощью алгоритмов МО: SVM, метод k-ближайших соседей, метод случайного леса, градиентный бустинг и метод экстремально случайных деревьев, с применением анализа главных компонент и селекции признаков.

Предложенное авторами решение позволило достичь точности в 80% для выявления состояния алкогольного опьянения, 94,7% в подтверждении его отсутствия, 75% для эмоции «радость», 66,7% для «нейтрального» состояния, 72,41% для «гнева» и 57,77% «грусти». Объединение признаков улучшило точность по сравнению с использованием отдельных наборов данных, что делает метод перспективным для определения ФС.

В исследовании [10] представлен алгоритм ADLAI, использующий глубокое обучение для определения алкогольного опьянения по 12-секундным речевым записям. Метод основан на анализе акустических призна-

ков речи, таких как тембр и ритм, с использованием свёрточных нейронных сетей для классификации. Тестирование на корпусе ALC показало значение по метрике UAR = 68.09% для выявления изменённого ФС при концентрации алкоголя в крови 0.5‰ и 75.7% для 0.12‰, что является лучшим результатом на ALC среди аналогичных методов по метрике UAR.

Исследование [11] посвящено методам определения изменённого ФС на примере корпуса ALC. Базовая система использует набор ComParE (6373 акустических признака, таких как основная частота F0, спектральные параметры и др.) и SVM, достигнув значения UAR 0.616 без нормализации. Нормализация по испытуемым повысила результат до 0.669, а кластерная нормализация до 0.665. Модель ResNet18 (18-слойная сеть с остаточными связями, где свёрточные слои выделяют признаки, слой глобального усреднения GAP формирует фиксированное представление, а два полносвязных слоя классифицируют данные) без нормализации достигла UAR 0.633. Z-Нормализация позволила повысить значение метрики до 0.677 (на 4.4 %), а кластерная нормализация до 0.671.

Можно сделать вывод, что с 2011 года, после публикации проекта ALC в открытом доступе, в научной среде произошло заметное развитие исследований по распознаванию ФС по голосу, что стало возможным благодаря обширному набору речевых данных. Этот ресурс послужил основой для создания и проверки множества методов, способствуя прогрессу в данной области.

Набор данных для экспериментальных исследований

В рамках одного из предыдущих исследований [2] был сформирован свой архив голосовых записей ALC-spkr-130, ориентированный на задачу аутентификации по голосовому паролю с защитой биометрического шаблона. Предлагаемый архив состоит из записей 130 испытуемых возраста от 18 до 50 лет в формате wav и частотой дискретизации 8 кГц. Для сбора архива голосовых данных было отобрано 130 участников, не страдающих серьёзными заболеваниями или неврологическими расстройствами. Запись проводилась в утренние часы после достаточного отдыха испытуемых, чтобы минимизировать влияние усталости. Участники по-

Таблица 1

Сравнение эффективности методов определения изменённого ФС

Описание архитектуры	Точность, %	UAR, %	База данных
Архитектура ResNeXt50 и мел-спектрограммы [5]		59,2	ALC
Метод адаптивной обратной фильтрации и модель Лильенкранца-Фанта [7]	69,3 – 77,0		Собственная
Метод случайного леса (10 деревьев) [8]	95,3		Собственная
Метод опорных векторов и метод случайного леса [9]	57,7 – 94,7		BESD, ALC
Алгоритм глубокого обучения на основе аудио для выявления алкогольного опьянения (ADLAIA) [10]		68,1 – 75,7	ALC
ResNet-18 + глобальный средний уровень пула + z-нормализация [11]		67,1	ALC
Система супервекторов модели Гауссовой смеси, объединенная с линейной машиной опорных векторов для классификации и дополнительными методами нормализации [12]		70,54	ALC
Модели Гауссовой смеси + линейная машина опорных векторов [13]	68,6	68,5	ALC
Двунаправленная рекуррентная нейронная сеть со стробируемыми рекуррентными единицами [14]	75,9	69,2	ALC

следовательно вводились в разные функциональные состояния, в каждом из которых повторяли >60 голосовые пароли, представляющие из себя фразы из 2 слов на русском языке.

Собранная база голосовых образов содержит записи испытуемых в 5 различных функциональных состояниях:

1. Нормальное состояние. В этом состоянии испытуемый не подвергался каким-либо внешним воздействиям.

2. Градации стадий алкогольной интоксикации. В соответствии с классификацией стадий алкогольного опьянения, приведённой в методических рекомендациях Минздрава СССР [15] выделяются 3 стадии, каждая из которых характеризует определённое функциональное состояние:

2.1. Состояние незначительной алкогольной интоксикации (интоксикация 1), при котором концентрация алкоголя в крови находится в диапазоне от 0.3‰ до 0.5‰. Она характеризуется отсутствием выраженных клинических проявлений воздействия алкоголя.

2.2. Лёгкое опьянение (интоксикация 2), при котором концентрация алкоголя находится в диапазоне от 0.5 до 1.5‰. Данное содержание токсина воздействует на когнитивные и моторные функции, способность к принятию решений, замедляет реакцию. Также в данном ФС наблюдается ощущение тепла, комфортного состояния, мышечное расслабление, понижение сдержанности. В рамках проводимого эксперимента у испытуемых данное значение составляло около 0,8‰.

2.3. Третья стадия (интоксикация 3). Концентрация алкоголя находится в диапазоне от 1.5 до 2.5‰. Характеризуется нарушением моторных функций, способности к рассуждению, нечеткой речью, реакцией зрачка на свет и эмоциональной неустойчивостью. В рамках эксперимента уровень ограничивался значением порядка 1,6‰.

3. Сонное. Это состояние характеризуется снижением общей активности и когнитивной продуктивности. Для его воспроизведения участники использовали натуральные растительные седативные средства.

К формированию базы голосовых образов привлечены русскоговорящие добровольцы без неврологических или речевых нарушений.

Архитектура нейросетевой модели для распознавания ФС пользователя

Эксперимент проводился AIC-spkr-130. Голосовые данные представляли собой сырые аудиосигналы с приведением частоты дискретизации к 16000 Гц. Предобработка включала удаление самых зашумлённых и нечётких записей, а также нормализацию амплитуды в диапазоне [-1, 1]. Извлечение признаков выполнялось комбинированным подходом: MFCC через библиотеку openSMILE захватывали спектральные характеристики речи, такие как формантные частоты, а высокоуровневые представления извлекались тремя предобученными моделями: Wav2Vec2 XLS-R, HuBERT-large и ECAPA-TDNN.

Традиционные мел-кепстральные коэффициенты обеспечивали базовое описание амплитудно-частотных свойств речи. Признаки, извлечённые моделями Wav2Vec2 XLS-R и HuBERT, формировались путём агрегации выходов последнего скрытого слоя с использованием усреднения и стандартного отклонения. Модели XLS-R и HuBERT формировали 2048-мерные вектора признаков, а ECAPA-TDNN - 192-мерные вектора. Итоговый набор получался путём объединения этих 3 наборов и содержал 4376 признаков.

Для классификации ФС были сформированы и обучены три модели. Первой был многослойный персептрон (MLP) [16], состоящий из входного слоя на 4376 признаков, трёх полносвязных слоёв с 512, 256 и 128 нейронами для постепенного уменьшения размерности и выделения ключевых признаков. В качестве функции активации применялась ReLu после каждого слоя. Оптимизация проводилась с помощью Adam с начальной скоростью обучения 0.0001. Обучение продолжалось 72 эпохи при размере батча 256. Параметры нейронной сети приведены в таблице 2.

В качестве ядра SVM использовались радиальные базисные функции (RBF). Для обучения и тестирования применялись только мел-кепстральные коэффициенты. Параметр регуляризации, отвечающий за величину штрафа за ошибки в классификации, был настроен на 10. Автоматически адаптируемый параметр ширины ядра определялся как функция дисперсии данных. Основные параметры SVM приведены в таблице 3.

Для обучения и тестирования модели CatBoost [17] (метод МО, использующий дере-

Таблица 2

Параметры модели MLP

Параметр	Значение
Входной слой	4376 признаков
Скрытые слои	512, 256, 128 нейронов
Dropout	0.3, 0.3, 0.2 соответственно
Функция активации	ReLU
Оптимизатор	Adam
Число эпох в обучении	72
Размер батча	256

Таблица 3

Параметры модели SVM

Параметр	Значение
Тип ядра	RBF
Признаки	MFCC
Регуляризация	10
Параметр ширины ядра	Адаптивный (по дисперсии данных)

Таблица 4

Параметры ансамблевой модели MO CatBoost

Параметр	Значение
Число признаков	4376
Функция потерь	Многоклассовая классификация
Глубина дерева	8
Количество итераций	3000
Скорость обучения	0.1
Бутстрэппинг	Байесовский (регуляризация 0.7)
Балансировка весов классов	Автоматическая
Начальное значение	Фиксированное

вья решений для задачи классификации), уже показавшей свою эффективность в задаче распознавания возраста и пола по голосу, использовался весь набор из 4376 признаков, включая мел-кепстральные коэффициенты и высокоуровневые представления. В качестве целевой функции при обучении использовалась многоклассовая функция потерь. Глубина дерева 8 и Байесовский бутстрэппинг

(bootstrapping) с параметром регуляризации 0.7 были подобраны опытным путём для избегания переобучения. А количество итераций 3000 и скорость обучения 0.1 позволило модели в достаточной степени обучиться. Автоматическая балансировка весов классов увеличивала вес наименьшего из классов (интоксикация 2). Значения параметров указаны в таблице 4.

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

Таблица 6

Средние метрики по классам для MLP и CatBoost

Класс	MLP precision	MLP recall	MLP F1	CatBoost precision	CatBoost recall	CatBoost F1
Инттоксикация 1	0.76	0.81	0.79	0.96	0.94	0.95
Инттоксикация 2	0.59	0.68	0.63	0.93	0.92	0.93
Инттоксикация 3	0.91	0.75	0.82	0.95	0.96	0.95
Нормальное	0.96	0.89	0.93	0.97	0.97	0.97
Сонное	0.87	0.93	0.90	0.96	0.97	0.97

Эксперимент проводился с использованием 5-блочной кросс-валидации, где данные разделялись на обучающую и тестовую выборки в соотношении 80/20 с сохранением пропорций классов (нормальное, интоксикация 1–3, сонное). Кросс-валидация обеспечивала оценку устойчивости моделей путём разделения очищенного набора данных на пять подмножеств, где каждое подмножество использовалось как тестовая выборка один раз в то время, как оставшиеся четыре – для обучения. Результаты всех 5 кросс-валидационных экспериментов и их усреднённые значения представлены в таблице 5. Ансамблевый алгоритм MO CatBoost оказался наиболее эффективным в задаче распознавания ФС: средняя точность 0,956, сбалансированная точность 0,952, макро-F1 0,954. MLP уступала с точностью 0.826, сбалансированной 0.818 и макро-F1 0,814. SVM, ограниченная только MFCC признаками, имела самые низкие показатели: точность 0,716, сбалансированная 0,7, макро-F1 0,7.

Детальный анализ по классам (табл. 6) показал, что CatBoost стабильно превосходит другие модели, особенно для интоксикации 1–3 (F1 0.92–0.96) и нормального состояния

(F1 0.97). MLP хорошо справлялась с нормальным и сонным состояниями (F1 0.89–0.95), но слабее для интоксикации 2 (F1 0.56–0.73). SVM показала умеренные результаты (F1 0.54–0.84), с наибольшими трудностями для интоксикации 2.

Для визуализации представлена гистограмма данных результатов (Рис. 2), где CatBoost демонстрирует стабильное превосходство.

Матрицы ошибок (Рис. 3–5) иллюстрируют распределение предсказаний: для CatBoost (Рис. 5) ошибки классификации минимальны, с высокой точностью для всех классов, тогда как MLP (Рис. 3) показывает заметное число ошибок между интоксикацией 1 и 2, а SVM (Рис. 4) – между всеми стадиями опьянения. По ошибочным решениям всех 3 классификаторов можно выявить общую закономерность. Она заключается в том, что состояние «интоксикация 2» может быть ошибочно классифицировано как «интоксикация 1» или «интоксикация 3». А состояние «интоксикация 3» может быть ошибочно классифицировано с наибольшей вероятностью как «интоксикация 1» или «интоксикация 2». Также можно выявить, что ансамблевый алго-

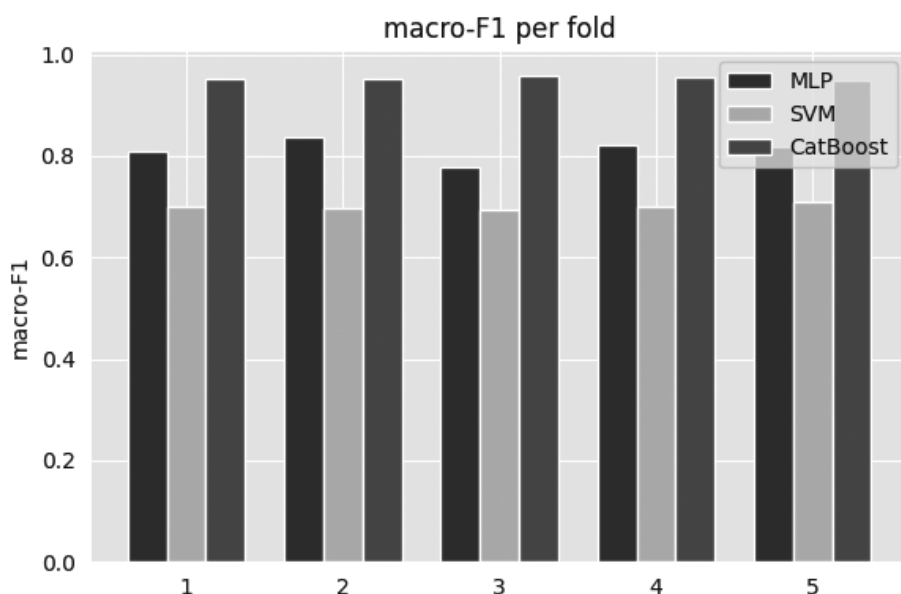


Рис. 2. Гистограмма метрики макро-F1 по 3 моделям 5 кросс-валидационных экспериментов

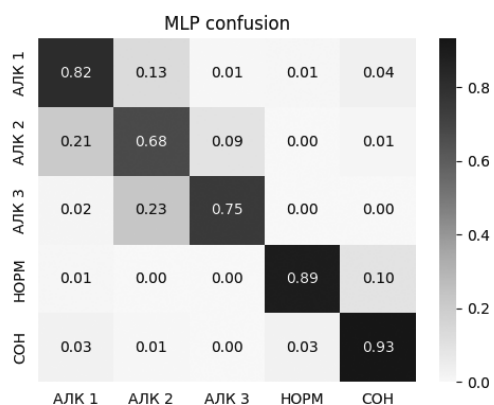


Рис. 3. Матрица ошибок модели MLP при классификации ФС

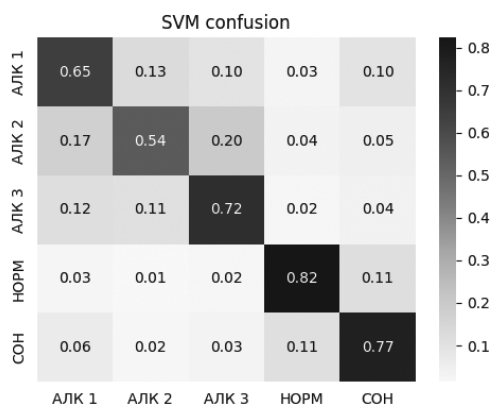


Рис. 4. Матрица ошибок модели SVM при классификации ФС

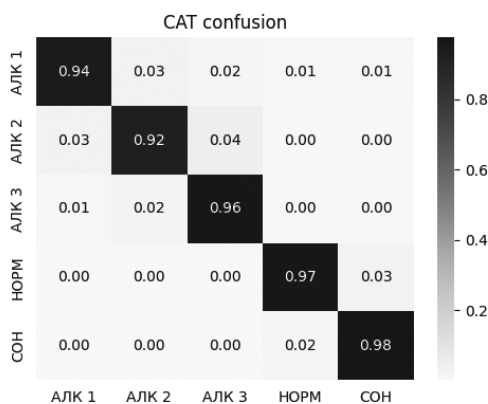


Рис. 5. Матрица ошибок модели CatBoost при классификации ФС

ритм МО CatBoost не допустил ошибок в решениях классификатора между такими состояниями как нормальное и интоксикация 2/3.

Исходя из полученных результатов, можно сделать вывод, что наиболее эффективным оказался подход с использованием объединённого набора признаков (мелкепстральные коэффициенты и высокоуровневые представления) и модели CatBoost, который позволил добиться результата 0.954 по метрике макро-F1.

Для анализа эффективности разработанной модели проведём сравнение с лучшими работами из уже проведённого литературного обзора (таблица 1). В работе [8] на собственной базе данных (голоса мужчин 22–34 лет) с методом случайного леса достигнута точность 95,3% для бинарной классификации (нормальное состояния или алкогольное опьянения) без градаций опьянения или сонного состояния. В [9] на архиве ALC с опорными векторами и случайным лесом получена точность 80%, также для бинарной задачи. Несмотря на различия в используемых базах голосовых образов (собственная в [8], ALC в [9], в настоящей работе – ALC-spkr-130), предложенная в данной работе модель обеспечивает точность 95,6% (сбалансированная точность 95,2%), но с учётом трёх стадий опьянения и сонного состояния, что повышает её возможности и способы применения.

Методика аутентификации пользователя по голосу с учётом его функционального состояния

Нейросетевой преобразователь «биометрия-код» (НПБК) представляет собой технологию, которая преобразует биометрический образ пользователя в бинарный код, используемый для аутентификации или генерации криптографических ключей, обеспечивая при этом защиту исходных биометрических данных от компрометации. В настоящей работе в качестве классификатора, осуществляющего подтверждение личности пользователя, предлагается применить уже апробированный НПБК, предложенный в [9], где используются два типа тригонометрических корреляционных нейронов для обработки голосовых образов, что позволяет генерировать стабильные ключи длиной 1024 бита с низким уровнем ошибок (EER = 2,1%).

НПБК рассматривается как технология, способная превращать биометрический образ в бинарный код. Этот код не является

простым решением о доступе (True или False), а представляет собой последовательность бит, которая может служить либо паролем для аутентификации, либо основой для генерации закрытого ключа электронной подписи. Объединяя НПБК с моделью классификации ФС, которая классифицирует состояние пользователя по пяти категориям, мы создаём систему, в которой доступ предоставляется с адаптацией к текущему состоянию пользователя. Предлагаемая методика аутентификации по голосу, основанная на ансамблевых методах МО, позволяет проводить не только верификацию личности, но и выполнять классификацию ФС пользователя, позволяя учитывать его при авторизации. Основная особенность данной методики заключается в возможности ограничения прав доступа пользователя при выявлении у него изменённого функционального состояния для минимизации рисков при работе с наиболее важными данными.

Основная идея комплексирования заключается в последовательном алгоритме: сначала голосовой сигнал поступает в НПБК для генерации бинарного кода и проверки легальности пользователя. Если пользователь признан легальным, сигнал обрабатывается моделью ФС для определения текущего состояния. Если выявленное состояние считается приемлемым (например, нормальное или интоксикация 1), код используется для аутентификации или служит для генерации ключевой пары закрытого ключа. В случае критического состояния, такого как интоксикация 3, система может ограничить или отложить генерацию кода, предлагая альтернативные меры, включая повторную аутентификацию через определённое время или уведомление администратора.

Подобное объединение может быть полезно в системах с дискреционным или мандатным доступом, где права пользователя динамически корректируются. Например, в корпоративной сети при нормальном состоянии пользователь получает полный доступ: чтение, запись, подпись документов. При интоксикации 2 или 3 доступ ограничивается чтением и просмотром, но запись или подпись блокируются, чтобы избежать ошибок под влиянием интоксикации. В сонном состоянии система может предоставить базовый доступ (ограничивающий действия с персональными данными, конфиденциальной информацией и выполнение финансовых или



Рис. 6. Схема алгоритма комплексирования решения НПБК и модели классификации ФС, обученных для генерации бинарного кода для аутентификации

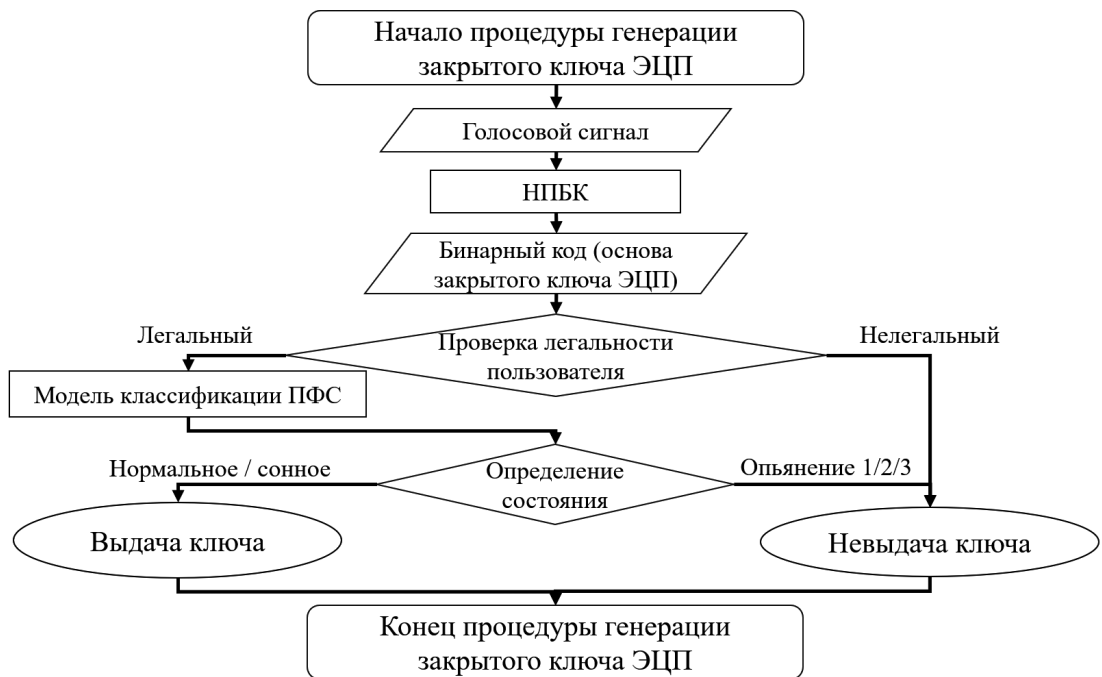


Рис. 7. Схема алгоритма комплексирования решения НПБК и модели классификации ФС, обученных для генерации закрытого ключа электронной подписи

административных операций), но полный только через какое-то число минут (время на "просыпание"), с автоматическим повторным запросом голосового образца. В банковских приложениях, где используется электронная подпись, голос генерирует закрытый ключ только в нормальном состоянии. Подобное ограничение позволит не совершать действий с повышенной ответственностью, например, взятие кредита на крупную сумму удалённо. Похожий способ применения распознавания ФС может быть полезен для приложений по заказу такси. Водитель, перед "выходом на линию" проходит верификацию личности и состояния машины через фото-подтверждение. Прохождение данной верификации с помощью голосового образа позволит не только убедиться в том, что за рулём сам водитель, но и определить его ФС, что очень важно. Такое применение биометрической системы добавляет дополнительный слой безопасности, делая биометрию не только идентификатором личности, но и состояния. На рисунках 6 и 7 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).

В данной работе приведены только основные способы объединения данных классификаторов. Варианты их совместного использования могут зависеть от специфики применения и прав, предоставляемых поль-

зователю. Так, в информационной системе для образовательных целей, направленной на получение знаний и освоение навыков, может и вовсе не требоваться проверка на ФС, а в оборонных системах, с их мандатным контролем, оно потребует дополнительных механизмов, таких как временные блокировки или логирование попыток в изменённом ФС.

Заключение

В результате настоящего исследования разработана методика аутентификации пользователя по голосу, учитывающая его ФС при выдаче прав доступа, что существенно повышает надёжность и безопасность систем аутентификации по голосу. Предлагаемая методика сочетает нейросетевой преобразователь «биометрия-код» с моделью классификации ФС, что позволяет не только подтвердить личность пользователя, но и адаптировать его права в зависимости от выявленного состояния для предотвращения пользовательских ошибок в критически важных модулях. Разработанная модель классификации ФС, основанная на ансамблевом алгоритме МО CatBoost, достигла высокой точности (метрика макро-F1 0.954), что позволяет различать сразу несколько уровней алкогольной интоксикации, а также нормальное и сонное состояния. Интеграция НПБК с моделью классификации ФС даёт возможность генерировать бинарный код для аутентификации или криптографического ключа, с динамическим ограничением прав в зависимости от состояния.

Литература

1. 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.
2. 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. [и др.] // Международное издание «Компьютерные инструменты в образовании», г. Санкт-Петербург, 2017 г., с. 29-47.
3. Siddiqui N., Pryor L., Dave R. User authentication schemes using machine learning methods—a review //Proceedings of International Conference on Communication and Computational Technologies: ICCCT 2021. – Springer Singapore, 2021. – pp. 703-723.
4. Сулавко А.Е., Иниватов Д.П., Стадников Д.Г., Чобан А.Г. Преобразователь образов голосовых паролей дикторов в криптографический ключ на основе комитета предварительно обученных сверточных нейронных сетей // Вопросы защиты информации. – 2021. - №4. – С. 23-33.
5. Rezvani S. Intoxication detection from audio using deep learning //Bachelor Thesis. – 2019.
6. Schiel F., Heinrich C., Barfusser S. Alcohol language corpus: the first public corpus of alcoholized German speech //Language resources and evaluation. – 2012. – Т. 46. – №. 3. – С. 503-521.
7. Sigmund M., Zelinka P. Analysis of voiced speech excitation due to alcohol intoxication //Information Technology and Control. – 2011. – vol. 40. – no. 2. – pp. 143-150.
8. Terlapu P. V. Intelligent Novel Approach for Identification of Alcohol Consumers using Incremental Hidden Layer Neurons ANN (IHLN-ANN)-Based Model on Vowelized Voice Dataset. – 2023.
9. Shenoi V. V., Kuchibhotla S., Kotturu P. An efficient state detection of a person by fusion of acoustic and alcoholic features using various classification algorithms //International Journal of Speech Technology. – 2020. – Т. 23. – pp. 625-632.
10. Bonela A. A. et al. Audio-based Deep Learning Algorithm to Identify Alcohol Inebriation (ADLAIA) //Alcohol. – 2023. – Т. 109. – С. 49-54.
11. Wang W., Wu H., Li M. Deep neural networks with batch speaker normalization for intoxicated speech detection //2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC). – IEEE, 2019. – С. 1323-1327.
12. Bone D. et al. Intoxicated speech detection by fusion of speaker normalized hierarchical features and GMM supervectors // Twelfth Annual Conference of the International Speech Communication Association. – 2011.
13. Bocklet T., Riedhammer K., Nöth E. Drink and Speak: On the automatic classification of alcohol intoxication by acoustic, prosodic and text-based features //Twelfth Annual Conference of the International Speech Communication Association. 2011.
14. Berninger K., Hoppe J., Milde B. Classification of speaker intoxication using a bidirectional recurrent neural network //Text, Speech, and Dialogue: 19th International Conference, TSD 2016, Brno, Czech Republic, September 12-16, 2016, Proceedings 19. – Springer International Publishing, 2016. – pp. 435-442.
15. Методические указания Министерства здравоохранения СССР от 03.07.1974 «О судебно-медицинской диагностике смертельных отравлений этиловым алкоголем и допускаемых при этом ошибок» [Электронный ресурс] // Законодательство в области судебно-медицинской экспертизы. 1974. URL: http://zakon.forens-med.ru/doc/mz/3_25_101.html (дата обращения: 05.10.2025).
16. Частикова В. А., Жерлицын С. А., Войлова Д. О. Нейросетевая система биометрической идентификации личности по голосу //Вестник Адыгейского государственного университета. Серия 4: Естественно-математические и технические науки. – 2023. – №. 1 (316). – С. 70-79.
17. Badr A. A., Abdul-Hassan A. K. CatBoost Machine Learning Based Feature Selection for Age and Gender Recognition in Short Speech Utterances //International Journal of Intelligent Engineering & Systems. – 2021. – Т. 14. – №. 3.
18. Sulavko A. et al. Biometric-based key generation and user authentication using voice password images and neural fuzzy extractor // Applied System Innovation. 2025.T. 8. №. 1. С. 13.

References

1. Inivatov, D. P. Analiticheskoe issledovanie problemy biometricheskoy identifikatsii i autentifikatsii sub#ektov po golosu / D. P. Inivatov // Voprosy zashhity informatsii. – 2023. – № 3(142). – С. 28-38. – DOI 10.52190/2073-2600_2023_3_28.

2. Vliyanie psihofiziologicheskogo sostojaniya diktora na parametry ego golosa i rezul'taty biometricheskoy autentifikacii po rechevomu parolju / A.E. Sulavko, A.V. Eremenko, R.V. Borisov, D.P. [i dr.] // Mezhdunarodnoe izdanie «Komp'juternye instrumenty v obrazovanii», g. Sankt-Peterburg, 2017 g., s. 29-47.
3. Siddiqui N., Pryor L., Dave R. User authentication schemes using machine learning methods—a review // Proceedings of International Conference on Communication and Computational Technologies: ICCCT 2021. – Springer Singapore, 2021. – pp. 703-723.
4. Sulavko A.E., Inivatov D.P., Stadnikov D.G., Choban A.G. Preobrazovatel' obrazov golosovyh parolej diktovor v kriptograficheskij ključ na osnove komiteta predvaritel'no obuchennyh svertochnyh nejronnyh setej // Voprosy zashhity informacii. – 2021. – №4. – S. 23-33.
5. Rezvani S. Intoxication detection from audio using deep learning // Bachelor Thesis. – 2019.
6. Schiel F., Heinrich C., Barfüsser S. Alcohol language corpus: the first public corpus of alcoholized German speech // Language resources and evaluation. – 2012. – T. 46. – №. 3. – S. 503-521.
7. Sigmund M., Zelinka P. Analysis of voiced speech excitation due to alcohol intoxication // Information Technology and Control. – 2011. – vol. 40. – no. 2. – pp. 143-150.
8. Terlapu P. V. Intelligent Novel Approach for Identification of Alcohol Consumers using Incremental Hidden Layer Neurons ANN (IHLN-ANN)-Based Model on Vowelized Voice Dataset. – 2023.
9. Sheno V. V., Kuchibhotla S., Kotturu P. An efficient state detection of a person by fusion of acoustic and alcoholic features using various classification algorithms // International Journal of Speech Technology. – 2020. – T. 23. – pp. 625-632.
10. Bonela A. A. et al. Audio-based Deep Learning Algorithm to Identify Alcohol Inebriation (ADLAIA) // Alcohol. – 2023. – T. 109. – S. 49-54.
11. Wang W., Wu H., Li M. Deep neural networks with batch speaker normalization for intoxicated speech detection // 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC). – IEEE, 2019. – S. 1323-1327.
12. Bone D. et al. Intoxicated speech detection by fusion of speaker normalized hierarchical features and GMM supervectors // Twelfth Annual Conference of the International Speech Communication Association. – 2011.
13. Bocklet T., Riedhammer K., Nöth E. Drink and Speak: On the automatic classification of alcohol intoxication by acoustic, prosodic and text-based features // Twelfth Annual Conference of the International Speech Communication Association. 2011.
14. Berninger K., Hoppe J., Milde B. Classification of speaker intoxication using a bidirectional recurrent neural network // Text, Speech, and Dialogue: 19th International Conference, TSD 2016, Brno, Czech Republic, September 12-16, 2016, Proceedings 19. – Springer International Publishing, 2016. – pp. 435-442.
15. Metodicheskie ukazaniya Ministerstva zdravoohraneniya SSSR ot 03.07.1974 «O sudebno-medicinskoj diagnostike smertel'nyh otravlenij jetilovym alkogolem i dopuskaemyh pri jetom oshibkah» [Elektronnyj resurs] // Zakonodatel'stvo v oblasti sudebno-medicinskoj jekspertizy. 1974. URL: http://zakon.forens-med.ru/doc/mz/3_25_101.html (data obrashheniya: 05.10.2025).
16. Chastikova V. A., Zherlicyn S. A., Vojlova D. O. Nejrosetevaja sistema biometricheskoy identifikacii lichnosti po golosu // Vestnik Adygejskogo gosudarstvennogo universiteta. Serija 4: Estestvenno-matematicheskie i tehničeskie nauki. – 2023. – №. 1 (316). – S. 70-79.
17. Badr A. A., Abdul-Hassan A. K. CatBoost Machine Learning Based Feature Selection for Age and Gender Recognition in Short Speech Utterances // International Journal of Intelligent Engineering & Systems. – 2021. – T. 14. – №. 3.
18. Sulavko A. et al. Biometric-based key generation and user authentication using voice password images and neural fuzzy extractor // Applied System Innovation. 2025.T. 8. №. 1. S. 13.

ИНИВАТОВ Даниил Павлович, ассистент кафедры «Комплексная защита информации» федерального государственного автономного образовательного учреждения высшего образования «Омский государственный технический университет». 644050, г. Омск, пр. Мира, 11. E-mail: daniilini@mail.ru

INIVATOV Daniil Pavlovich, Assistant of Integrated data protection department of the Federal State Autonomous Educational Institution of Higher Education “Omsk State Technical University”. 11 Mira Ave., Omsk, 644050, Omsk. E-mail: daniilini@mail.ru