

**«Национальный исследовательский ядерный университет «МИФИ»
(НИЯУ МИФИ)**

На правах рукописи



Алюшин Александр Михайлович

**МЕТОДЫ, МОДЕЛИ И АЛГОРИТМЫ
АУДИОМАРКИРОВАНИЯ НОСИТЕЛЕЙ
ДОКУМЕНТИРОВАННОЙ ИНФОРМАЦИИ**

Специальность 2.3.6 - Методы и системы защиты информации,
информационная безопасность

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Москва – 2026

Работа выполнена в федеральном государственном автономном образовательном учреждении высшего образования «Национальный исследовательский ядерный университет «МИФИ» (НИЯУ МИФИ)

Научный руководитель: Дворянкин Сергей Владимирович – доктор технических наук, профессор, профессор кафедры «Стратегии и технологии кибербезопасности» Института интеллектуальных кибернетических систем НИЯУ МИФИ

Официальные оппоненты: Шелухин Олег Иванович – доктор технических наук, профессор, заведующий кафедрой «Информационная безопасность» Московского технического университета связи и информатики (МТУСИ)

Безродный Борис Федорович – доктор технических наук, профессор, ведущий научный сотрудник отдела разработки технической документации специального программного обеспечения Управления комплексов средств автоматизации Центра специальных систем АНО «Институт инженерной физики»

Хохлов Николай Степанович – доктор технических наук, профессор, профессор кафедры «Инфокоммуникационные системы и технологии» Воронежского института Министерства внутренних дел Российской Федерации

Защита диссертации состоится 15 апреля 2026 года в 15 часов 00 минут на заседании диссертационного совета МИФИ.2.05 федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский ядерный университет «МИФИ» (115409, г. Москва, Каширское шоссе, 31).

С диссертацией можно ознакомиться в библиотеке и на сайте <https://ds.mephi.ru> федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский ядерный университет «МИФИ».

Автореферат разослан «__» _____ 2026 г.

Ученый секретарь
диссертационного совета, к.т.н.



Кессаринский Л.Н.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы

Вызовы, связанные с фальсификацией и неправомерным использованием документированной информации (ДИ): текстовых, видео-, фото- и фоно- документов с соответствующими реквизитами, являются в настоящее время значимыми информационными угрозами. По данным ГИАЦ МВД РФ только в январе – июле 2025 г. зарегистрировано 424,9 тыс. преступлений, совершенных с использованием информационно-телекоммуникационных технологий. Более 12% из них были связаны с указанными угрозами. Для их нейтрализации, среди прочих, сегодня активно развиваются технологии маркирования ДИ и её носителей, под которыми подразумевается процесс нанесения изображений маркеров (меток) на объекты защиты для их последующей идентификации, верификации и отслеживания перемещений.

В качестве маркеров ДИ в системах цифрового (электронного), аналогового (бумажного) и смешанного документооборота используются уникальные графические образы в виде бланков и печатей предприятий, рукописной подписи, водяных знаков, голограмм и т.п., которые идентифицируют только сам носитель информации (НИ) (бумага, диск, флеш-накопитель, сигнал и др.) с размещённой на нем ДИ и не обеспечивают выявление её изменений со стороны злоумышленника (ЗЛ).

Использование в смешанном документообороте в качестве маркера электронной подписи (ЭП) и различных видов штрих-кодов на практике бывает затруднительным, особенно в условиях различных мешающих факторов, приводящих к замятости, закрашиванию, частичному повреждению и, в итоге, к неработоспособности такого маркера. Кроме того, в современном маркировании отсутствует возможность фиксации биометрических данных, по которым можно оценить психоэмоциональное состояние автора, подписывающего документ. Помимо этого, низка информационная эффективность самих маркеров, требуется спецоборудование для их вставки-считывания в документ. Таким образом, возникает насущная необходимость создания такого маркера, который мог бы удовлетворять следующим требованиям:

- содержать в своем составе наиболее важную (значимую) часть защищаемой ДИ для определения факта возможного её изменения ЗЛ, а также персонализированные признаки её автора или пользователя;
- быть экономичным, дополнительным к имеющимся элементом защиты ДИ, устойчивым к помехам и искажениям и использующим для подтверждения подлинности распространенную офисную технику;
- использоваться со всеми видами ДИ, включая мультимедийный контент (МК), для предотвращения подмены, подделок и неправомерного использования.

Решение предлагается в развитии технологии аудиомаркирования, которая заключается в преобразовании защищаемой информации по схеме «звук (речь) – изображение (маркер) – звук (речь)», что позволяет в качестве маркера ДИ использовать графический образ (ГО) корректно рассчитанной спектрограммы речевого сигнала (РС) и её модификации до требуемых свойств маркера.

Данная работа состоит из нескольких взаимосвязанных компонент, задействованных в указанной схеме аудиомаркирования, базирующихся на опыте ведущих отечественных и зарубежных ученых из различных областей. Так вопросы организации защищенного документооборота поднимались в работах Конявского В.А., Королева И.Д., Щербакова А.Ю., Парфенова В.Г., Савельева А.В. и др.

Вопросам цифровой обработки речевой информации (РИ), ее спектрального анализа, методам повышения разборчивости и защиты РИ от фальсификации посвящены работы ведущих отечественных исследователей – Бонч-Бруевича А.М., Голубинского А.Н., Дворянкина С.В., Женило В.Р., Козлачкова С.Б., Минаева В.А., Пирогова А.А., Покровского Н.Б., Хорева А.А., Чудновского Л.С., Шелухина О.И. Из зарубежных ученых, в первую очередь, следует выделить работы таких ученых, как Griffin D., Lim J., McAulay R.J., Quatieri T. и др.

Работы, связанные с методами стеганографии, встраивания контрольной информации в фотоизображения, видео- и аудиосигналы для подтверждения их авторства, подлинности, корректности использования проводились Грибуниным В.Г., Коржиком В.И., Шелухиным О.И., Han-Yang Lo, Topiwala S., Kim J.R. и др.

Проведенный анализ работ ученых в указанных областях позволил сформировать научную основу настоящего исследования. Несмотря на значительный прогресс в разработке методов защиты ДИ, комплексный анализ выявил ограниченность и недостаточную эффективность существующих решений, в том числе для маркирования различных видов НИ, обеспечивающих повышенную устойчивость к современным видам киберпреступности. В связи с этим была сформулирована **научная задача**, которая заключается в совершенствовании существующих и создании новых технологий защиты документооборота и мультимедийного контента на основе предлагаемых методов и решений аудиомаркирования.

Поэтому разработка новых методов, моделей и алгоритмов аудиомаркирования различных НИ, таких, как цифровые и бумажные документы, цифровой мультимедийный контент и др., является востребованной, а тема исследований диссертационной работы – актуальной.

Цель диссертационного исследования – повышение уровня защищенности документированной информации и мультимедийного контента от фальсификации и неправомерного использования посредством совершенствования существующих и разработки новых методов, моделей и алгоритмов аудиомаркирования.

Объект исследования – системы и процессы защищенного информационного обмена документированной информацией и мультимедийным контентом (аудио, видео, текст, фото).

Предмет исследования – методы, модели и алгоритмы звукового маркирования носителей документированной информации и мультимедийного контента, обеспечивающие защиту от фальсификации и неправомерного использования.

Основные задачи исследования

В диссертационной работе поставлены следующие задачи:

- анализ методов и средств обеспечения защищенности документированной информации от существующих и перспективных угроз;
- создание методов и моделей формирования графических образов аудиомаркирования на основе спектрально-временных представлений речевых сигналов;
- улучшение процессов звукового маркирования документированной информации в документообороте различных видов;
- создание алгоритмов внедрения маркировочных меток в аудио- и фотодокументы;
- проектирование архитектуры и разработка программно-аппаратного комплекса, реализующего предложенные методы, модели и алгоритмы звукового маркирования.

Методы исследования, которые были использованы в работе: системный и спектральный анализ, цифровая обработка сигналов и изображений, распознавание образов, компьютерная стеганография.

Научная новизна работы заключается в совершенствовании существующих и разработке новых методов, моделей и алгоритмов аудиомаркирования для защиты ДИ от фальсификации, а мультимедийного контента от неправомерного использования, в том числе:

1. Предложена обобщенная концептуальная модель анализа-синтеза речи для формирования-считывания аудиомаркеров, используемых в качестве дополнительного элемента защиты документированной информации и мультимедийного контента, реализующая усиленную тройственную связь между защищаемой информацией, её носителем и биометрическими признаками автора (пользователя).

2. Предложен новый метод представления аудиомаркера в виде бинарного графического образа (бинарной модели), что повышает его устойчивость к воздействию различного рода мешающих факторов, обеспечивает заданный уровень разборчивости, позволяет размещать его на различных носителях информации и использовать общедоступные технические средства для реализации технологии защитного аудиомаркирования.

3. Разработаны алгоритмы применения маркировочных звуковых меток в качестве цифровых водяных знаков в аудио- и фотодокументах для противодействия их неправомерному использованию и защиты авторских прав.

Теоретическая значимость результатов исследования состоит в разработке новых методов, моделей и алгоритмов звукового маркирования носителей информации, используемых в качестве дополнительных экономических элементов защиты подлинности, авторства и целостности документов и цифрового контента от угроз несанкционированного доступа (НСД) и неправомерного использования. В том числе, метода взаимного преобразования изображений полутоновой и бинарной спектрограмм, позволяющего привлечь к синтезу и анализу голосовой подписи широкий потенциал средств цифровой обработки бинарных и контурных изображений, и тем самым повысить эффективность аудиомаркирования, расширить круг его возможных приложений типа шумоподавления и реконструкции искаженных РС, аудиокодирования, стеганографии, технического закрытия и др.

Практическая значимость работы обусловлена тем, что в ней обозначены пути решения важной научно-практической задачи по разработке методов, моделей и алгоритмов защиты документированной информации от фальсификации, подделок и неправомерного использования.

Основные результаты, имеющие практическое значение:

- подтверждена возможность эффективного использования технологии аудиомаркирования для верификации подлинности и авторства ДИ от фальсификации;

- разработаны методы и алгоритмы для практической реализации технологии аудиомаркирования ДИ на распространенной офисной технике и за приемлемое время;

- разработан комплекс специализированных программ для реализации предложенных методов аудиомаркирования, защищенных свидетельствами о государственной регистрации программ для ЭВМ.

Полученные результаты были использованы в рамках гранта № 15130ГУ/2020 Фонда содействия инновации, а также гранта ИБ МТУСИ № 40469-06/23-К Министерства Цифрового развития РФ.

Соответствие диссертации паспорту научной специальности

Представленная диссертация удовлетворяет п. 4, п. 15 паспорта специальности 2.3.6 – Методы и системы защиты информации, информационная безопасность: п. 4. Системы документооборота (вне зависимости от степени их компьютеризации) и средства защиты циркулирующей в них информации; п. 15. Принципы и решения (технические, математические, организационные и др.) по созданию новых и совершенствованию существующих средств защиты информации и обеспечения информационной безопасности.

Основные положения, выносимые на защиту:

- обобщенная концептуальная модель защитного аудиомаркирования по схеме «звук – изображение – звук», позволяющая формировать/считывать аудиомаркеры для эффективного противодействия фальсификации и подделке значимой информации, ее неконтролируемому использованию;

- метод представления и применения полутоновых и бинарных сонограмм в качестве голосовой подписи, связывающей защищаемую ДИ с элементом ее защиты (аудиомаркером) и личными данными автора (собственника) документа;

- новый метод представления речевой информации в виде графического образа бинарной модели спектрограммы, повышающий её устойчивость к деструктивному воздействию мешающих факторов, позволяющий сохранить итоговую разборчивость на уровне 80% от исходной;

- алгоритмы звуковой маркировки цифровых изображений и аудиосигналов в мультимедийном контенте для защиты от неправомерного использования.

Достоверность и обоснованность результатов достигается за счет трех факторов: корректного применения математических методов, согласованности расчетных данных с ожидаемыми результатами и подтверждения вычислительными экспериментами.

Апробация результатов

Основные положения и результаты диссертационной работы докладывались и обсуждались на следующих международных и всероссийских научных конференциях:

- XXIV Всероссийская научно-практическая конференция «Проблемы информационной безопасности в системе высшей школы», Москва, 1 февраля 2017 г.;
- The 11th IEEE International Conference on Application of Information and Communication Technologies (AICT), Moscow, 20-22 September 2017;
- Международная молодежная научно-практическая конференция студентов, аспирантов и молодых ученых «Информационная безопасность в банковско-финансовой сфере», Москва, 28-30 ноября 2017 г.;
- XXI Международная конференция по мягким вычислениям и измерениям, Санкт-Петербург, 23-25 мая 2018 г.;
- The International Science and Technology Conference «FarEastCon 2019», Vladivostok, Russia, October 2019;
- XXIII International Conference on Soft Computing and Measurements (SCM 2020), St. Petersburg, Russia, 27-29 May 2020;
- XXV International Conference on Soft Computing and Measurements (SCM 2022), St. Petersburg, Russia, 25-27 May 2022;
- III Всероссийская научная школа-семинар «Современные тенденции развития методов и технологии защиты информации», Москва, МТУСИ, 25-27 октября 2023 г.;
- Международная научно-практическая конференция «Применение современных технологий в противодействии киберпреступности: проблемы и пути решения», Алматы, 30 ноября 2023 г.;
- Вторая Всероссийская научно-техническая конференция «Кибернетика и информационная безопасность», Москва, НИЯУ МИФИ, 22-23 октября 2024 г.;
- Всероссийская научно-практическая конференция «Кадровое обеспечение информационной безопасности Российской Федерации», Москва, МТУСИ, 16-18 апреля 2025;
- XVI Международная научно-практическая конференция студентов, аспирантов и молодых ученых «Информационные технологии в науке, бизнесе и образовании. Стратегические вызовы цифровой трансформации в социотехнических системах», Москва, МГЛУ, 25 ноября 2025 г.;
- Третья Всероссийская научно-техническая конференция «Кибернетика и информационная безопасность», Москва, НИЯУ МИФИ, 3-4 декабря 2025 г.

Публикации

По основным положениям диссертационной работы опубликованы 15 научных работ, из них 6 статей в рецензируемых журналах, входящих в Перечень ВАК, 9 публикаций – в материалах международных конференций в изданиях, входящих в базы цитирования Scopus и Web of Science. Получены 4 свидетельства о государственной регистрации программ для ЭВМ (приравниваются к публикациям в изданиях, рекомендованных ВАК Министерства образования и науки Российской Федерации).

Личный вклад соискателя

Основные научные результаты получены автором лично: при разработке усовершенствованного метода представления речевой информации в виде бинарной спектрограммы; методов и алгоритмов обработки изображений динамических спектрограмм, включая их использование в процессах создания, распознавания и синтеза речевой информации; при создании программных комплексов в стационарном и мобильном (носимом) исполнении, реализующих распознавание и синтез маркеров РС по изображениям их спектрограмм.

Внедрение результатов работы

Разработанные в диссертационной работе методические и алгоритмические средства использованы в Национальном институте радио и инфокоммуникационных технологий (АО «НИРИТ») в качестве эффективного инструмента защиты авторских прав на мультимедийный контент за счет повышения уровня его защищенности от несанкционированного копирования и размножения. Получен акт об использовании результатов диссертационной работы.

Разработанные в диссертационном исследовании преобразования для получения аудиомаркеров в оттенках серого и в бинарном вариантах, а также разработанный веб-ресурс с соответствующими преобразованиями внедрены в учебный процесс Института интеллектуальных наук Федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный лингвистический институт» (ФГБОУ ВО МГЛУ). Получен акт о внедрении результатов диссертационной работы.

Материалы диссертационного исследования, а именно: методы преобразований изображений из графической формы в звуковую и обратно, бинарная модель представления звуковой информации, позволяющие организовывать на рабочих местах предприятия технологию защитного маркирования значимой информации, внедрены в практическую деятельность ООО «Института проблем безопасности и анализа информации». Получена справка о внедрении результатов диссертационной работы.

Структура и объем диссертации

Диссертация состоит из введения, четырех разделов, заключения, списка использованной литературы и двух приложений. Общий объем работы составляет 184 страницы, содержит 84 рисунка и 40 таблиц. Список литературы представлен 126 наименованиями.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы диссертационной работы, сформулированы цель и задачи исследования, обоснованы научная новизна, теоретическая и практическая значимость работы, сформулированы основные положения, выносимые на защиту.

В первом разделе рассмотрены основные носители информации и выявлена их роль в информационном обмене. Показано, что, несмотря на бурное развитие цифровых технологий, смешанный документооборот продолжает оставаться часто используемым видом НИ.

Выявлены основные угрозы (Таблица 1) фальсификации документированной информации и нелегитимного использования цифрового мультимедийного контента.

Таблица 1 – Дестабилизирующие факторы, потенциальные угрозы

№	Дестабилизирующие факторы, потенциальные угрозы
1.	Искажение, потеря части информации бумажных документов в результате старения, порчи
2.	Преднамеренное искажение информации электронного/бумажного документа злоумышленниками
3.	Получение подписи автора под принуждением
4.	Перехват, похищение всего документа
5.	Подделка подписи автора документа
6.	Нахождение автора документа в неадекватном состоянии (стресса, заторможенности, усталости)
7.	Оформление ряда важных документов только на бумажных носителях
8.	Другие угрозы и требования (полный список см. в диссертационной работе)

Проанализированы основные методы маркирования носителей информации, выявлены их преимущества и недостатки. Проведена классификация и анализ эффективности применяемых на практике подходов к подтверждению валидности документов на электронных (бит четности, контрольная сумма, CRC-коды, код Рида-Соломона, ЭП) и бумажных (одномерное и двумерное штриховое кодирование (Рисунок 1,а), полиграфические методы маркирования (Рисунок 1,б), встроенные магнитные полосы и ASIC, личная подпись и отпечатки пальцев) носителях информации.



Рисунок 1 – а) Двумерные штрих-коды;
б) Полиграфические методы поддержания валидности документов.

Проведенный анализ позволил обосновать актуальность и перспективность поиска и применения маркеров иной природы, интегрирующих в своем составе основные достоинства и возможности рассмотренных подходов.

Рассмотрены существующие методы защиты документов, выделены их основные преимущества и недостатки (Рисунок 2).

В результате анализа существующих в настоящее время дестабилизирующих факторов и потенциальных угроз (Таблица 1) для защищенного документооборота сформулированы основные требования, которые заключаются в установлении связи между защищаемой информацией, размещенной на НИ, технологией ее защиты и личными данными автора НИ. Под ними, в первую оче-

редь, подразумевают биометрические данные голоса, а также другие биометрические данные (пульс, давление, температура и др.) и звуковые метки, что делает фальсификацию и подделку ДИ крайне затруднительными или невозможными.










Документ	Защита	Защищенный документ	Свойства способов защиты
	 Подпись		+ Связь защиты с автором документа x Связь защиты с текстом документа x Надежность инструмента защиты x Фиксирование психоэмоционального состояния автора
	 Электронная подпись		+ Связь защиты с автором документа + Связь защиты с текстом документа x Надежность инструмента защиты x Фиксирование психоэмоционального состояния автора
	 Голосовая подпись		+ Связь защиты с автором документа + Связь защиты с текстом документа + Надежность инструмента защиты + Фиксирование психоэмоционального состояния автора

Рисунок 2 – Методы защиты документов, их преимущества и недостатки

Сделан вывод о том, что эффективным способом защиты информации, который удовлетворяет поставленным требованиям, является технология голосовой подписи (ГП) (Рисунок 3), создаваемая на основе образов речевых спектрограмм, содержащих важную контекстную информацию из документа или мультимедийного контента.

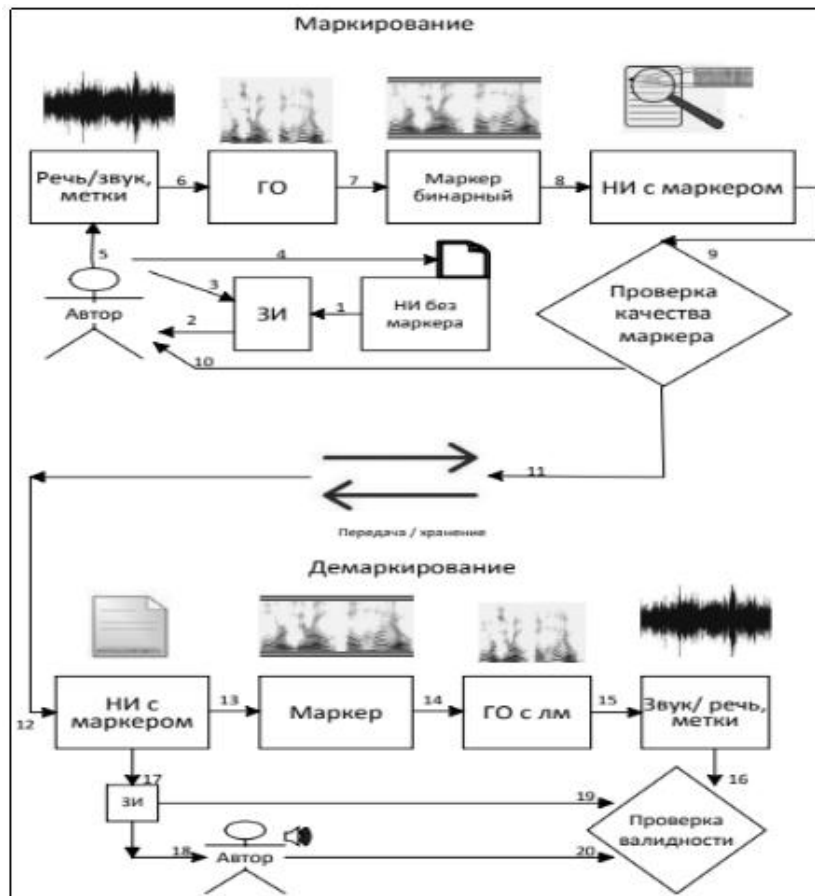


Рисунок 3 – Схема маркирования и демаркирования носителя информации аудиомаркером голосовой подписи

Данная технология аудиомаркирования (Рисунок 3) заключается в следующем: из ДИ выделяется та часть данных (1), которую необходимо защитить. Далее эти данные используются (2) автором для озвучивания и получения звукового файла (5) нужного формата и качества. Затем, с помощью спектрального преобразования (6) строится графический образ (ГО), при добавлении границ к которому (7) и будет готова голосовая подпись, которая добавляется к защищаемому носителю информации (8), размещается на нем, и таким образом участвует в информационном документированном обмене (11).

При необходимости автор может изменить защищаемую информацию (ЗИ) (3) или внести изменения как до получения финальной ГП (4), так и после получения (10) (например, если не устроило качество (9) звучания РС из маркера). Этапы демаркирования и получения звуковой информации из ГП выполняются в обратной последовательности (12,13,14,15). В заключительной стадии осуществляется проверка соответствия как смыслового содержимого полученного звукового файла (16) с защищаемой информацией (19), так и соответствия голоса автора (17,18,20) голосу в звуковом файле.

Изменение ЗИ требует изменения ГП, а сделать это без биометрических данных автора документа и без его личного участия практически невозможно.

Предлагаемый метод защиты ДИ не является альтернативой известных криптографических методов, а служит их дополнением. Совместное применение криптоалгоритмов и предлагаемого метода существенно увеличит степень защиты информации, особенно той, для которой величина предполагаемого ущерба при нарушении защиты потенциальным ЗЛ значительна.

Рассмотренный подход защитного аудиомаркирования основывается на обобщенной концептуальной модели по схеме преобразования «звук – изображение (аудиомаркер) – звук» (Рисунок 4), позволяющей применять современные методы цифровой обработки изображений и машинного обучения для аудиосигналов с целью противодействия фальсификации и подделке значимой информации или защиты авторских прав мультимедийного контента.

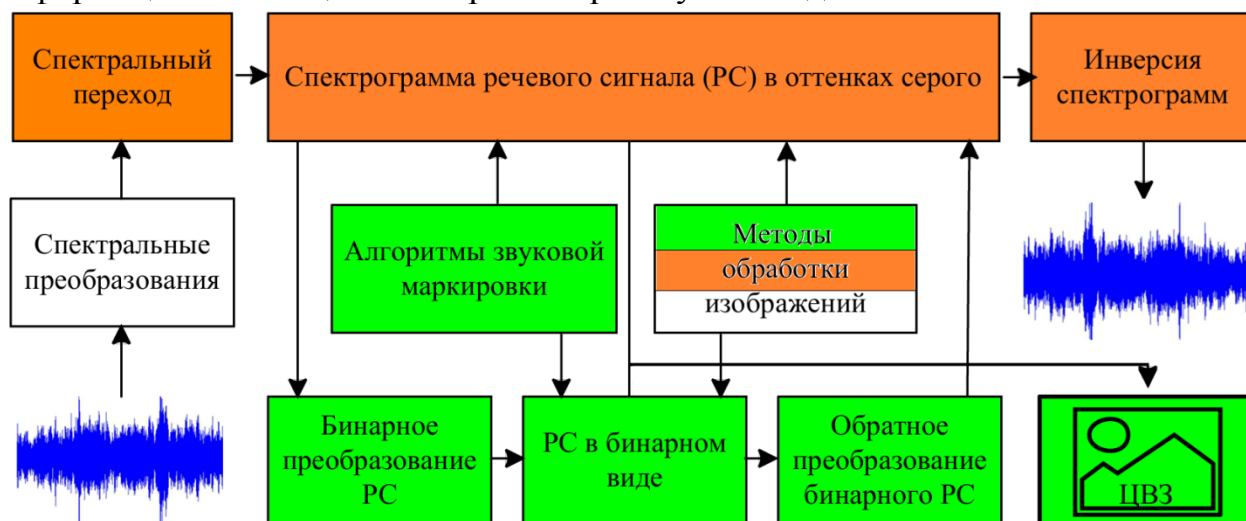


Рисунок 4 – Обобщенная концептуальная модель (оранжевым цветом выделено то, что было уточнено в диссертационном исследовании, зеленым – новые составляющие, научная новизна)

Предложена система метрик (расстояние Минковского, метрика ESTOI, метод PEAQ) для разрабатываемой технологии маркирования. Метрика ESTOI использовалась для определения разборчивости речевого сигнала до и после преобразования, а метод PEAQ (оценка восприятия качества аудиосигнала) (ГОСТ Р 58245— 2018/ IEC/TR 62251:2003), рекомендованный ITU-R BS.1387-1 для оценки общего качества аудосигналов. Расстояние Минковского применялось для измерения скрытности (похожести) изображения со встроенной информацией с исходным эталонным изображением.

Расстояние Минковского порядка p между двумя точками $x, y \in \mathfrak{R}^n$ определяется как:

$$\rho(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{1/p} \quad (1)$$

Метрика Extended Short-Time Objective Intelligibility (ESTOI):

$$d = \frac{1}{JM} \sum_{j,m} d_{j,m}, \quad d_{j,m} = \frac{(x_{j,m} - \mu_{x_{j,m}})^T (y_{j,m} - \mu_{y_{j,m}})}{\|x_{j,m} - \mu_{x_{j,m}}\| \|y_{j,m} - \mu_{y_{j,m}}\|}, \quad (2)$$

где $d_{j,m}$ – частные показатели разборчивости, d – итоговый показатель разборчивости, величины μ характеризуют средние значения соответствующих векторов. Эти метрики были использованы для оценивания результатов основных этапов работы.

Во втором разделе рассматривались задачи формирования описаний аудиосигналов в виде графических образов спектрограмм, а также синтеза искусственной речи на их основе.

Показаны преимущества спектрального представления сигнала с помощью разделения частот как более информативного способа представления информации о сигнале, простоты анализа и др.

Проведено обоснование выбора аналитического преобразования, предназначенного для создания аудиоподписи из акустического (речевого) сигнала. Рассмотрены различные формы представления РС во временной и частотной областях.

Проведено сравнение достоинств и недостатков преобразований Фурье, вейвлетов, Эрмита, Фурье-Гаусса, позволяющих получить спектральные характеристики речевых сигналов. Показано, что наиболее предпочтительным (для работы с речью в спектральной плоскости) является кратковременное преобразование Фурье со взвешиванием усеченным окном Гаусса, обеспечивающее наиболее равномерную спектральную плотность для получаемых изображений спектрограмм. Это дает возможность применять специальные алгоритмы для работы с гармонической структурой РС, а также снизить требования по разрешающей способности к применяемым средствам копирования и сканирования бумажных документов, защищенных с помощью ГП.

Тогда спектр сигнала можно описать через круговую частоту ω и время t :

$$S(\omega, t) = \int_{-\infty}^t s(\tau) w(t - \tau) e^{-i\omega\tau} d\tau, \quad (3)$$

где $w(t)$ в данной формуле является оконной функцией (ОФ), которая подбирается в зависимости от контекста поставленной задачи. Для получения наибольшей информации о сигнале на практике используют амплитудный спектр $|S(\omega, t)|$:

$$|S(\omega, t)| = \sqrt{a^2(\omega, t) + b^2(\omega, t)}, \quad (4)$$

где $a(\omega, t)$ и $b(\omega, t)$ – это вещественная и мнимая часть временного спектра,

$$a(\omega, t) = \int_{-\infty}^t s(\tau)w(t-\tau)\cos\omega\tau d\tau; \quad b(\omega, t) = \int_{-\infty}^t s(\tau)w(t-\tau)\sin\omega\tau d\tau.$$

С помощью данных преобразований можно получать изображения спектрограмм речевого сигнала (сонограмм) (Рисунок 5). Каждый пиксель на спектрограмме показывает интенсивность частоты в конкретный момент времени (черный – наибольшая мощность, белый – слабая). После проведения спектрального анализа формируется матрица S комплексных чисел, которая используется для построения спектрограммы, где интенсивность пикселя I с частотой ω на временном срезе t определяется по следующей зависимости:

$$I(\omega, t) = 40\log(1 + |S(\omega, t)|). \quad (5)$$

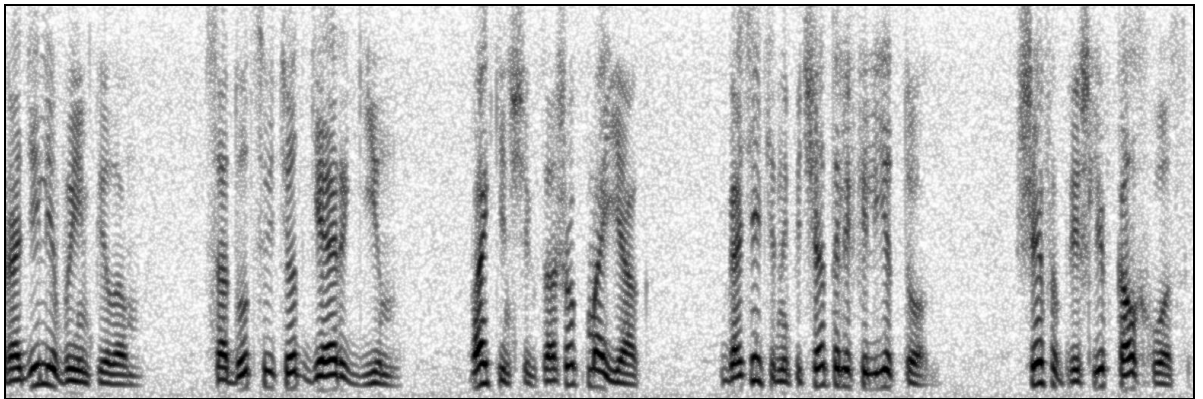


Рисунок 5 – Изображение спектрограммы речевого сигнала

При использовании быстрого преобразования Фурье (БПФ) качество получаемых спектрограмм РС (в первую очередь, контрастность выделенных гармоник основного тона) во многом определяется типом используемой ОФ $w(i)$.

Декодировка ГП осуществлялась на тех же принципах, что и прямое КФА, но в обратном порядке.

Обоснованы оптимальные параметры прямого и обратного Фурье-преобразования, использующегося при создании и декодировке ГП, с целью обеспечения необходимых показателей качества РС. Показано, что использование окон Гаусса и Хэмминга дает наилучшие результаты в контексте качества сонограмм. Для получения качественных спектрограмм и фазограмм при использовании БПФ рекомендуется использовать базу БПФ с $N=1024$, с шагом анализа (Рисунок 6), равным 50 отсчетов для сигналов с частотой 8000 Гц и 140 отсчетов для сохранения изображений.

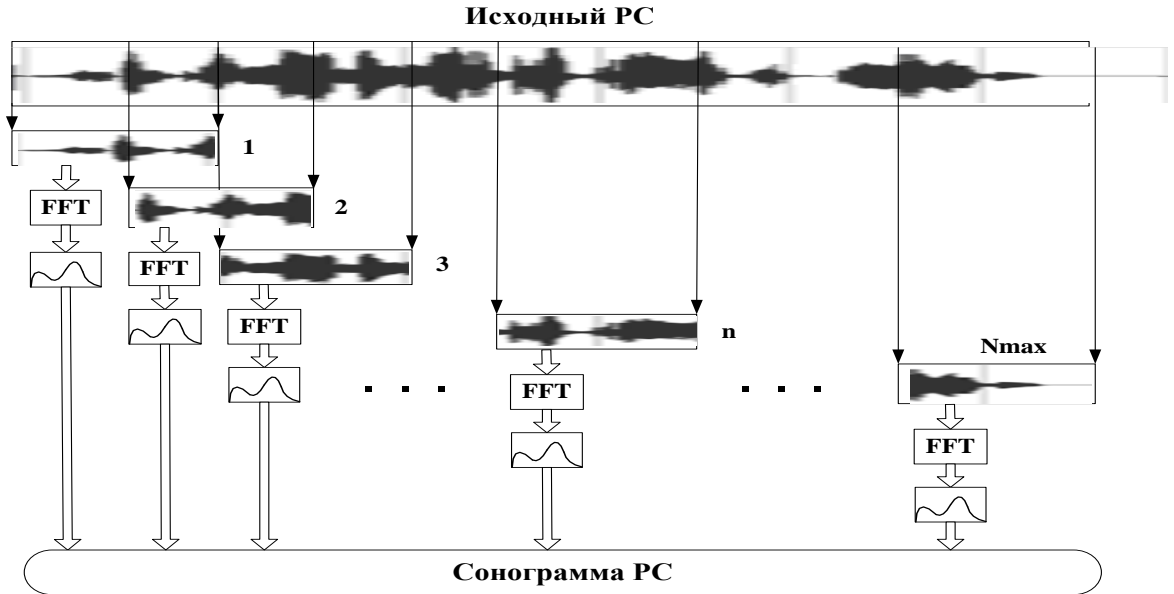


Рисунок 6 – Иллюстрация суммирования-наложения окон анализа на сигнал

Особое внимание в ходе инверсии спектрограмм по трекам микрогармоник было уделено оценке искусственной фазы, которая рассчитывалась по следующему линейному закону:

$$\varphi_i(0) = \text{random}(512); \varphi_i(t) = \varphi_{\min(i \pm 5)}(t - \Delta t) + \Delta t \frac{2\pi i}{N};$$

если $\varphi_i(t) \geq 1024$, то $\varphi_i(t) = \varphi_i(t) - 1024$. (6)

Такой расчет фазы обеспечивает более низкую алгоритмическую сложность и высокое качество реконструкции для сохранения разборчивости речи (Таблица 2) по сравнению с известными однопроводными подходами, как например, метод McAulay–Quatieri, имеющий недостаток, связанный со сложным расчетом \arctan фазы, а также чрезвычайно ресурсоёмким и медленным на практике алгоритмом Гриффина-Лима, который итеративно реконструирует сигнал по амплитудному спектру, минимизируя расстояние D_M (повторяя итерации до 10-100 раз).

$$D_M(x(n), x_i(n)) = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \int_{-\pi}^{\pi} [|X(mS, \omega)| - |X_i(mS, \omega)|]^2 d\omega, \quad (7)$$

где $|X(mS, \omega)|$ – спектр исходного сигнала $x(n)$, а $|X_i(mS, \omega)|$ – спектр i -й оценки сигнала $x_i(n)$ через проецирование гибридного спектра во временную область по формуле (8).

$$x_{i+1}(n) = \frac{\sum_{m=-\infty}^{\infty} w(n-mS) \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{X}_i(mS, \omega) e^{j\omega n} d\omega}{\sum_{m=-\infty}^{\infty} w^2(n-mS)}, \quad \text{где } \hat{X}_i(mS, \omega) = |X(mS, \omega)| \cdot \frac{X_i(mS, \omega)}{|X_i(mS, \omega)|} \quad (8)$$

где $\hat{X}_i(mS, \omega)$ – это спектр $x_i(n)$ с ограничением по амплитуде, а $w(n-mS)$ оконная функция.

Предложен алгоритм адаптивного применения оконной функции в зависимости от частоты основного тона, обладающий рядом преимуществ по сравне-

нию с традиционными методами взвешивания сигнала в окне. Среди них – снижение вычислительной сложности, а также повышение точности определения частоты основного тона и его гармоник.

Учтены особенности применения изображения спектрограммы в качестве ГП в формате Grey Scale и формате Black and White для использования полиграфической техники различного качества и смартфонов, что позволило существенным образом снизить погрешности считывания ГП при различных искажающих факторах.

Это позволило осуществлять переход от полутонового к бинарному виду ГП, не изменяя размера изображения с незначительной потерей качества передаваемой в ГП информации.

Для этого разработан метод получения спектрограммы ГП в бинарном виде (Рисунок 7), который заключается в следующем.

1. Вначале строится спектрограмма в формате Grey Scale с найденными оптимальными параметрами.

2. Затем производится редактирование спектрограммы, очистка от помех и установка пороговых фильтров.

3. Далее на спектрограмме выделяется гармоническая структура речи (треки локальных максимумов – микрогармоник).

4. На последнем этапе происходит замена пикселей серого цвета в точках локальных максимумов на черные вертикальные отрезки длины L , которые определяются функцией преобразования:

$$L = L_{\min} + \frac{I_{\text{лок.макс}}}{255} \cdot (L_{\max} - L_{\min}). \quad (9)$$

$$I_{\text{лок.макс}}(k, t) = \left\{ \begin{array}{l} S(k, t) > P; \\ S(k, t) > S(k \pm 1, t \pm 1) \quad \forall \Delta t, k \neq 0, k \neq 255 \end{array} \right\}, \quad (10)$$

где L_{\min} , L_{\max} – минимальная и максимальная длина вертикального отрезка, k – частота на спектрограмме, t – временной срез на спектрограмме, $I_{\text{лок.макс}}(k, t)$ – положение k -ого локального максимума на временном срезе t ; P – значение порога; $S(k, t)$ – интенсивность пикселя на спектрограмме.

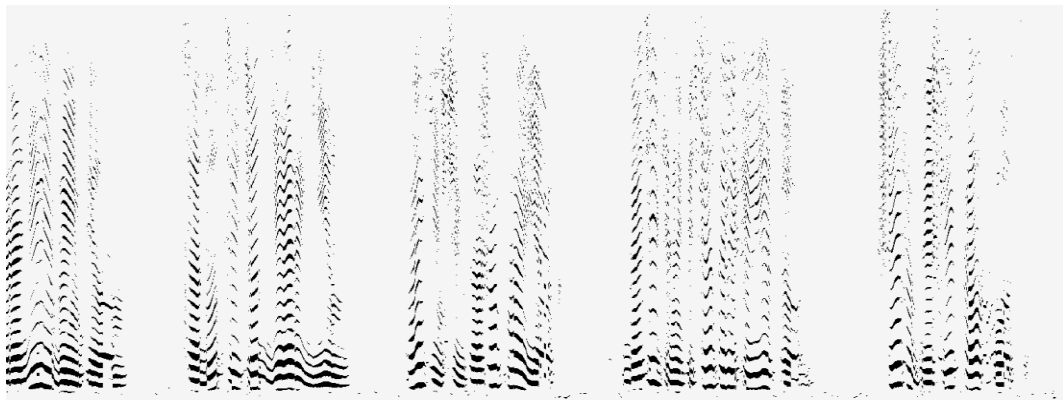


Рисунок 7 – Модель изображения ГП в графическом формате В&W (без границ для распознавания)

Проведено сравнение результатов синтеза по метрике ESTOI для определения разборчивости речи (PP) (Таблица 2).

Таблица 2 – Результаты сравнения синтезов для спектрограммы в оттенках серого и в бинарном видах по метрике ESTOI

Вид синтеза	Результат ESTOI для аудиомаркера в оттенках серого	Результат ESTOI для бинарного аудиомаркера
По всему спектральному срезу с исходной фазой	0.99	0.83
По всему спектральному срезу с искусственной фазой	0.95	0.80

Результаты эксперимента показывают, что синтез речи с искусственной фазой по бинарной спектрограмме обеспечивает PP по метрике ESTOI на уровне 0.80, а по спектрограмме в градациях серого – на уровне 0.95.

Разработана комплексная модель ГП с добавлением в нее других биопараметров автора, помимо речевой составляющей, которая в случае ее формирования на основе энергетического спектра, имеет вид:

$$\left| Si \left(\frac{F_0}{N}, t \right) \right| = PSi(t); \quad i = 1, \dots, N, \quad (11)$$

где F_0 – усредненное значение частоты основного тона голоса автора; N – число параметров, описывающих функциональное и психоэмоциональное состояние (ФПЭС) автора документа; $PSi(t)$ – функция, характеризующая временное изменение i -го биопараметра первого класса.

Использование биометрической информации с использованием расширенного набора параметров, включая биометрические данные, характеризующие ФПЭС автора, позволяет существенным образом повысить защищенность важных документов от подделки и фальсификации.

Решение указанных задач заложило основу разработки необходимых практических методик и алгоритмов и программного обеспечения, представленных в разделах 3, 4 диссертационного исследования.

Третий раздел посвящен разработке методов и алгоритмов звукового маркирования текстовых документов, цифровых изображений и аудиосигналов для их защиты от неправомерного использования.

Предложен метод подтверждения валидности документов с помощью ГП, дополнительно защищенной традиционными методами защиты информации (ЗИ), которые позволяют реализовать эффективное маскирование информативной части ГП. Так на Рисунке 8 представлен результат применения известного криптографического метода защиты данных для бинарной ГП.

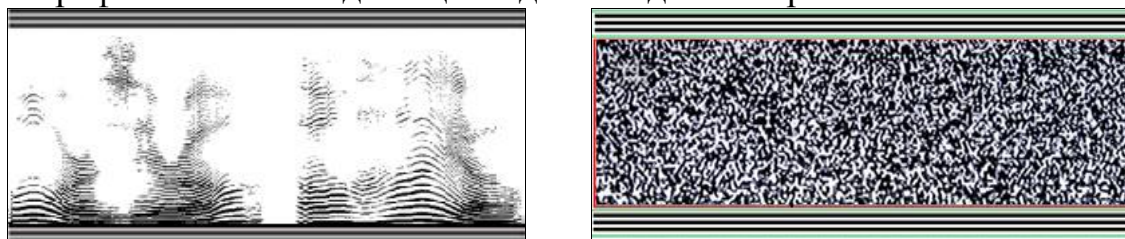


Рисунок 8 – Изображения ГП с использованием методики маскирования информативной части (слева – исходная ГП, справа – с защитой)

Сравнение рисунков оригинальной сонограммы с маскированием показывает, что применение такой защиты не позволяет идентифицировать даже структуру данных, так как получаемое изображение информативной части ГП практически является двухмерным изображением белого шума.

Разработанный бинарный маркер может быть интегрирован в QR-код. Так, в QR-код (Рисунок 9) размером 177 на 177 можно интегрировать не менее 2,5 секунд речи телефонного качества.



Рисунок 9 – Пример встраивания бинарного маркера в QR-код

Дополнительно рассмотрены существующие способы маркирования аудио-сигналов. С учетом их анализа были разработаны методы и алгоритмы маркирования звуковой информации. В качестве критерия оценки внесенных искажений в сигнал в результате маркирования был выбран параметр объективной разницы качества ODG, измеренный методом PEAQ. Данный показатель в дальнейшем был использован при тестировании алгоритмов аудиомаркирования.

В диссертации разработаны следующие алгоритмы маркирования аудиоданных:

- алгоритм маркирования размножением шумовой составляющей. Изображение спектрограммы такого подхода представлено на Рисунке 10,а;
- алгоритм маркирования удалением спектральной составляющей. Изображение спектрограммы такого подхода представлено на Рисунке 10,б;
- алгоритм маркирования вставкой информации на определенных частотах. Изображение спектрограммы такого подхода представлено на Рисунке 10,в;
- алгоритм внедрения контрольной информации в аудиосигнал посредством модификации фазы. Изображение спектрограммы такого подхода представлено на Рисунке 10,г.

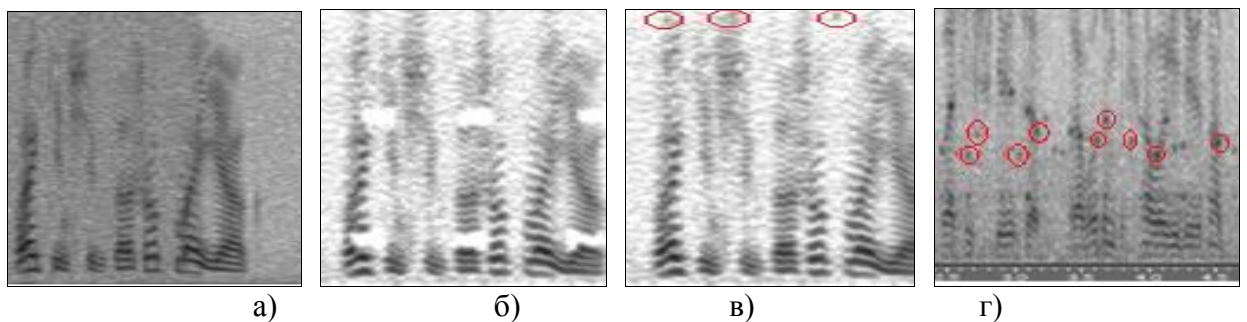


Рисунок 10 – Примеры спектрограмм со встроенными метками с помощью алгоритмов: а) маркирования размножением шумовой составляющей;

б) маркирования удалением спектральной составляющей;

в) маркирования вставкой информации на определенных частотах; г) изменения фаз

Сравнительные характеристики описанных выше алгоритмов встраивания меток в речевые сигналы представлены в Таблице 3.

Таблица 3. Сравнение различных методов аудиомаркирования речевых сигналов

Алгоритмы маркирования	Алгоритмы маркирования размножением шумовой составляющей	Алгоритмы маркирования удалением спектральной составляющей	Алгоритмы маркирования вставкой информации на определенных частотах	Алгоритмы маркирования посредством модификации фазы
Среднее значение параметра ODG анализируемых сигналов	-1.71	-0.53	-0.84	-1.83
Мера близости спектрограммы эталонного сигнала и маркированного (по Минковскому)	60% вставки зрительно заметны	84% вставки зрительно заметны	92% вставки зрительно заметны	98% вставки зрительно не заметны
Вероятность обнаружения меток (ВОМ) в цифровом сигнале	0.8	0.9	0.9	0.9
ВОМ в повторно записанном сигнале (ПЗС) на расстоянии 0.5 м от источника	0.3	0.8	0.9	0.8
ВОМ в ПЗС на расстоянии 2 м от источника	0.1	0.2	0.8	0.7
ВОМ в ПЗС на расстоянии 5 м от источника	< 0.1	< 0.1	0.6	0.5

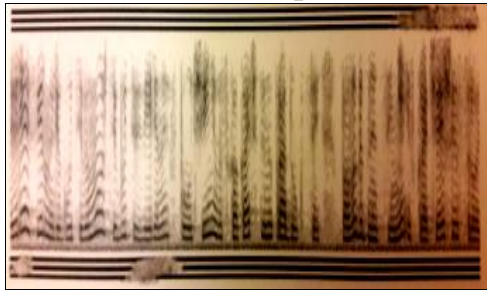
Из анализа результатов экспериментов (Таблица 3) можно сделать вывод, что метод внедрения контрольной информации в аудиосигнал посредством модификации фазы его опорных звуковых отпечатков и алгоритм маркирования вставкой информации на определенных частотах имеют наилучшие характеристики при детектировании аудиомаркера после перезаписи. Однако метод, основанный на модификации фазы, более скрытен для распознавания на спектрограмме.

В четвертом разделе решены методические и алгоритмические задачи, которые связаны с особенностями практической реализации технологии голосовой подписи, и проведена оценка результатов экспериментального тестирования.

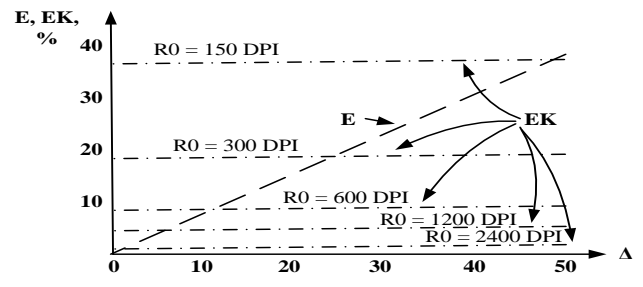
По итогу практической реализации технологии был разработан программный комплекс ДАК для стационарного компьютера, а также приложения для мобильных устройств android. Для этого были решены следующие задачи:

- разработана структура ГП с полями для меток позиционирования, обеспечивающая достоверное распознавание ГП в документе;
- разработан алгоритм распознавания меток позиционирования ГП, дано обоснование оптимальной формы графического представления границ ГП на основе интегрального критерия: минимум времени обработки и минимум занимаемой площади в составе ГП;
- показана высокая эффективность алгоритма распознавания границ ГП в условиях ухудшении качества носителя (Рисунок 11,а);

– разработаны методы и алгоритмы синтеза и декодирования ГП на основе комплексных моделей представления с включением биометрических данных.



а)



б)

Рисунок 11 – а) Возможность распознавания ГП при потертости границ;
б) Величина ошибки при декодировании ГП в зависимости от износа бумажного носителя

Проведен анализ условий создания, использования и хранения документов с ГП с учетом факторов, влияющих на качество и износ бумажного носителя. Разработаны критерии оценки применимости различных форм представления ГП, позволяющие определить границы их наиболее целесообразного использования в зависимости от разрешающей способности полиграфического оборудования и техники мобильного считывания. Так на Рисунке 11,б представлены зависимости величины ошибки E (12) от величины Δ , которая характеризует уровень старения бумажной основы, при различных значениях параметра R_0 (разрешающая способность) и относительной ошибки EK (13); I_{\max} – максимальное значение яркости пикселей изображения; N_m – средняя ширина линий спектрограммы, входящей в состав ГП при ее бинарном представлении.

$$E = \frac{2\Delta}{I_{\max}} 100\% = \frac{2\Delta}{255} 100\% \approx 0,78\Delta\% . \quad (12)$$

$$EK = \frac{1}{N_m} 100\% . \quad (13)$$

Представленные на Рисунке 11,б зависимости дают возможность обоснованно выбрать способ графического представления (в тонах серого, либо в бинарном виде) спектрограммы в качестве ГП.

Кроме того, в 4-м разделе представлены результаты практической реализации разработанных методик и алгоритмов, ориентированных на использование в составе мобильных устройств. Основное внимание уделено маркированию речевой информации как наиболее сложному и информационно емкому компоненту ГП.

Показано, что разработанные методы, модели и алгоритмы могут быть использованы в системах защиты документов от подделок и фальсификации независимо от типа документооборота. Также могут найти свое применение в области стеганографии, либо применяться в качестве высокоэффективных технологий аудиокомпрессии с шумоподавлением и реконструкцией искаженных участков.

В заключении сформулированы основные результаты диссертационной работы.

ОСНОВНЫЕ ВЫВОДЫ И РЕЗУЛЬТАТЫ РАБОТЫ

В работе получены следующие результаты.

1. Проведен системный анализ технологий защитного маркирования, направленных на противодействие ключевым угрозам: фальсификации документированной информации и неправомерному использованию цифрового контента. На его основе обоснована целесообразность и исследована возможность применения технологии аудиомаркирования для повышения уровня защищенности документированной информации.

2. Предложена обобщенная концептуальная модель анализа-синтеза речи для формирования и считывания аудиомаркеров. Данная модель реализует тройственную связь в виде: «защищаемая информация – её носитель – биометрические признаки автора (пользователя)», что принципиально повышает достоверность верификации и устойчивость системы защиты в целом.

3. В рамках данной модели предложено и обосновано использование аудиомаркеров в качестве голосовой подписи, размещенной на носителе информации в виде стикера или рисунка, что позволяет использовать для его применения общедоступные технические средства.

4. Предложен метод представления аудиомаркера в виде бинарного графического образа (бинарной модели), который существенно повышает устойчивость к различным внешним негативным факторам, что обеспечивает возможность размещения на разных носителях информации (бумажных, цифровых).

5. Разработаны и исследованы специализированные алгоритмы применения маркировочных звуковых меток в качестве цифровых водяных знаков. Данные алгоритмы повышают защищенность документированной информации и мультимедийного контента от фальсификации и неправомерного использования.

6. Осуществлена практическая реализация разработанных методов и алгоритмов в составе различных программно-аппаратных комплексов для подтверждения подлинности документов и авторских прав. Так время распознавания аудиомаркера с мобильного устройства до начала его прослушивания не превышает 2–4 секунды.

7. Сформированы практические рекомендации по применимости алгоритмов, методов и программного обеспечения для реализации предложенных способов аудиомаркирования.

8. Проведенная экспериментальная апробация прикладного программного обеспечения показала возможность сохранения разборчивости исходного сигнала в бинарной ГП на уровне 80% от первоначальной, а полутоновой ГП в оттенках серого на уровне 95% от исходной.

Основные публикации по теме диссертации

В изданиях, рекомендованных ВАК Минобрнауки России:

1. Алюшин, А. М. Физиологические аспекты построения сонограмм и реконструкции спектра искаженных речевых вокализмов / А. М. Алюшин, М. В. Алюшин, С. В. Дворянкин, Н. С. Дворянкин // Безопасность информационных технологий. – 2025. – Т. 32. – № 2. – С. 32–47. (специальность 2.3.6, К2).
2. Алюшин, А. М. Быстрый синтез аудиосигналов по изображениям спектрограмм в задачах защиты речевой информации / С. В. Дворянкин, Н. С. Дворянкин, А. М. Алюшин // Вопросы кибербезопасности. – 2024. – №. 5(63). – С. 34-46. (специальность 2.3.6, К1)
3. Алюшин, А. М. Анализ перспектив использования фитнес-браслетов в качестве источника биометрической информации при синтезе биоподписи важного документа / А. М. Алюшин, С. В. Дворянкин // Безопасность информационных технологий. – 2024. – Т. 31. – № 1. – С. 63–74. (специальность 2.3.6, К2)
4. Алюшин, А. М. Методы биомаркирования защищаемых объектов / В. А. Минаев, С. В. Дворянкин, А. М. Алюшин // Информация и безопасность. – 2023. – Т. 26. – № 3. – С. 321-328. (специальность 2.3.6, К2)
5. Алюшин, А. М. Использование речевых технологий для защиты документооборота / А. М. Алюшин, С. В. Дворянкин // Безопасность информационных технологий. – 2017. – Т. 24. – № 2. – С. 6–15. (специальность 2.3.6, К2)
6. Алюшин, А. М. Особенности распознавания изображений речевой подписи на мобильных устройствах / А. М. Алюшин, Н. С. Дворянкин // Безопасность информационных технологий. – 2015. – № 4. – С. 38–45. (специальность 2.3.6, К2)

Научные работы в сборниках трудов международных конференций, индексируемые в Scopus:

1. Alyushin, A. M. A Biologically Inspired Approach to Protecting and Verifying the Authenticity of Important Documents / A. M. Alyushin, V. M. Alyushin, S. V. Dvoryankin, L. V. Kolobashkina // Studies in Computational Intelligence. – 2024. – Vol. 1130. – P. 50–55. https://doi.org/10.1007/978-3-031-50381-8_7
2. Alyushin, A. M. Universal indicator of information content of the speech signature of a document / A. M. Alyushin // Proceedings of 2022 25th International Conference on Soft Computing and Measurements, SCM 2022, St. Petersburg, May 2022. – 2022. – P. 159-162. <https://ieeexplore.ieee.org/document/9794888>
3. Alyushin, A. M. Correctness of the information contained in the biosignature document of the document / A. M. Alyushin, M. V. Alyushin, S. V. Dvoryankin // Proceedings of 2022 25th International Conference on Soft Computing and Measurements, SCM 2022, St. Petersburg, May 2022. – 2022 – P. 163-166. <https://doi.org/10.1109/SCM55405.2022.9794840>
4. Alyushin, A. M. Improving the reliability of the voice signature when it is used to protect paper documents in the information and communication system of enterprise management / A. M. Alyushin // Smart Innovation, Systems and Technologies. – 2020. – Vol. 172. – Q3. – P. 673–681. DOI:[10.1007/978-981-15-2244-4_63](https://doi.org/10.1007/978-981-15-2244-4_63)

5. Alyushin, A. M. Document protection technology in the digital economics using cognitive biometric methods / A. M. Alyushin // Procedia Computer Science. – 2020. – Vol. 169. – Q2. – P. 887–891. DOI: [10.1016/j.procs.2020.02.147](https://doi.org/10.1016/j.procs.2020.02.147)
6. Alyushin, A. M. Biologically inspired physical model of the vocal tract for the tasks of recognition of the current psycho-emotional state of a person / A. M. Alyushin // Advances in Intelligent Systems and Computing. – 2020. – Vol. 948. – P. 15–21. https://doi.org/10.1007/978-3-030-25719-4_3
7. Alyushin, A. M. Intelligent processing of speech information in the tasks of noise reduction for communication tools at the objects of the digital economy / A. M. Alyushin, N. M. Leonova, A. D. Modyaev // In Proc. of the 2020 23rd International Conference on Soft Computing and Measurements, SCM 2020, St. Petersburg, May 2020. – 2020. – P. 230–233. <https://doi.org/10.1109/SCM50615.2020.9198803>
8. Alyushin, A. M. Human face thermal images library for laboratory studies of the algorithms efficiency for bioinformation processing / A. M. Alyushin, M. V. Alyushin, L. V. Kolobashkina // In Proc. of the 11th IEEE International Conference on Application of Information and Communication Technologies, AICT 2017, Moscow, Russia. – 2019. – P. 1–5. <https://doi.org/10.1109/ICAICT.2017.8687078>
9. Alyushin, A. M. Increase a person's face vibrational image information content due to compensation of components caused by his speech activity/ A. M. Alyushin // Procedia Computer Science. – 2018. – Vol. 145.– Q2.– P. 37–42. DOI: [10.1016/j.procs.2018.11.006](https://doi.org/10.1016/j.procs.2018.11.006)

Свидетельства о регистрации программы для ЭВМ

1. Свидетельство № 2024684239. Российская Федерация. Свидетельство о государственной регистрации программы для ЭВМ «Программа для маркирования сигналов шумовой составляющей» / А. М. Алюшин. – Заявка № 2024681755 от 23.09.2024; дата гос. регистрации в Реестре программ для ЭВМ 15.10.2024. – 1 с.
2. Свидетельство № 2024690107. Российская Федерация. Свидетельство о государственной регистрации программы для ЭВМ «Программа для маркирования сигналов посредством удаления спектральных данных» / А. М. Алюшин. – Заявка № 2024681741 от 23.09.2024; дата гос. регистрации в Реестре программ для ЭВМ 12.12.2024. – 1 с.
3. Свидетельство № 2024683747. Российская Федерация. Свидетельство о государственной регистрации программы для ЭВМ «Программа для маркирования сигналов через вставку информации на определенных частотах» / А. М. Алюшин. – Заявка № 2024681737 от 23.09.2024; дата гос. регистрации в Реестре программ для ЭВМ 14.10.2024. – 1 с.
4. Свидетельство № 2022612581. Российская Федерация. Свидетельство о государственной регистрации программы для ЭВМ «Программный комплекс обработки аудиомаркеров»/ А. М. Алюшин. – Заявка № 2022612375 от 16.02.2022; дата гос. регистрации в Реестре программ для ЭВМ 28.02.2022. – 1 с.