

На правах рукописи



Волков Никита Андреевич

**МЕТОДИКА ОЦЕНКИ ЗАЩИЩЕННОСТИ
АКУСТИЧЕСКОЙ РЕЧЕВОЙ ИНФОРМАЦИИ ОТ УТЕЧКИ
ПО ТЕХНИЧЕСКИМ КАНАЛАМ С ПРИМЕНЕНИЕМ
СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ**

**Специальность 2.3.6. Методы и системы защиты
информации, информационная безопасность**

АВТОРЕФЕРАТ

**диссертации на соискание ученой
степени кандидата технических наук**

Самара – 2025

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

Научный руководитель: кандидат технических наук, доцент, доцент кафедры «Электронные системы и информационная безопасность» федерального государственного бюджетного образовательного учреждения высшего образования «Самарский государственный технический университет»,

г. Самара

Иванов Андрей Валерьевич

Официальные оппоненты:

Дворянкин Сергей Владимирович, доктор технических наук, профессор, профессор кафедры стратегии и технологии кибербезопасности (№43) института интеллектуальных кибернетических систем федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский ядерный университет «МИФИ»

Аникин Игорь Вячеславович, доктор технических наук, профессор, заведующий кафедрой систем информационной безопасности федерального государственного бюджетного образовательного учреждения высшего образования «Казанский национальный исследовательский технический университет им. А.Н. Туполева – КАИ»

Ведущая организация: Федеральное государственное бюджетное учреждение науки "Санкт-Петербургский Федеральный исследовательский центр Российской академии наук", г. Санкт-Петербург

Защита диссертации состоится 12.03.2026 г. в 15 ч. 15 мин. на заседании диссертационного совета 24.2.415.04 на базе федерального государственного автономного образовательного учреждения высшего образования «Томский государственный университет систем управления и радиоэлектроники», по адресу: 634050, г. Томск, пр. Ленина, 40, ауд. 201.

С диссертацией можно ознакомиться в библиотеке федерального государственного автономного образовательного учреждения высшего образования «Томский государственный университет систем управления и радиоэлектроники» и на сайте <https://postgraduate.tusur.ru/urls/rkemp34y>

Автореферат разослан «___»_____2026 года.

Ученый секретарь
диссертационного совета,
к.т.н., доцент

Костюченко Евгений Юрьевич

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Одной из ключевых задач в области обеспечения информационной безопасности является защита акустической речевой информации при проведении переговоров. Защищаемая информация может быть получена вследствие утечки информации по техническим каналам. Технический канал утечки информации (ТКУИ) представляет собой совокупность источника информации (диктора), физической среды распространения сигнала и технических средств перехвата информации (приёмника). Среди различных видов ТКУИ в данной работе рассматриваются акустические каналы утечки информации в помещении, где в качестве передаваемого сигнала выступает акустическая речевая информация.

Для оценки защищенности информации от утечки по техническим каналам в помещении, которое предназначено для проведения закрытых переговоров, чаще всего используется общепринятая инструментально-расчетная методика, предложенная Хоревым А.А., Желязняком В.К. и Макаровым Ю.К., основанная на результатах экспериментальных исследований Н.Б. Покровского, называемых формантным методом. Критерием оценки защищенности акустической речевой информации является значение разборчивости речи (W). В данном случае под оценкой разборчивости речи понимается определение доли речевой информации, которая может быть воспринята злоумышленником за пределами помещения, где ведутся закрытые переговоры.

Общепринятая инструментально-расчетная методика, по сути, является частным случаем методики Н.Б. Покровского, а именно:

- 1) рассматривается только словесная разборчивость W ;
- 2) измерение сигналов производится в октавных полосах частот;
- 3) сигнал измеряется на расстоянии 1 м от микрофона (в исследованиях Н.Б. Покровского - 8 см).

Однако условия проведения экспериментов в исследованиях Н.Б. Покровского (оценивалось качество телефонных линий связи) существенно отличались от современных условий защиты информации, а также от современных возможностей средств измерительной техники и обработки сигналов. Поскольку общепринятая инструментально-расчетная методика является частным случаем методики Покровского Н.Б., она также сталкивается с рядом ограничений при применении в современных условиях. Основные из этих ограничений:

- 1) не рассматривались условия с высоким уровнем маскирующих шумов, что оказывает влияние на зависимости восприятия речевых сигналов;
- 2) не учитывались различия в спектрах речи дикторов, все основывалось на усредненном спектре речи;
- 3) маскировка речи учитывалась только с помощью помех на основе белого шума (с изменением спектра путем применения фильтров), исключая возможность оценки защищенности с помощью помех типа «речевой хор» (помеха, состоящая из смеси звуков речи, формируемая по определенным правилам, описанным в работах Авдеева В.Б., Зельманского О.Б., Трушина В.А. и др).

Учесть перечисленные недостатки становится возможным при использовании для оценки защищенности акустической речевой информации методики на основе многоканальной системы сверточных нейронных сетей. В связи с этим была сформулирована **научная задача**, заключающаяся в разработке и обосновании методики оценки защищенности акустической речевой информации при помощи многоканальной системы сверточных нейронных сетей в помещениях, предназначенных для проведения закрытых переговоров, с учётом высокого уровня маскирующих помех, различия в спектрах речи дикторов, различных спектров помех на основе белого шума, а также помехи типа «речевой хор».

Стоит отметить, что в работе рассматривается случай с проведением оценки защищенности акустической речевой информации только с применением средств активной защиты (САЗ), генерирующих маскирующие помехи.

Проведенные в ходе работы исследования показали, что предложенная методика оценки защищенности акустической речевой информации является полноценным решением поставленной научной задачи.

Степень проработки темы исследования. Вопросам оценки защищенности акустической речевой информации посвящены работы таких отечественных и зарубежных исследователей, как: Хорев А.А., Железняк В.К., Макаров Ю.К., Покровский Н.Б., Авдеев В.Б., Дворянкин С.В., Иванов А.В., Козлачков С.Б., Трушин В.А., Тупота В.И., Claudio S., Hasegawa-Johnson M., Houtgast T., Killion M., Lu L., Mueller G., Shifas M., Steeneken H., Stylianou Y., Verhave J. и многих других. Для применения общепринятой методики на практике используют инструментальный подход, основанный на измерении интегральных уровней в октавных полосах тестовых сигналов, фоновых шумов и шумов от САЗ, создающих маскирующие помехи, а также акустического сигнала за ограждающей конструкцией в виде смеси тестового сигнала и шума. В соответствии с зависимостями, полученными экспериментальным путем Покровским Н.Б., производится расчет разборчивости речи и, на её основе, оценка защищенности акустической речевой информации. Использование методов искусственного интеллекта позволяет перейти от указанной методики с измерением уровней сигнала и шума к методам соответствия разборчивости речи распределению энергии речевого и маскирующего сигналов в акустическом спектре сигнала. Существующие обученные алгоритмы распознавания речи для решения поставленной задачи не подходят так как они не предназначены для работы в условиях высоких уровней шумов (отношение сигнал/шум меньше -10 дБ).

Цель диссертационной работы: Повышение достоверности оценки защищенности акустической речевой информации за счет учета реальных условий, а именно высокого уровня маскирующих помех, различия в спектрах речи дикторов, различных спектров помех на основе белого шума, а также помехи типа «речевой хор» с использованием методики, основанной на многоканальной системе сверточных нейронных сетей.

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

1. Провести анализ современного состояния исследований в области оценки защищенности акустической речевой информации.
2. Разработать методику предобработки данных для оценки разборчивости акустической речевой информации с применением многоканальной системы сверточных нейронных сетей.
3. Разработать модель оценки разборчивости акустической речевой информации на основе многоканальной системы сверточных нейронных сетей.
4. Разработать методику оценки защищенности акустической речевой информации на основе анализа распределения энергии речевого и маскирующего сигнала в акустическом спектре.
5. Разработать структуру программно-аппаратного комплекса оценки защищенности акустической речевой информации на основе многоканальной системы сверточных нейронных сетей.

Объектом исследования является акустическая речевая информация.

Предметом исследования являются методики оценки защищенности акустической речевой информации на основе спектрограмм и графиков мел-частотных кепстральных коэффициентов с использованием сверточных нейронных сетей.

Методы исследования. Применялись положения теории разборчивости речи, теории измерений, методы обработки речевых сигналов, классификации и идентификации образов, глубокого обучения, теории вероятностей и математической статистики, распознавания образов, аппарат искусственных нейронных сетей (ИНС).

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

Научную новизну работы составляют:

1. Методика предобработки данных для оценки разборчивости акустической речевой информации, отличающаяся созданием зашумленных аудиозаписей речи со множеством незначительных стохастических вариаций в амплитуде шумового сигнала, а также классификацией аудиозаписей по уровням разборчивости речи на основе распределения энергии речевого и маскирующего сигнала в акустическом спектре, что позволяет сформировать наборы данных для обучения многоканальной системы сверточных нейронных сетей при оценке защищенности акустической речевой информации от утечки по техническим каналам.

2. Модель оценки разборчивости акустической речевой информации на основе многоканальной системы сверточных нейронных сетей, которая в отличие от традиционных подходов использует отдельные каналы обработки спектрограмм и графиков мел-частотных кепстральных коэффициентов, позволяя учитывать высокий уровень маскирующих помех, различные спектры помех на основе белого шума, помехи типа «речевой хор» и различия в спектрах речи дикторов при определении уровня разборчивости речи на основе мажорантной оценки.

3. Методика оценки защищенности акустической речевой информации, отличающаяся от общепринятой инструментально-расчетной методики переходом от измерения интегральных уровней в октавных полосах акустических сигналов к анализу распределения энергии речевого и маскирующего сигнала в акустическом спектре, а также наличием тестового сигнала в виде последовательности аудиозаписей речи дикторов, что обеспечивает повышение достоверности оценки в наиболее значимом диапазоне показателя защищенности акустической речевой информации.

Теоретическая значимость диссертационной работы заключается в методике предобработки данных для оценки разборчивости акустической речевой информации, разработке модели оценки разборчивости акустической речевой информации на основе многоканальной системы сверточных нейронных сетей, разработке методики оценки защищенности акустической речевой информации.

Практическая значимость работы заключается в разработке программно-аппаратного комплекса, позволяющего проводить оценку защищенности помещений при проведении закрытых переговоров в условиях высокого уровня маскирующих помех, воздействия различных спектров помех на основе белого шума и помехи типа «речевой хор», а также учета различия в спектрах речи дикторов. Проведённые испытания подтвердили точность распознавания программно-аппаратным комплексом словесной разборчивости речи (более 90%) и достоверность предложенной методики в реальных условиях эксплуатации. Разработанный комплекс может использоваться как в учебном процессе при подготовке специалистов по защите информации, так и в практической — организациями - лицензиатами ФСТЭК России, профессионально занимающимися оценкой защищенности информации, разработкой методик, созданием программно-аппаратных комплексов оценки защищенности акустической речевой информации, а также организациями, занимающимися проектированием переговорных комнат.

Положения, выносимые на защиту:

1. Методика предобработки данных для оценки разборчивости акустической речевой информации, позволяющая сформировать наборы данных для обучения многоканальной системы сверточных нейронных сетей при оценке защищенности акустической речевой информации от утечки по техническим каналам (**пп. 10, 11 паспорта специальности 2.3.6**).

2. Модель оценки разборчивости акустической речевой информации на основе многоканальной системы сверточных нейронных сетей, позволяющая учитывать высокий уровень маскирующих помех, различные спектры помех на основе белого шума, помехи типа «речевой хор» и различия в спектрах речи дикторов при определении уровня разборчивости речи на основе мажорантной оценки (**пп. 10, 11 паспорта специальности 2.3.6**).

3. Методика оценки защищенности акустической речевой информации, позволяющая повысить достоверность оценки защищенности и обеспечить отклонение значения защищенности от истинного значения (результаты субъективных испытаний) не более 4% в наиболее значимом диапазоне значений показателя оценки защищенности. Возможность учета помехи типа «речевой

хор» при оценке защищенности позволяет расширить область применения данной методики (**пп. 10, 11 паспорта специальности 2.3.6**).

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

Апробация результатов работы. О полученных результатах работы докладывалось на следующих конференциях: XIV Всероссийская научная конференция молодых ученых «Наука. Технологии. Инновации-2020» (г. Новосибирск); XXI всероссийский конкурс-конференция студентов и аспирантов по информационной безопасности «SIBINFO-2021» (г. Томск); XV Международная научно-техническая конференция «Динамика систем, механизмов и машин» (г. Омск); XV Всероссийская научная конференция молодых ученых «Наука. Технологии. Инновации-2021» (г. Новосибирск); 10-ая Всероссийская молодежная школа-семинар по проблемам информационной безопасности «ПЕРСПЕКТИВА - 2023» (г. Красноярск); XXIV Всероссийский конкурс-конференция студентов и аспирантов по информационной безопасности «SIBINFO-2024» (г. Томск); VII Всероссийская молодежная научно-практическая конференция с международным участием «Информационные технологии обеспечения комплексной безопасности в цифровом обществе» (г. Уфа); AISMA-2024: International Workshop on Advanced in Information Security Management and Applications (г. Ставрополь).

Соответствие паспорта специальности. Тема и содержание диссертации соответствуют паспорту специальности 2.3.6. Методы и системы защиты информации, информационная безопасность, пункту 10: «Модели и методы оценки защищенности информации и информационной безопасности объекта», а также пункту 11: «Модели и методы оценки эффективности систем (комплексов), средств и мер обеспечения информационной безопасности объектов защиты».

Публикации. По теме диссертации лично и в соавторстве опубликовано 11 печатных работ, 4 из которых изданы в журналах, рекомендованных ВАК; 2 научные публикации индексируются в международной информационно-аналитической системе научного цитирования Scopus.

Объем и структура работы. Диссертация состоит из введения, 6 глав, заключения, списка литературы (123 наименования) и 5 приложений. Общий объем диссертации составляет 172 страницы, включающих в себя 21 таблицу и 27 рисунков.

ОСНОВНЫЕ ПОЛОЖЕНИЯ РАБОТЫ

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

Первая глава посвящена анализу современного состояния исследований в области оценки защищенности акустической речевой информации.

Рассмотрены отечественные и зарубежные методики оценки защищенности помещений от утечки речевой информации по акустическим каналам. Среди отечественных выделяется инструментально-расчетная методика, основанная на результатах экспериментальных исследований Н.Б. Покровского, а среди зарубежных методов оценки разборчивости речи выделяются Articulation Index (AI), Speech Transmission Index (STI) и другие.

Одним из перспективных направлений в области обработки акустических сигналов является использование нейронных сетей. Их применение позволяет реализовать адаптивные подходы, устойчивые к изменяющимся условиям и к работе в диапазоне низкого отношения сигнал/шум. При этом необходимо обеспечить высокую точность распознавания при отношении сигнал/шум ниже -10 дБ, определить формат представления данных акустических сигналов для обучения и выбрать наиболее подходящую архитектуру нейронной сети для оценки разборчивости речи.

В качестве входных данных для нейросетевых систем обработки акустических сигналов, как правило, используются частотно-временные представления сигналов в виде спектрограмм и графиков мел-частотных кепстральных коэффициентов (MFCC). Эти параметры позволяют отразить спектральную структуру речи, её динамику и особенности в условиях различных шумов,

что делает их удобными для последующего анализа нейросетевыми моделями. Для реализации инструмента определения соответствия значения разборчивости речи доли примеси шумового сигнала в речевом сигнале и распределения энергии по акустическому спектру записи предлагается применить сверточную нейронную сеть, которая работает с изображениями.

В связи с этим представляется актуальной задача разработки методики оценки защищенности акустической речевой информации на основе многоканальной системы сверточных нейронных сетей с использованием в качестве входных данных спектрограмм либо графиков MFCC в виде файлов изображений для последующего анализа разборчивости в зашумленной аудиозаписи речи.

Во второй главе рассматривается методика предобработки данных для оценки разборчивости акустической речевой информации (рис.1).

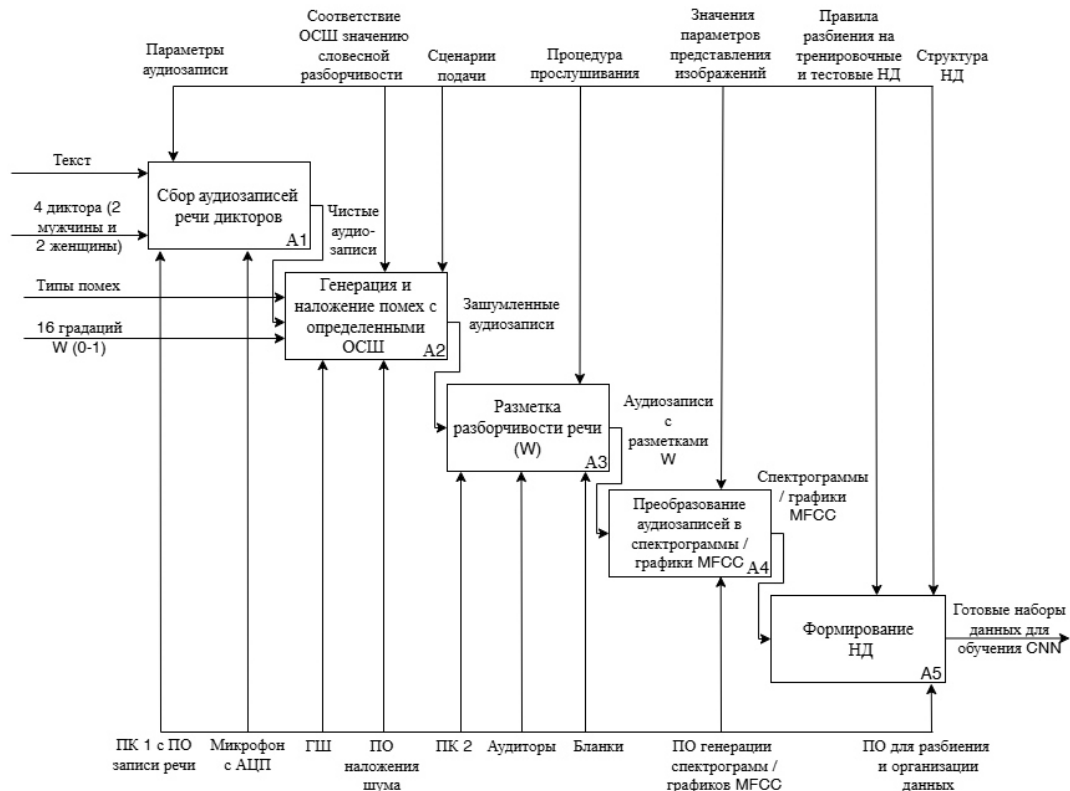


Рисунок 1. Методика предобработки данных для оценки разборчивости акустической речевой информации

Для создания обучающих выборок для обучения сверточных нейронных сетей сначала были сформированы в формате .wav аудиозаписи дикторов с наложением белого шума, розового шума, формантоподобной помехи и помехи типа «речевой хор». Для учёта различия в спектрах речи дикторов использовались 4 вида голосов – мужские и женские голоса с низким и высоким тембрами. Длительность каждой аудиозаписи составила 30 секунд - этого времени достаточно для корректного распознавания распределения энергии в сигнале сверточной нейронной сетью, что подтверждено экспериментально. Для записи использовался измерительный микрофон с предусилителем, с модулем АЦП для акустики и программное обеспечение, имеющее функционал обработки звука.

На этапе наложения помех на аудиозаписи речи дикторов было сгенерировано по 720 вариаций каждой из помех на основе белого шума, различающихся незначительными стохастическими флуктуациями амплитуды, характерными для случайных сигналов. Эти вариации обеспечивают дополнительное разнообразие обучающих данных. Каждый вариант помехи накладывался отдельно на аудиозаписи речи с различным отношением сигнал/шум, соответствующим определённым градациям разборчивости речи. При определении отношения сигнал/шум предварительно удалялись паузы между словами в аудиозаписи речи. Соотношение сигнал/шум к градации разборчивости определялось экспериментально при прослушивании зашумленных

записей 6 аудиторами, учитывая требования ГОСТ Р 50840–95 и ГОСТ 16600–72. В результате аудиозаписи были проградуированы по шкале разборчивости речи (доля акустической речевой информации, которая может быть правильно воспринята злоумышленником) в градациях от 0 до 0,50 с шагом в 0,05 и от 0,50 до 1 с шагом в 0,10.

Аналогичным образом были сформированы вариации помехи типа «речевой хор». В течение 1 минуты дикторы читали разные связные тексты. Каждая помеха создавалась путем наложения 6 голосов (3 мужских и 3 женских) с временной задержкой. Всего было создано 720 вариантов «речевого хора», совпадающих по количеству с наборами данных для других видов помех. Полученные зашумленные аудиозаписи проградуированы по той же шкале разборчивости речи, что и при наложении помех на основе белого шума. В итоге, для каждой комбинации типа голоса и помехи было сформировано по 11 520 зашумленных аудиозаписей, при этом для прослушивания аудиторами выбирались лишь отдельные записи, достаточные для оценки всех условий эксперимента.

Были рассмотрены параметры, используемые при формировании обучающей выборки, представленные в табл. 1. Для определения значений параметров для обучения сверточной нейронной сети в каждом варианте обучающей выборки изменяли только одно значение параметра данных из диапазона значений. Удовлетворяющими требованиям считались те значения, при которых во время обучения сверточной нейронной сети точность распознавания нейронной сетью была более или равна 90%.

Таблица 1. Параметры, используемые для формирования обучающей выборки

Параметр	Значения
Длительность аудиозаписи	от 5 до 30 секунд с шагом в 5 секунд
Параметры спектрального анализа	база БПФ для спектрограмм - 1024, 2048, 4096, 8192, 16384; параметры спектральной огибающей (фреймы) для MFCC - от 10 до 40 мс с шагом в 10 мс
Размер изображения в пикселях	от 100х100 пикселей до 1200х1200 пикселей с шагом в 100х100 пикселей
Тип масштабирования генерируемых спектрограмм и графиков MFCC	линейное или логарифмическое масштабирование частоты

В процессе формирования набора данных для обучения на основе зашумленных аудиозаписей были сгенерированы спектрограммы и графики MFCC в формате .png. В каждом наборе данных содержалось 11520 изображений, из которых 9216 использовались для обучения, а 2304 – для тестирования.

Были подробно рассмотрены существующие архитектуры сверточных нейронных сетей и выделена архитектура ResNet, обладающая параметрами, которые были выбраны на основе рекомендаций её создателей, предложивших такие конфигурации для улучшения обучения глубоких нейронных сетей и предотвращения исчезновения градиента, что подтверждено их успешным применением в задачах классификации изображений.

После проведения обучений сверточных нейронных сетей с варьированием параметров обучающих выборок был выполнен анализ результатов, по итогам которого выбраны значения параметров обучающих выборок. В результате получены 32 сверточные нейронные сети, обученные на наборах данных, сформированных из зашумленных аудиозаписей мужской и женской речи с низким и высоким тембрами:

- 1) 16 сверточных нейронных сети ResNet, обученных на спектрограммах;
- 2) 16 сверточных нейронных сети ResNet, обученных на графиках MFCC.

В качестве примера в табл. 2 показаны результаты обучения сверточных нейронных сетей и используемые в них значения параметров обучающих выборок, сформированных на основе зашумленных белым шумом и «речевым хором» аудиозаписей мужской речи с низким тембром.

Таблица 2. Результаты обучения сверточных нейронных сетей

Мужская речь с низким тембром					
Графики MFCC (белый шум)					
Длительность аудиозаписи, с	Длина фрейма, мс	Размер изображения, пикселей	Тип масштабирования изображения	Точность распознавания нейронной сетью, %	Коэффициент потерь при обучении
10	20	600х600	Логарифмическое по частоте	97,61	0,1543

Таблица 2. Результаты обучения сверточных нейронных сетей

<i>Мужская речь с низким тембром</i>					
<i>Спектрограммы («речевой хор»)</i>					
Длительность аудиозаписи, с	База БПФ, точек	Размер изображения, пикселей	Тип масштабирования изображения	Точность распознавания нейронной сетью, %	Коэффициент потерь при обучении
15	2048	600x600	Логарифмическое по частоте	96,13	0,1323

Анализ полученных результатов показал, что при выбранных значениях параметров обучающих выборок достигаются более высокие значения метрики точности распознавания, используемой для оценки качества работы сверточных нейронных сетей, по сравнению с другими протестированными вариантами. Высокие значения метрики подтверждают корректность выбора параметров обучающих выборок, способов представления и визуализации данных, а также настроек процесса обучения.

В табл. 3 представлены результаты проверки обученных сверточных нейронных сетей на тестовых наборах данных для каждого вида зашумленной аудиозаписи речи диктора.

Таблица 3. Результаты проверки обученных сверточных нейронных сетей

Вид голоса	Наложённый шум	Метод визуализации аудиозаписи	Точность распознавания нейронной сетью, %	Коэффициент потерь при обучении
Мужской с низким тембром	Белый	График MFCC	97,61	0,1543
	Розовый	График MFCC	96,16	0,1812
	Формантоподобная помеха	График MFCC	97,01	0,1974
	«Речевой хор»	Спектрограмма	96,13	0,1323
Мужской с высоким тембром	Белый	График MFCC	95,33	0,1421
	Розовый	График MFCC	94,45	0,1541
	Формантоподобная помеха	График MFCC	95,17	0,1497
	«Речевой хор»	Спектрограмма	91,1	0,2192
Женский с низким тембром	Белый	График MFCC	94,54	0,1671
	Розовый	График MFCC	96,64	0,1212
	Формантоподобная помеха	График MFCC	93,02	0,1888
	«Речевой хор»	Спектрограмма	91,13	0,3213
Женский с высоким тембром	Белый	График MFCC	94,29	0,1946
	Розовый	График MFCC	92,15	0,2131
	Формантоподобная помеха	График MFCC	94,89	0,1501
	«Речевой хор»	Спектрограмма	91,87	0,2752

Сравнительный анализ различных наборов данных показал, что при использовании набора данных с помехами на основе белого шума модель, использующая графики MFCC, достигала более высоких значений метрики точности распознавания по сравнению с моделью, использующей спектрограммы. Напротив, для набора данных с помехами типа «речевой хор» более высокие значения метрики точности распознавания демонстрировала модель, использующая спектрограммы, по сравнению с моделью, использующей графики MFCC. Во всех приведенных выше случаях также учитывалось значение коэффициента потерь при обучении, которое находилось в допустимых границах (не более 0,33), так как при его увеличении наблюдается переобучение модели сверточной нейронной сети, приводящее к снижению значений метрики точности распознавания.

Таким образом, при разработке методики оценки защищенности акустической речевой информации на основе анализа распределения энергии речевого и маскирующего сигнала в акустическом спектре в случае применения различных спектров помех на основе белого шума будут использоваться графики MFCC, а в случае применения помех типа «речевой хор» будут использоваться спектрограммы.

В третьей главе рассмотрена модель оценки разборчивости акустической речевой информации на основе многоканальной системы сверточных нейронных сетей.

Пусть у нас есть аудиозапись речи диктора α , представленная в виде изображения:

$$\alpha \in A(m, n): m \in M, n \in N, \quad (1)$$

где M – множество спектрограмм, N – множество графиков MFCC.

Каждой аудиозаписи речи диктора, представленной в виде изображения, можно поставить в соответствие элемент ω множества Ω , так как каждой аудиозаписи соответствует уникальная комбинация параметров:

$$\Omega = M \times N \times D \times H, \quad (2)$$

где $D = \{d_1, d_2, d_3, d_4\}$ – множество голосов дикторов; $H = \{h_1, h_2, h_3, h_4\}$ – множество типов маскирующих помех.

Заметим, что во время реальных испытаний невозможно точно определить, какая конкретно помеха использовалась для зашумления речевого сигнала, более того, ввиду некорректной настройки САЗ, действительная используемая помеха может не соответствовать ни одному элементу из множества H . Для решения данной проблемы используется оператор

$$T: M \rightarrow H, \quad (3)$$

находящий аппроксимацию $\hat{h}$ реальной помехи h и реализованный на практике в виде сверточной нейронной сети. Таким образом в дальнейшем для описания записи речи диктора α в множестве Ω будем использовать $\hat{h}$.

Введем $R = \{r_1, r_2, \dots, r_{16}\}$ – множество дискретных уровней разборчивости акустической речевой информации, задающих градацию уровней разборчивости W (от 0 до 0,50 с шагом в 0,05 и от 0,50 до 1 с шагом в 0,10). Диапазон меньше 0,50 разборчивости акустической речевой информации является наиболее значимым диапазоном значений показателя оценки защищенности акустической речевой информации. Это обусловлено тем, что составление краткой справки-аннотации о содержании перехваченного разговора невозможно при разборчивости от 0,40 до 0,50, при разборчивости от 0,10 до 0,39 значительно затруднено установление даже предмета ведущегося разговора, а при разборчивости менее 0,10 это практически невозможно даже при использовании современной техники фильтрации помех. Диапазон более 0,50 разборчивости акустической речевой информации менее требователен к дискретности оценки, потому как очевидно, что в данном случае информация не защищена, и оценка разборчивости акустической речевой информации в этом диапазоне необходима оператору только чтобы внести соответствующие корректировки при настройке САЗ.

Тогда для оценки уровня разборчивости акустической речевой информации в одних и тех же условиях будет достаточно использовать 4 аудиозаписи речи дикторов $\omega_i, i = \overline{1,4}$ одного и того же текста, записанные при помощи всех возможных дикторов из множества D . Будем считать, что помеха для них одинакова и равна $\hat{h}$, полученной при анализе тестовой аудиозаписи с установленной помехой без диктора.

Для оценки уровня разборчивости используются 4 сверточные нейронные сети F_k по одной для каждого вида помехи, причем стоит учесть, что нейронная сеть определяющая разборчивость для аудиозаписей, зашумленных при помощи помехи типа «речевой хор» использует в качестве входных данных спектрограммы, а остальные – графики MFCC:

$$F_k: N \rightarrow R^{16}, k \in \{1,2,3\}, \quad (4)$$

$$F_4: M \rightarrow R^{16}, \quad (5)$$

где выходной элемент p представляет собой вектор вероятностей того, к какому уровню разборчивости соответствует анализируемая запись:

$$p = (p_1, p_2, \dots, p_{16}), \sum_{i=1}^{16} p_i = 1, p_i \geq 0. \quad (6)$$

Выбранная в соответствии с используемой помехой нейронная сеть f_k анализирует записи $\omega_i, i = \overline{1,4}$ и возвращает 4 вектора p^j , после чего выбирается индекс i_{max} , соответствующий мажорантной оценке уровня разборчивости $r_{i_{max}}$ из наиболее вероятных в векторе p^j :

$$i_{max}^j = \arg \max_{i=\overline{1,16}} p_i^j \quad (7)$$

$$i_{max} = \max(i_{max}^j), j = \overline{1,4}. \quad (8)$$

Решающее правило определения разборчивости имеет вид:

$$r_{i_{max}} = W_{i_{max}}, \quad (9)$$

что позволяет получить достоверную оценку уровня разборчивости акустической речевой информации W . Такой механизм выбора логично вытекает из концепции оценки защищенности

через наихудший сценарий, хотя в публикациях многих исследователей (Хорев А.А., Сагдеев К.М. и др.) прямое указание на это отсутствует, но подразумевается методикой расчёта показателей защищенности через крайние условия.

Таким образом, разработанная модель оценки разборчивости акустической речевой информации на основе многоканальной системы сверточных нейронных сетей позволит реализовать новую методику оценки защищенности акустической речевой информации.

В четвертой главе предложена методика оценки защищенности акустической речевой информации, основанная на анализе распределения энергии речевого и маскирующего сигнала в акустическом спектре, отличающаяся также наличием тестового сигнала в виде последовательности аудиозаписей речи дикторов. При её использовании, как и в общепринятой инструментально-расчётной методике, оператор проводит оценку разборчивости акустической речевой информации в контрольных точках помещения.

Основное отличие предлагаемой методики заключается в используемом тестовом сигнале и способе анализа полученных данных: вместо генератора тестовых сигналов применяется источник речевого тестового сигнала – последовательность аудиозаписей 4 дикторов (мужских и женских голосов с низким и высоким тембрами), обеспечивающий условия, максимально приближенные к реальной речевой информации, а роль измерительного устройства выполняет модуль оценки разборчивости на основе сверточной нейронной сети. Такой подход позволяет определять значение разборчивости речи W по спектрограммам или графикам MFCC на основе анализа распределения энергии речевого и маскирующего сигнала в акустическом спектре. На рис. 2 представлена предлагаемая методика оценки защищенности акустической речевой информации с применением многоканальной системы сверточных нейронных сетей в нотации IDEF0.

Показатель защищенности акустической речевой информации p необходимо получить из определенного ранее значения разборчивости речи при помощи следующего выражения:

$$p = 1 - W_{max}, \quad (10)$$

где W_{max} - мажорантная оценка уровня разборчивости речи.

Для оценки защищенности акустической речевой информации используется пороговая функция:

$$y = \begin{cases} 1, & \text{если } p > p_p, \\ 0, & \text{если } p \leq p_p, \end{cases} \quad (11)$$

где 1 – ограждающая конструкция защищена от утечки акустической речевой информации, 0 – ограждающая конструкция не защищена от утечки акустической речевой информации, p - показатель защищенности акустической речевой информации, полученный из определенного ранее значения разборчивости речи, p_p – порог защищенности акустической речевой информации. Значение порога защищенности выбирается исходя из установленных требований к уровню безопасности и условий проведения оценки. В различных исследованиях приводятся разные варианты пороговых значений. Так, например, основываясь на работе Хорева А.А., приведена подобная градация в значениях разборчивости речи для наиболее значимого диапазона значений показателя оценки защищенности акустической речевой информации. Можно произвести расчет аналогичной таблицы в значениях пороговой защищенности для наиболее значимого диапазона (табл. 4):

Таблица 4. Диапазоны значений пороговой защищенности

Цель защиты помещений от утечки акустической речевой информации	Значения разборчивости речи, W	Значения пороговой защищенности, p_p
Невозможно установление предмета ведущегося разговора при использовании техники фильтрации помех	0 – 0,10	0,90 – 1
Невозможно установление предмета ведущегося разговора в помещении	0,11 – 0,20	0,80 – 0,89
Затруднено установление предмета ведущегося разговора в помещении	0,21 – 0,40	0,60 – 0,79
Невозможно составить краткую справку-аннотацию о содержании перехваченного разговора	0,41 – 0,50	0,50 – 0,59

Таким образом, разработанная методика была реализована с помощью программно-аппаратного комплекса оценки защищенности акустической речевой информации на основе многоканальной системы сверточных нейронных сетей.

1) Субъективная оценка разборчивости акустической речевой информации, путем прослушивания 6 аудиторами за ограждающей конструкцией тестового сигнала с использованием САЗ (результаты приведены в табл. 5 и 6);

2) Инструментально-расчетная оценка разборчивости акустической речевой информации с использованием САЗ (результаты приведены в табл. 5 и 6).

Была проведена статистическая обработка результатов, которая включала расчёт математического ожидания, выборочного среднеквадратичного отклонения среднего, определения доверительных интервалов при доверительной вероятности 95 % для 1152 измерений. Полученные оценки разборчивости речи обеспечивают несмещённую оценку стандартной ошибки среднего даже при ограниченном числе измерений и подтверждают корректность выполненных измерений.

Из табл. 5 и 6 видно, что результаты разработанной методики оказались приближены более к результатам субъективной оценки, нежели чем к инструментально-расчетной оценке по формантному методу.

Таблица 5. Результаты измерений разборчивости речи при зашумлении белым шумом и розовым шумом

№ п/п	Контрольные точки помещения	Разборчивость речи, W					
		Белый шум			Розовый шум		
		При проведении субъективной оценки	При проведении инструментально-расчетной оценки	При использовании разработанной методики	При проведении субъективной оценки	При проведении инструментально-расчетной оценки	При использовании разработанной методики
1	Дверь между помещением и коридором	0,36	0,14	0,40	0,32	0,27	0,35
2	Стена между помещением и коридором	0,28	0,12	0,30	0,27	0,25	0,30
3	Стена между помещением и соседним кабинетом	0,26	0,11	0,30	0,24	0,24	0,25
4	Вентиляция	0,32	0,13	0,35	0,29	0,26	0,30

Таблица 6. Результаты измерений разборчивости речи при зашумлении формантоподобной помехой и помехой типа «речевой хор»

№ п/п	Контрольные точки помещения	Разборчивость речи, W					
		Формантоподобная помеха			«Речевой хор»		
		При проведении субъективной оценки	При проведении инструментально-расчетной оценки	При использовании разработанной методики	При проведении субъективной оценки	При проведении инструментально-расчетной оценки	При использовании разработанной методики
1	Дверь между помещением и коридором	0,24	0,27	0,25	0,12	0,75	0,10
2	Стена между помещением и коридором	0,17	0,22	0,20	0,09	0,73	0,10
3	Стена между помещением и соседним кабинетом	0,19	0,24	0,20	0,06	0,68	0,05
4	Вентиляция	0,22	0,25	0,25	0,11	0,74	0,10

В соответствии с (10) значение показателя защищенности акустической речевой информации определяется на основании значения разборчивости речи. Соответственно, все выше рассмотренные методики необходимо сравнивать по значениям показателя защищенности.

Необходимо провести анализ достоверности оценки защищенности акустической речевой информации при помощи разработанной методики в наиболее значимом диапазоне значений показателя оценки защищенности. Как было показано ранее в табл. 4 данный диапазон разделен на 5 поддиапазонов значений: от 0,90 до 1; от 0,80 до 0,89; от 0,70 до 0,79; от 0,60 до 0,69; от 0,50 до 0,59. Следовательно, необходимо дополнительно получить результаты оценки значений разборчивости речи W , соответствующие каждому из указанных поддиапазонов показателя оценки защищенности, что достигается путем регулировки уровня маскирующей помехи при помощи САЗ.

Также, для анализа достоверности оценки защищенности акустической речевой информации при помощи разработанной методики используются следующие показатели:

1) погрешности $\Delta p_{\text{ИРО}}$ (12) и $\Delta p_{\text{ПАК}}$ (13):

$$\Delta p_{\text{ИРО}} = |p_{\text{СУБ}} - p_{\text{ИРО}}|, \quad (12)$$

$$\Delta p_{\text{ПАК}} = |p_{\text{СУБ}} - p_{\text{ПАК}}|, \quad (13)$$

где $p_{\text{СУБ}}$ – показатель защищенности при проведении субъективной оценки защищенности акустической речевой информации;

$p_{\text{ИРО}}$ – показатель защищенности при проведении общепринятой инструментально-расчетной оценки защищенности акустической речевой информации;

$p_{\text{ПАК}}$ – показатель защищенности при проведении оценки защищенности акустической речевой информации при помощи разработанной методики.

2) значения достоверности $V_{\text{ИРО}}$ (14) и $V_{\text{ПАК}}$ (15):

$$V_{\text{ИРО}} = \left(1 - \frac{\Delta p_{\text{ИРО}}}{p_n}\right) * 100\%, \text{ при } \frac{\Delta p_{\text{ИРО}}}{p_n} \leq 1, \quad (14)$$

$$V_{\text{ПАК}} = \left(1 - \frac{\Delta p_{\text{ПАК}}}{p_n}\right) * 100\%, \text{ при } \frac{\Delta p_{\text{ПАК}}}{p_n} \leq 1, \quad (15)$$

где $\Delta p_{\text{ИРО}}$ – погрешность, полученная при использовании инструментально-расчётной методики, $\Delta p_{\text{ПАК}}$ – погрешность, полученная при использовании разработанной методики, $p_n = 1$ – нормирующее значение (верхний предел шкалы значений защищенности речевой акустической информации).

В качестве истинного значения принимается значение показателя защищенности $p_{\text{СУБ}} = 1 - W_{\text{СУБ}}$, где $W_{\text{СУБ}}$ – значение разборчивости акустической речевой информации в результате субъективных испытаний.

В табл. 7 и 8, как пример, приведены результаты численной оценки защищенности помещения в случаях зашумления белым шумом и помехой типа «речевой хор»:

Таблица 7. Анализ достоверности оценки защищенности акустической речевой информации при помощи разработанной методики (случай с зашумлением белым шумом)

№ п/п	Диапазоны значений показателя защищенности p	Белый шум									
		$W_{\text{СУБ}}$	$W_{\text{ИРО}}$	$W_{\text{ПАК}}$	$p_{\text{СУБ}}$	$p_{\text{ИРО}}$	$p_{\text{ПАК}}$	$\Delta p_{\text{ИРО}}$	$\Delta p_{\text{ПАК}}$	$V_{\text{ИРО}}, \%$	$V_{\text{ПАК}}, \%$
1	0,90-1	0,04	0,07	0,05	0,96	0,93	0,95	0,03	0,01	97	99
2	0,80-0,89	0,17	0,10	0,20	0,83	0,90	0,80	0,07	0,03	93	97
3	0,70-0,79	0,26	0,11	0,30	0,74	0,89	0,70	0,15	0,04	85	96
4	0,60-0,69	0,36	0,14	0,40	0,64	0,86	0,60	0,22	0,04	78	96
5	0,50-0,59	0,47	0,17	0,50	0,53	0,83	0,50	0,30	0,03	70	97

Таблица 8. Анализ достоверности оценки защищенности акустической речевой информации при помощи разработанной методики (случай с зашумлением помехой типа «речевой хор»)

№ п/п	Диапазоны значений показателя защищенности p	«Речевой хор»									
		$W_{\text{СУБ}}$	$W_{\text{ИРО}}$	$W_{\text{ПАК}}$	$p_{\text{СУБ}}$	$p_{\text{ИРО}}$	$p_{\text{ПАК}}$	$\Delta p_{\text{ИРО}}$	$\Delta p_{\text{ПАК}}$	$V_{\text{ИРО}}, \%$	$V_{\text{ПАК}}, \%$
1	0,90-1	0,06	0,68	0,05	0,94	0,32	0,95	0,62	0,01	38	99
2	0,80-0,89	0,12	0,75	0,10	0,88	0,25	0,90	0,63	0,02	37	98
3	0,70-0,79	0,23	0,81	0,25	0,77	0,19	0,75	0,58	0,02	42	98
4	0,60-0,69	0,33	0,84	0,35	0,67	0,16	0,65	0,51	0,02	49	98
5	0,50-0,59	0,48	0,86	0,50	0,52	0,14	0,50	0,38	0,02	62	98

Сравнительный анализ результатов оценки защищенности акустической речевой информации, приведённых в табл. 7 и 8, показал, что разработанная методика обеспечивает более достоверные значения показателя защищенности по сравнению с общепринятой инструментально-расчётной методикой. В частности, при воздействии маскирующей помехи типа «речевой хор» общепринятая методика завышает оценку защищенности, поскольку интерпретирует данный тип помехи как белый шум с огибающей спектра речи, что приводит к искажению результатов оценки.

Разработанная методика, напротив, позволяет получить значения показателя защищенности, более точно отражающие действительное состояние защищённости акустической речевой информации и выявляющие случаи потенциальной утечки информации. Из результатов видно, что разработанная методика имеет большую достоверность по сравнению с общепринятой инструментально-расчётной методикой. Максимальное отклонение показателя защищенности акустической речевой информации, полученного при помощи разработанной методики, от значения, полученного в ходе субъективных испытаний, составило не более 4% в наиболее значимом диапазоне значений показателя оценки защищенности.

Таким образом, полученные результаты подтверждают адекватность разработанной методики и её практическую применимость в реальных условиях эксплуатации. Данная методика учитывает недостатки общепринятой инструментально-расчётной методики, а именно: позволяет учитывать условия с высоким уровнем маскирующих помех, учитывает различия в спектрах речи дикторов, а также учитывает различные спектры помех на основе белого шума, а также помеху типа «речевой хор».

В заключении приводятся основные результаты и выводы, полученные автором.

В приложениях приведены программные коды на языке программирования Python для обучения архитектуры сверточной нейронной сети ResNet на наборах данных, статистическая обработка результатов измерений разборчивости речи рассмотренными методиками, а также акты внедрения.

ОСНОВНЫЕ ВЫВОДЫ И РЕЗУЛЬТАТЫ РАБОТЫ

1. Разработана методика предобработки данных для оценки разборчивости акустической речевой информации, позволяющая сформировать наборы данных для обучения многоканальной системы сверточных нейронных сетей при оценке защищенности акустической речевой информации от утечки по техническим каналам. В результате применения методики было подготовлено по 11520 зашумленных аудиозаписей, охватывающих 4 типа маскирующих помех и 4 типа голосов дикторов. На их основе обученные модели сверточных нейронных сетей показали высокую точность классификации: свыше 97% при работе с графиками MFCC в условиях спектров помех на основе белого шума, и более 96% при использовании спектрограмм в условиях помехи типа «речевой хор».

2. Разработана модель оценки разборчивости акустической речевой информации на основе многоканальной системы сверточных нейронных сетей, позволяющая учитывать высокий уровень маскирующих помех, различные спектры помех на основе белого шума, помехи типа «речевой хор» и различия в спектрах речи дикторов при определении уровня разборчивости речи на основе мажорантной оценки. Разработанная модель объединила возможности современных методов машинного обучения и обработки акустических сигналов. Она способна автоматически анализировать спектрограммы и графики MFCC, определяя уровни разборчивости речи с учетом воздействия маскирующих помех.

3. Разработана методика оценки защищенности акустической речевой информации, позволяющая повысить достоверность оценки защищенности и обеспечить отклонение значения защищенности от истинного значения (результаты субъективных испытаний) не более 4% в наиболее значимом диапазоне значений показателя оценки защищенности речевой информации, а возможность учета помехи типа «речевой хор» при оценке защищенности позволяет расширить область применения данной методики.

Результаты диссертационного исследования приняты к внедрению в производственные и бизнес-процессы компании ООО СпецНТЦ «Преграда» г. Самары, а также в учебные процессы ФГБОУ ВО «Самарский государственный технический университет» и ФГБОУ ВО «Новосибирский государственный технический университет».

Перспективы дальнейшего развития темы. В дальнейшем планируется совершенствование разработанной методики оценки защищенности акустической речевой информации для применения в условиях виброакустического канала утечки информации.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

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

1. Волков Н. А., Иванов А. В. Подход к оценке защищенности речевой акустической информации с применением нейронных сетей // Вестник СибГУТИ: научный журнал – Новосибирск: Изд-во СибГУТИ, 2024. Т. 18, № 2 – с. 43-56.

2. Волков Н. А., Иванов А. В. Определение параметров обучающей выборки для нейронной сети, применяемой при определении процента разборчивости речи в задачах оценки защищенности речевой акустической информации // Вопросы защиты информации: Науч.-практ. журн. / ФГУП «НТЦ оборонного комплекса «Компас», 2024. Вып. 4 (147). - с. 44-51.

3. Волков Н. А., Иванов А. В. К вопросу определения параметров обучающей выборки для обучения нейронной сети в задачах оценки защищенности речевой акустической информации с применением речевого хора // Инфокоммуникационные технологии: Науч.-практ. журн. / ПГУТИ, 2025. Том 23, № 1 (89) - с. 86-96.

4. Волков Н.А., Иванов А.В., Карпова Н.Е. Методика оценки защищенности речевой акустической информации от утечки по техническим каналам с применением сверточных нейронных сетей // Вопросы защиты информации: Науч.-практ. журн. / ФГУП «НТИЦ оборонного комплекса «Компас», 2025. Вып. 3 (150). – с. 22-32.

Статьи, индексируемые в международной базе Scopus:

5. Salimov S. R., Volkov N. A. and Ivanov A. V. The Use of Speech Recognition Systems to Select a Useful Signal in Noisy Speech at a Low Signal-To-Noise Ratio, 2021 Dynamics of Systems, Mechanisms and Machines (Dynamics), Omsk, Russian Federation, 2021, pp. 1-4.

6. Volkov, N., Ivanov, A. (2024). On the Use of Convolutional Neural Networks in the Tasks of Assessing the Security of Speech Acoustic Information. In: Lapina, M., Raza, Z., Tchernykh, A., Sajid, M., Zolotarev, V., Babenko, M. (eds) AISMA-2024: International Workshop on Advanced Information Security Management and Applications. AISMA 2024. Lecture Notes in Networks and Systems, vol 863. Springer, Cham., pp.320-327.

Прочие публикации:

7. Волков Н. А., Иванов А. В. О применении методов шумоочистки в задачах речевой акустической информации / Наука. Технологии. Инновации: материалы всероссийской научной конференции молодых ученых. Часть 1 – Новосибирск: Изд-во НГТУ, 2020. – 262 с – С. 12-14.

8. Волков Н. А., Иванов А. В. Применение методов шумоочистки для обработки речевой акустической информации / Сборник избранных статей научной сессии ТУСУР. – 2021. – № 1-3. – С. 34-37.

9. Волков Н. А. Об использовании современных систем распознавания речи для выделения полезного сигнала в зашумленной речи / Наука. Технологии. Инновации: XV Всероссийская научная конференция молодых ученых, посвященная Году науки и технологий в России: Сборник научных трудов в 10 частях, Новосибирск, 06–10 декабря 2021 года. Том Часть 1. – Новосибирск: Новосибирский государственный технический университет, 2021. – С. 33-37.

10. Волков Н. А., Иванов А. В., Карпова Н. Е., Чуваков А. В. Исследование зашумленной аудиозаписи речи для обработки в автоматизированной системе / Безопасность цифровых технологий. – 2022. – № 2 (105). – С. 9–20.

11. Волков Н.А., Иванов А.В. Использование сверточных нейронных сетей для оценки защищенности речевой акустической информации // Сборник материалов VII Всероссийской молодежной научно-практической конференции с международным участием «Информационные технологии обеспечения комплексной безопасности в цифровом обществе», Уфимский университет науки и технологий, г. Уфа, 2024. – С. 64-68.

Диссертант

Волков Н.А.

Волков Никита Андреевич

**МЕТОДИКА ОЦЕНКИ ЗАЩИЩЕННОСТИ АКУСТИЧЕСКОЙ РЕЧЕВОЙ
ИНФОРМАЦИИ ОТ УТЕЧКИ ПО ТЕХНИЧЕСКИМ КАНАЛАМ С
ПРИМЕНЕНИЕМ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ**

**Специальность 2.3.6. Методы и системы защиты информации,
информационная безопасность**

АВТОРЕФЕРАТ

**диссертации на соискание ученой степени
кандидата технических наук**

Формат 60x90 1/16. Объем: усл. печ. л. 1,0.

Тираж 100 экз. Заказ № 400.

Отдел типографии и оперативной полиграфии
ФГБОУ ВО «Самарский государственный технический университет»
443100, г. Самара, ул. Молодогвардейская, 244.