На правах рукописи

ВАСИЛЬЕВ Роман Александрович

БИОМЕТРИЧЕСКАЯ ИДЕНТИФИКАЦИЯ ПОЛЬЗОВАТЕЛЕЙ ИНФОРМАЦИОННЫХ СИСТЕМ НА ОСНОВЕ КЛАСТЕРНОЙ МОДЕЛИ ЭЛЕМЕНТАРНЫХ РЕЧЕВЫХ ЕДИНИЦ

Специальность: 05.13.19 – Методы и системы защиты информации, информационная безопасность

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Саров - 2016

Работа выполнена в Саровском физико-техническом институте - филиале ФГАОУ ВО «Национальный исследовательский ядерный университет «МИФИ» в лаборатории «Безопасность информационных и технических систем» при кафедре «Радиофизика и электроника».

Научный руководитель: кандидат технических наук, доцент
Николаев Дмитрий Борисович

Официальные оппоненты: доктор технических наук, профессор
Мартынов Александр Петрович
Российский федеральный ядерный центр, Всероссийский научно-исследовательский институт экспериментальной физики, начальник научно - исследовательского отдела

доктор технических наук, профессор
Сидоркина Ирина Геннадьевна
ФГБОУ ВО «Поволжский государственный технологический университет», декан факультета Информатики и вычислительной техники

Ведущая организация: ФГАОУ ВО «Нижегородский государственный университет им. Н.И. Лобачевского» (г. Нижний Новгород)

Защита диссертации состоится «27» января 2017 г. в 10^{00} часов на заседании диссертационного совета Д-212.288.07 при ФГБОУ ВО «Уфимский государственный авиационный технический университет» по адресу: 450008, г. Уфа, ул. К. Маркса, д. 12.

С диссертацией можно ознакомиться в библиотеке ФГБОУ ВО «Уфимский государственный авиационный технический университет» и на сайте *http://www.ugatu.su/*.

Автореферат разослан «___» _____ 20__ года.

Ученый секретарь
диссертационного совета
д.т.н., доцент И.Л. Виноградова

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. В последние годы для идентификации личности человека наиболее перспективным считается применение биометрических технологий, особенно при запросах конфиденциальной информации по телефону, в системах разграничения доступа, при управлении различными устройствами, в криминалистике и т.д.

Наиболее широкое применение в биометрической идентификации получили следующие параметры человека: особенности геометрии лица, отпечатки пальца, геометрия ладони рук, сетчатка и радужная оболочка глаза, голосовые характеристики, особенности подписи и клавиатурный подчерк. В некоторых случаях применение биометрических характеристик человека осложнено. Геометрии лица свойственна низкая уникальность, для анализа сетчатки и радужной оболочки глаза требуется дорогостоящее оборудование. Параметрам клавиатурного подчерка и подписи свойственна низкая стабильность и зависимость от эмоционального состояния человека. При применении сканеров отпечатков пальцев и геометрии ладони рук возможны вопросы чистоты контактных площадок и соблюдения санитарных норм. Однако широкое применение биометрических систем влечет за собой повышенный интерес со стороны злоумышленников, направленный на разработку атак по их взлому. Наиболее часто применяемой является атака, суть которой заключается в том, что в систему передаются биометрические признаки, предъявленные ранее, например, силиконовый муляж пальца или магнитофонная запись парольной фразы. Таким образом, разработку систем биометрической идентификации необходимо вести с учетом защиты от этих атак. Свести к минимуму недостатки указанных выше методов биометрической идентификации пользователей позволит разработка новых методов и алгоритмов идентификации, основанных на предъявлении случайно сформированных ключевых признаков из биометрической базы эталонов пользователей.

В связи с этим решаемая в диссертационной работе задача, заключающаяся в разработке алгоритмического и программного обеспечения системы идентификации, позволяющей предотвратить атаку на биометрическую систему и проводить текстонезависимую идентификацию по голосу на основе кластерной модели элементарных речевых единиц, является актуальной.

Степень разработанности темы. Исследованиями проблемы биометрической идентификации занимается ряд отечественных ученых: Аграновский А.В., Леднов Д.А., Балакирев Н.Е., Малков М.А., Галунов В.И., Соловьев А.Н., Кульбак С.,Винцюк, Т. К.,Савченко В.В., Маковкин К.А., Иванов А.И., и зарубежных специалистов: Дуглас А.Рейнолдс, Патрик Дж. Кенни, Маркел Дж.Д., Грэй А.Х., Анн К. Сурдал, Эрик Келлер, Фредерик Джелинек, Харри Френсис Холлен, Джон Р. Вакка, Джон Чирилло, Ловвер Б.Т. Большую работу в направлении исследования атак на голосовые биометрические системы провела группа исследователей под руководством

Томи Кинунен в Университете Восточной Финляндии. Существует множество компаний, успешно занимающихся разработкой программно-аппаратных комплексов идентификации по голосу, среди которых ООО «Центр речевых технологий» (разработана система «VoiceKey» и система «ИКАР Лаб»), ООО «ГритТек» (создана система «GritTec Speaker-ID»). Однако существующие разработки обладают рядом недостатков. В связи с чем актуальны исследования по созданию метода формирования эталонов голоса пользователя и повышение эффективности методов статистического анализа фонем, а также созданию более эффективного алгоритма идентификации пользователей по голосу.

Объектом исследования диссертационной работы является система биометрической идентификации пользователей по голосу.

Предметом исследования диссертационной работы являются методы и алгоритмы биометрической идентификации пользователей по голосу с применением кластерной модели элементарных речевых единиц.

Цель диссертационной работы – повысить эффективность идентификации пользователей информационных систем по голосу путем разработки методов и алгоритмов решения данной задачи на основе кластерной модели элементарных речевых единиц.

Для достижения указанной цели в диссертации были представлены и решены следующие **задачи**:

1. Исследование методов, алгоритмов, систем идентификации пользователей по голосу и анализ уязвимости современных голосовых биометрических систем к различным способам фальсификации индивидуальных голосовых характеристик;
2. Разработка метода формирования голосовых эталонов пользователей на основе кластерной модели элементарных речевых единиц;
3. Разработка метода статистического анализа фонем и принципа накопления информации для решения задачи идентификации по голосу;
4. Разработка алгоритмов идентификации пользователей по индивидуальным характеристикам голоса в условиях вариативности речи с учетом возможности защиты от различных видов атак на систему биометрической идентификации;
5. Разработка программного комплекса для биометрической идентификации пользователей информационных систем по голосу, экспериментальные исследования разработанного комплекса идентификации, представление рекомендаций по его практическому применению в реальных условиях эксплуатации.

Методы исследования. В диссертационной работе используются методы теории информации и теории вероятностей, теории распознавания образов, спектрального анализа, теории речеобразования, теория сигналов, методы проектирования программного и информационного обеспечения, технологии объектно-ориентированного программирования.

Научная задача: разработать методы и алгоритмы текстонезависимой идентификации пользователей информационных систем по голосу на основе кластерной модели элементарных речевых единиц в условиях малой обучающей выборки с учетом возможности защиты от различных видов атак на систему биометрической идентификации.

Научная новизна результатов исследования:

1. Предложен метод формирования голосовых эталонов пользователя, основанный на кластерной модели элементарных речевых единиц в информационной метрике Кульбака-Лейблера, отличающийся от известных методов определением информационного центра эталона голоса пользователя с последующей кластеризацией голосовых эталонов, что позволяет уменьшить количество ошибок при идентификации пользователей информационных систем по голосу в среднем в 1,5 раза;
2. Предложен метод статистического анализа фонем и принцип накопления информации, основанные на цифровом программном обнаружителе и критерии Неймана-Пирсона, отличающиеся от других методов применением статистического анализа элементарных речевых единиц для принятия решения по идентификации, что обеспечивает уменьшение количества ошибок идентификации более чем в 4,5 раза;
3. Предложены алгоритмы идентификации пользователей информационных систем по индивидуальным характеристикам голоса, основанные на совместном использовании метода статистического анализа фонем и кластерной модели элементарных речевых единиц в метрике Кульбака-Лейблера, отличающиеся повышенной защищенностью от различных видов атак на систему биометрической идентификации, позволившие идентифицировать пользователей с вероятностью ошибок первого и второго рода 0,025 и 0,005.

Практическая значимость научной работы. Полученные результаты позволили решить проблему надежности идентификации по голосу. Применение полученных результатов позволит повысить надежность процесса идентификации в информационных системах от различных атак. Практическую ценность представляют:

- разработанный программный комплекс для биометрической идентификации пользователей информационных систем по голосу и реализованный в комплексе метод формирования голосовых эталонов пользователя на основе кластерной модели элементарных речевых единиц, благодаря которому было уменьшено количество ошибок идентификации пользователей информационных систем в среднем в 1,5 раза;
- результаты экспериментальных испытаний по идентификации пользователей, полученные в «информационной системе идентификации пользователей по голосу», базирующейся на применении описанных выше методов и алгоритмов, отличающийся от существующих систем возможностью текстонезависимой идентификации с защитой от атак, что позволяет подтвердить высокую надежность процедуры идентификации при

влиянии на пользователя внешних факторов с вариативностью речи и предотвратить попытки атак на систему идентификации.

Полученные результаты применимы как в системах защиты информации от несанкционированного доступа, использующие параметры голоса для идентификации пользователей, так и в системах разграничения доступа в помещения с голосовой идентификацией. Разработанный алгоритм идентификации так же можно применять в системах криминалистической (фоноскопической) экспертизы, использующих в качестве доказательной базы голос подозреваемого.

Внедрение результатов работы. Результаты диссертационного исследования использованы в практической деятельности Нижегородского НТЦ ФГУП «НПП «Гамма». Выполнено внедрение эскизного проекта разработанного программного комплекса идентификации пользователей по голосу для усиления механизмов идентификации системы разграничения доступа к информации.

В Федеральной службе по интеллектуальной собственности (Роспатент) получено свидетельство о государственной регистрации программы для ЭВМ №2015663306 от 15.12.2015г. «Программа идентификации дикторов по голосу».

Решением Ученого совета Саровского физико-технического института-филиала ФГАОУ ВО «Национальный исследовательский ядерный университет «МИФИ» результаты диссертационной работы внедрены в учебный процесс лаборатории «Безопасность информационных и технических систем». В рамках учебного процесса был разработан новый учебный курс, поставлена серия лабораторных работ по данному курсу, издано учебное пособие по курсу с применением системы идентификации пользователя по голосу.

Соответствие диссертации паспорту научной специальности.

Диссертация соответствует п. 11 паспорта специальности 05.13.19 – «Методы и системы защиты информации, информационная безопасность» - п. 11. Технологии идентификации и аутентификации пользователей и субъектов информационных процессов. Системы разграничения доступа.

Достоверность результатов подтверждена результатами экспериментов, проведенных в разработанной «информационной системе идентификации дикторов по голосу», а также использованием признанной методики статистической обработки данных.

Апробация работы. Основные положения работы доложены и обсуждены на V Международной научно-практической конференции «Информационные технологии в науке, бизнесе и образовании» (г. Москва, 2012г.), Международной научно-технической конференции «Информационные системы и технологии» (г. Нижний Новгород, 2013г.), II Международной научно-практической конференции «Технические науки – основа современной инновационной системы» (г. Йошкар-Ола, 2013г.), XI

Международной научно-технической конференции «Новые информационные технологии и системы» (г. Пенза, 2014г).

Публикации. Результаты исследований опубликованы в 6 журналах, рекомендованных ВАК, материалы диссертационной работы докладывались и обсуждались на 4 международных научно-практических конференциях и 8 всероссийских научно-технических конференциях. Получено свидетельство о государственной регистрации программы для ЭВМ.

Личный вклад автора. Основные результаты и положения, рассмотренные в диссертационной работе, получены автором лично. Методы и алгоритмы идентификации разработаны и экспериментально исследованы лично автором. Научный руководитель участвовал в постановке цели и задач исследований.

Основные положения, выносимые на защиту:

1. Метод формирования голосовых эталонов пользователя, включающий построение информационного центра эталона голоса пользователя с последующей кластеризацией голосовых эталонов, основанных на кластерной модели элементарных речевых единиц в информационной метрике Кульбака-Лейблера, позволивший уменьшить количество ошибок при идентификации пользователей информационных систем по голосу в среднем в 1,5 раза;
2. Метод статистического анализа фонем и принцип накопления информации, на основе цифрового программного обнаружителя и критерия Неймана-Пирсона, при помощи которого снижено количество ошибок идентификации пользователей по голосу не менее чем в 4,5 раза по сравнению с существующими методами;
3. Алгоритмы идентификации пользователей информационных систем по индивидуальным характеристикам голоса, основанные на совместном использовании метода статистического анализа фонем и кластерной модели элементарных речевых единиц в метрике Кульбака-Лейблера, позволяющие идентифицировать пользователей с вероятностью ошибок первого рода 0,025 и второго рода 0,005, с учетом возможного влияния различных видов атак на систему биометрической идентификации;
4. Программный комплекс для идентификации пользователей информационных систем по голосу, в основе которого лежат предложенные методы и алгоритм идентификации, позволяющий повысить защищенность процесса идентификации от внешних атак в системах разграничения доступа.

Объем и структура работы. Диссертационная работа включает введение, четыре главы, заключение, список используемой литературы и приложения. Вся работа изложена на 153 страницах текста, включающих в себя 3 страницы приложений, 69 рисунков, 16 таблиц. Количество библиографических ссылок – 101.

СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во введении дается обоснование актуальности темы применения систем голосовой биометрии в информационных системах с разграничением доступа. Сформулирована цель работы и приведено описание основных полученных результатов. Дана характеристика научной новизны и практической значимости диссертационной работы, а также представлена аннотация диссертационной работы по главам.

В первой главе рассматриваются и анализируются наиболее известные из литературы подходы к задаче идентификации пользователя по голосу, а именно: методы динамического программирования, векторное квантование, смеси Гауссовских процессов, скрытая Марковская модель, теоретико-информационный подход. Анализируются известные разработки ведущих компаний в данной области («Центр речевых технологий», «ГритТек»).

Описаны различные виды фальсификации индивидуальных голосовых характеристик и атак на систему голосовой идентификации: методы, основанные на приеме имперсонализации; методы, основанные на записи голосовых биометрических характеристик человека и их дальнейшем повторе; методы, основанные на технологии преобразования речи злоумышленника в речь другого человека; методы, основанные на технологии синтеза речи. Приведены известные решения по увеличению защищенности голосовых биометрических систем к описанным видам атак.

По результатам проведенного анализа выделен ряд нерешенных на данный момент проблем в области идентификации пользователей по голосу. Среди них главной является проблема малых выборок наблюдений и текстозависимая идентификация, как следствие – невысокая точность и надежность при проведении идентификации по голосу. Указанная проблема в значительной мере может быть ослаблена с использованием адаптивной кластерной модели элементарных речевых единиц (ЭРЕ), применяющий принципы информационной теории восприятия речи (ИТВР) профессора Савченко В. В. Показано, что применение кластерной модели элементарных речевых единиц позволяет существенно повысить точность и надежность идентификации при малых вычислительных затратах по сравнению с разработанными ранее методами. Это дает возможность применения кластерной модели в системах с текстонезависимой идентификацией для повышения защищенности от различных атак на голосовые системы.

Во второй главе решается задача повышения точности идентификации пользователей по голосу (ИДГ), разрабатывается метод формирования голосовых эталонов пользователя, основанный на кластерной модели элементарных речевых единиц в информационной метрике Кульбака-Лейблера, так же разрабатывается метод статистического анализа фонем и принцип накопления информации, на основе цифрового программного обнаружителя и критерия Неймана-Пирсона.

Для разработки метода формирования голосовых эталонов пользователя выбрано определение фонемы, с точки зрения ИТВР – это

множество ЭРЕ, объединенных в группу по критерию минимума информационного рассогласования (МИР). На основе данного определения был разработан метод кластеризации ЭРЕ и формирования словаря эталонов пользователя на основе ЭРЕ типа отдельных фонем из слитной речи.

Первый этап кластеризации – это деление имеющегося фрагмента речевого сигнала на короткие (т порядка 10 мс) сегменты (отрезки) приблизительной его стационарности с одновременным вычислением по каждому сегменту данных значений авторегрессионных (АР)-параметров ЭРЕ. Для этого применялась известная рекуррентная вычислительная процедура Берга-Левинсона. Для представления принципа МИР в частотной области (энергетический спектр) использовалось Гауссово распределение сигнала $\mathbf{P}(X_r)$ с автокорреляционной матрицей (АКМ) $\mathbf{K}_r$ ленточной структуры, где выражение для решающей статистики сводится к виду:

$$\rho_{x,r} \overset{\Delta}{=} \frac{1}{F}\sum_{f=1}^{F}\left(\frac{G_x(f)}{G_r(f)} + \ln\frac{G_r(f)}{G_x(f)}\right) - 1 \to \min\Big|_{r=\overline{1,R}}. \qquad (1)$$

Здесь $G_x(f)$ – выборочная оценка спектральной плотности мощности (СПМ) сигнала X в функции дискретной частоты f; $G_r(f)$ – СПМ r-го сигнала из словаря эталонов; F – верхняя граница частотного диапазона сигнала или используемого канала связи. В выражении (2) представлена известная формулировка критерия МИР на основе АР-модели речевого сигнала:

$$x(n) = \sum_{i=1}^{P} a(i)x(n-i) + \varepsilon(n). \qquad (2)$$

Здесь $x(n)$ – значение n-го отсчета речевого сигнала, $\mathbf{a} = \{a(i)\}$ – вектор его АР–коэффициентов, P – порядок АР-модели, а $\varepsilon(n)$ – порождающий процесс типа белого гауссова шума (БГШ) с нулевым значением математического ожидания и фиксированной дисперсией σ^2.

При дополнительном условии нормировки АР-модели сигналов типа ЭРЕ по дисперсиям их порождающего шума выражение для решающей статистики МИР (1) приобретает предельно простой вид, позволяющий вычислить вероятность информационного рассогласования (ВИР) через энергетический спектр:

$$\rho_{x,r} = \frac{1}{F}\sum_{f=1}^{F}\frac{\left|1 + \sum_{m=1}^{P} a_r(m)\exp(-j\pi mf/F)\right|^2}{\left|1 + \sum_{m=1}^{P} a_x(m)\exp(-j\pi mf/F)\right|^2} - 1, \qquad (3)$$

Здесь $\rho_{x,r}$ - величина ВИР ЭРЕ, $a_r(m)\exp(-j\pi mf/F)$ - векторы АР коэффициентов. Это стандартная формулировка метода обеляющего фильтра (МОФ) в задаче ИДГ на основе выборочной оценки ВИР между сигналом X на входе и r-м сигналом из словаря в частотной области. Преимуществом данной интерпретации критерия МИР является, прежде всего, возможность его эффективной реализации в адаптивном варианте на основе быстрых

вычислительных процедур авторегрессионного анализа, таких как метод Берга. На следующем этапе вычислений предложен метод редукции данных, основанный на объединении однородных смежных сегментов в одну ЭРЕ без потери полезной информации. Решение об объединении принималось по критерию МИР: $\rho(\mathbf{x}_l, \mathbf{x}_{l+1}) \le \rho_1, \quad \mathbf{x}_l \cap \mathbf{x}_{l+1} = \mathbf{X}_L,$ (4)

Здесь ρ_1 - некоторый пороговый уровень, характеризующий допустимую степень неоднородности одноименных сегментов речевого сигнала, $\mathbf{X}_L$ - реализация ЭРЕ, $\mathbf{x}(l),\ l = 1,2,...$ – вектор отсчетов речевого сигнала (фрейма) на интервале его стационарности τ. Пример сегментации слова «раз» приведен на рисунке 1.

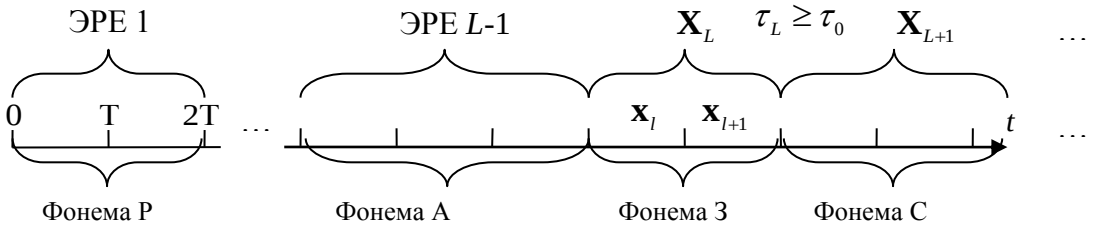

Рисунок 1 - Сегментация слова «раз».

Кроме того, для отбора надежных фонем дополнительно применялась процедура проверки ЭРЕ в отношении ее длительности $\tau_L \ge \tau_0 = n\mathrm{T}$, где τ_0 - порог по длительности, кратный периоду сегментации T.

На втором этапе все множество выделенных ЭРЕ было разбито на R подмножеств – путем последовательной группировки подобных речевых единиц в одну фонему-кластер: $\rho(\mathbf{X}_L, \mathbf{X}_{L+1}) \le \rho_2$. (5)

Здесь ρ_2 - второй пороговый уровень по ВИР, характеризующий допустимую степень неоднородности в метрике Кульбака-Лейблера одноименных ЭРЕ.

После этого на основе критерия МИР в пределах каждого r-го кластера вычисляется ИР в кластере: $\rho_k^{(r)} = \sum_{j=1}^{J_r} \rho^{(r)}(\mathbf{X}_k, \mathbf{X}_j), \quad k = \overline{1, J_r},$ (6)

Здесь $\rho^{(r)}(\mathbf{X}_к, \mathbf{X}_j)$ – величина ИР по Кульбаку-Лейблеру.

На третьем этапе определялась минимальная ВИР информационного центра-эталона (ИЦ-эталона) ЭРЕ: $\mathbf{X}_r^* = \arg\min_k \rho_k^{(r)}, r = \overline{1, R}$ (7)

Множество (по числу фонем R) отобранных таким образом ИЦ-эталонов и составляет, в конечном итоге, искомый результат автоматической обработки речевого сигнала: словарь эталонов голоса пользователя, или фонетическую базу данных, состав которой зависит от фонетических особенностей конкретного пользователя.

Так же разработан метод на основе цифрового программного обнаружителя (ЦПО) и критерия Неймана-Пирсона, позволяющий проводить статистический анализ и накопление информации о фонемах пользователя для решения задачи ИДГ.

Для описания метода обозначим $p_ш$ - вероятность события u_i =1, наблюдающегося при превышении шумом порога квантования U_0 на i-ой

позиции (фонема пользователя проходит по своим параметрам пороговое значение), а $q_ш$ - вероятность противоположного события u_i=0 ($q_ш$=1- $p_ш$). Аналогичные вероятности при приеме смеси сигнала с шумом обозначим $p_{сш}$ и $q_{сш}$. Относительно характера выборки U_1, U_2, U_3,…U_N можно выдвинуть две гипотезы: H_0 - выборка порождена шумом (фонемами «чужих» пользователей) и H_1 - выборка порождена смесью сигнала с шумом (фонемы «определяемого» пользователя).

Оптимальной процедурой вынесения решения является вычисление отношения правдоподобия $\Lambda(U_1, U_2, U_3,\ldots U_N)$ (подсчет количества распознанных ЭРЕ из БД) и сравнение его с порогом Λ_0, величина которого при использовании критерия Неймана-Пирсона определяется заданной вероятностью ложной тревоги (отсутствие ЭРЕ в БД). В рассматриваемом случае $\Lambda(U_1, U_2, U_3,\ldots U_N)= P_{сш} / P_ш$, где $P_{сш}$ и $P_ш$ - соответственно вероятности того, что выборка порождена смесью сигнала с шумом (определение и подсчет ЭРЕ в БД) или только шумом (отсутствие ЭРЕ в БД). Вследствие независимости событий u_i (i=1…N) вероятность того, что данная выборка шума содержит k единиц, которые занимают определенные позиции, составляет (согласно биномиальному закону распределения) $P_ш = p_ш^k q_ш^{N-k}$. Аналогично при приеме смеси прямоугольной последовательности импульсных сигналов с шумом $P_{сш} = p_{сш}^k q_{сш}^{N-k}$. Поэтому отношение правдоподобия определяется:

$$\Lambda(U_1,U_2,U_3...U_N) = \frac{p_{сш}^k (1-p_{сш})^{N-k}}{p_ш^k (1-p_ш)^{N-k}} \qquad (8)$$

Гипотеза H_1 принимается, если $\Lambda(U_1, U_2, U_3,\ldots U_N)= \Lambda_0$. Решая это выражение относительно k, получим:

$$k > \frac{\ln\left[\Lambda_0\left(\frac{1-p_ш}{1-p_{сш}}\right)^N\right]}{\ln\frac{p_{сш}(1-p_ш)}{p_ш(1-p_{сш})}} = k_0 \qquad (9)$$

Следовательно, оптимальное правило обнаружения сигнала по выражению (9) (подсчет количества совпадений ЭРЕ конкретного пользователя) по выборке его N дискретных значений заключается в сравнении числа единиц в выборке с пороговым числом k_0. Если $k>k_0$ (количество совпадающих ЭРЕ больше порога 60%) то принимается решение о приеме сигнала (идентификация пользователя), в противном случае об его отсутствии (пользователь не идентифицирован). В качестве примера, положим N=56 и $p_ш = 0{,}1$. В результате согласно (8) получим оптимальный пороговый уровень порядка $k_0 \approx 38$. Т.е. примерно 2/3 совпадений фонем пользователя – это оптимальный порог для разработанного выше метода.

В третьей главе для практической реализации разработанного метода формирования голосовых эталонов пользователя, основанного на кластерной модели элементарных речевых единиц в информационной метрике

Кульбака-Лейблера, в соответствии с выражениями (3…7) предложен алгоритм выделения ИЦ - эталона голоса пользователя.

Для реализации метода статистического анализа фонем и принципа накопления информации, основанных на цифровом программном обнаружителе и критерии Неймана-Пирсона, разработан алгоритм статистического подсчета и анализа всех выделенных фреймов, где в соответствии с выражением (9) по наиболее часто повторяющимся фреймам (не менее 60%) принимается решение о принадлежности данного голоса конкретному пользователю.

Для разработки программного обеспечения «информационной системы идентификации пользователей по голосу» («ИС ИДГ») сформирована блок-схема модулей и подсистем, из которых будет состоять «ИС ИДГ».

На рисунке 2 изображена архитектура функционирования подсистемы идентификации по голосу, реализующая в себе предложенные выше методы и алгоритмы.

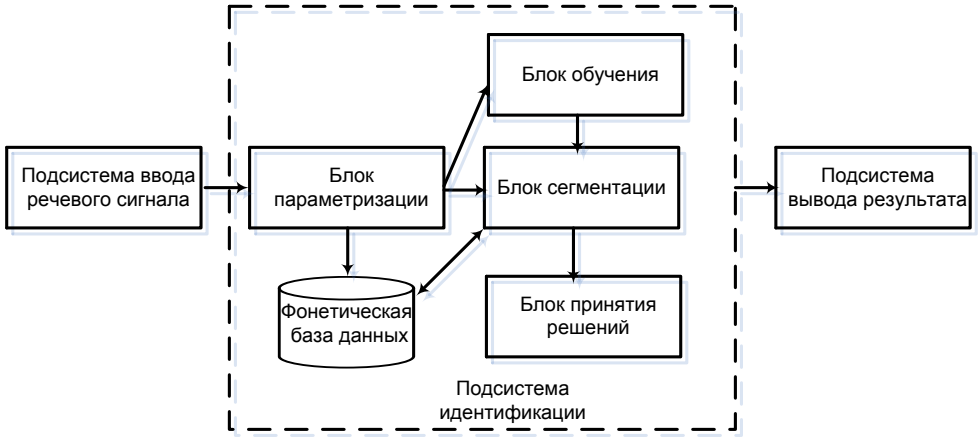

Рисунок 2 - Архитектура подсистемы голосовой идентификации.

Для сведения описанных выше блоков и подсистем в единую программную реализацию предложена схема функционирования программной оболочки «ИС ИДГ» (рисунок 3).

На данной схеме в модулях идентификации и распознавания выполняется процедура анализа и подсчета повторяющихся фреймов пользователя, выполняется проверка содержания (распознавание) фреймов в соответствии с произнесенным словом. При наибольшем соответствии фреймов определенному пользователю (не менее 60%) и совпадении распознанных слов со словами в окне системы, принимается решение о предоставлении доступа пользователю.

Так же определены и классифицированы типовые сценарии атак на различные компоненты обобщенной голосовой биометрической системы: атака на устройство ввода биометрической информации, атака на канал связи, атака на компонент обработки данных, атака на базу данных речевых шаблонов, атака на компонент принятия решения, атака на интерфейс вывода результата. Отмечено, что только атака на устройство ввода биометрической

информации не является общей для всех биометрических систем и требует уникальных методов противодействия.

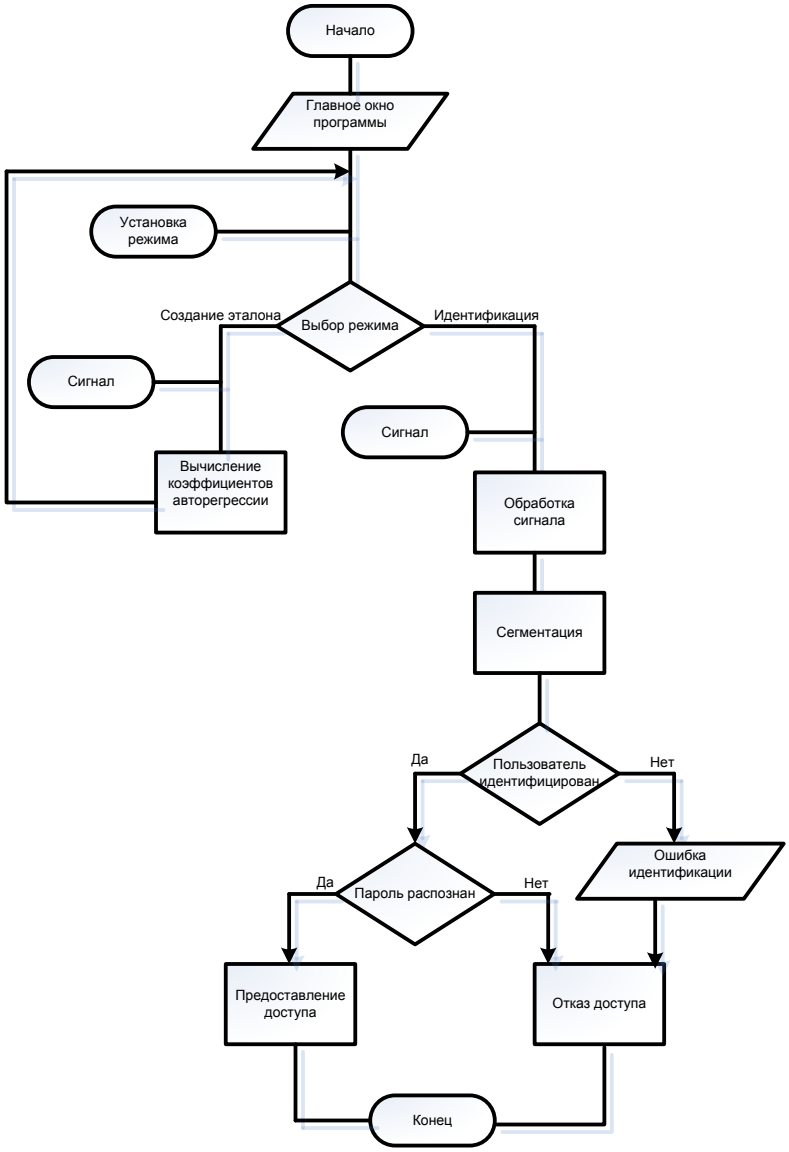

Рисунок 3 - Блок-схема функционирования программы «ИС ИДГ»

В стандартных системах идентификации существует следующий недостаток: нарушитель может записать на диктофон эталонное голосовое сообщение, а далее получить доступ по записи голоса. Для исключения данного недостатка в структуру «ИС ИДГ» встроена схема модуля идентификации пользователей с защитой от атак (рисунок 4).

Рисунок 4 - Интерфейс модуля идентификации с защитой от атак.

На 1-м этапе записываются не связанные между собой короткие голосовые сообщения в виде отдельных слов (существительное, глагол, наречие, например «идентификация происходит надежно»). Под цифрой 1 - обозначен микрофон, 2 - канал связи, 3 - аналого-цифровой преобразователь, 4 - запоминающее устройство. На 2-м этапе информационная система выбирает несколько записанных пользователем звуковых сообщений и предоставляет возможность в выбранном порядке произнести данные слова с указанием текста в окне системы. Пользователь произносит данные слова для последующего анализа в устройстве анализа данных – 5. Далее устройство идентификации – 6, выполняет процедуру анализа и подсчета повторяющихся фреймов пользователя, выполняет проверку содержания (распознавание) фреймов в соответствии с произнесенным словом, принимается решение об успешной идентификации.

В четвертой главе решаются вопросы реализации методов и алгоритмов идентификации в разработанном программном комплексе «информационная система идентификации пользователей по голосу» («ИС ИДГ»). Интерфейс «ИС ИДГ» состоит из главной формы, в которой отображаются пользователи, внесенные в базу данных. В данном меню возможен выбор режим работы, сохранение и последующее отображения данных. В меню «идентификация» осуществляется последовательное сегментирование звукового сигнала, выделение фреймов пользователя и его последующая идентификация с возможностью распознавания выделенных фреймов для повышения надежности текстонезависимой идентификации с защитой от атак. Общий вид интерфейса показан на рисунке 5.

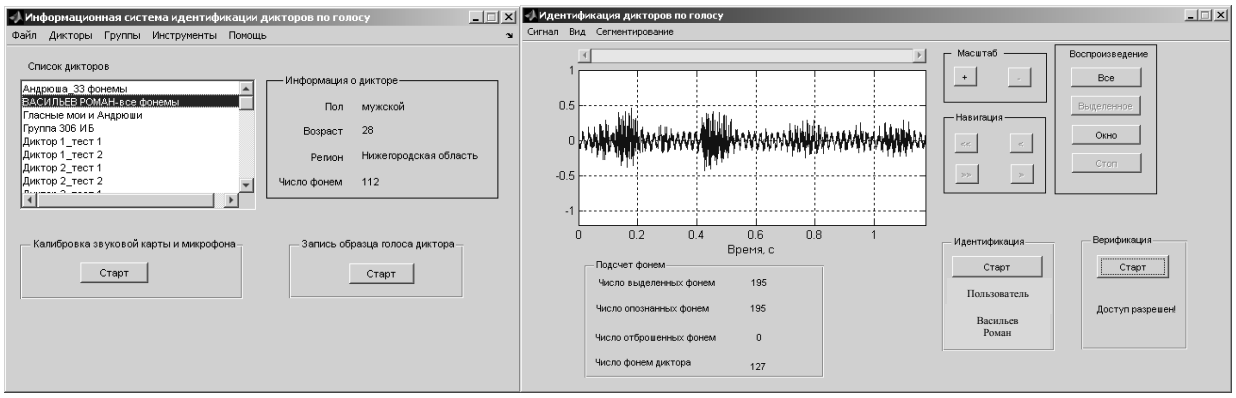

Рисунок 5 - Интерфейс «ИС ИДГ».

Тестирование «ИС ИДГ» проводилось в соответствии с правилами тестирования систем идентификации, установленными в стандарте ИСО/МЭК 19795-1-2007 «Автоматическая идентификация. Идентификация биометрическая. Эксплуатационные испытания и протоколы испытаний в биометрии. Часть 1. Принципы и структура» и ГОСТ 16600-72 «Передача речи по трактам связи. Требования к разборчивости речи и методы артикуляционных измерений». В соответствии с данным стандартом необходимо проводить испытания в зависимости от возрастных, гендерных, физиологических, эмоциональных состояний испытуемой группы. Программа экспериментальных испытаний приведена в таблице 1.

Таблица 1 - Программа экспериментальных испытаний информационной системы идентификации пользователей по голосу.

№	Название эксперимента
№1	Выявление различия фонем пользователей для процедуры идентификации
№2	Выявление различия в произношении пользователей для проведения процедуры идентификации по отдельным фразам
№3	Проведение текстонезависимой идентификации пользователей по голосу
№4	Проверка возможности идентификации пользователей различных национальностей
№5	Определение влияния физического (в том числе состояния здоровья) и эмоционального состояния пользователей на процесс идентификации
№6	Исследование вероятности правильной идентификации при использовании технологий клонирования и пародирования речи (voice changing) для модификации «подделки» голоса пользователя.

Таким образом, стандарт ИСО/МЭК 19795-1 был адаптирован под испытания систем идентификации человека по речевому сигналу. Так же по результатам эксперимента сделан вывод, что при увеличении числа идентифицируемых пользователей качество идентификации не снижается, в связи с чем в разработанной информационной системе нет явного ограничения на количество пользователей. Представленные выше эксперименты проведены в разработанной «ИС ИДГ» и в аналогичных программно-аппаратных и программных комплексах идентификации по голосу российских производителей, среди которых система «VoiceKey» - ООО «Центр речевых технологий», система «ИКАР Лаб» - ООО «Центр речевых технологий», GritTec Speaker-ID – ООО“ГритТек”, таблица 2.

Проведя анализ полученных результатов, можно сделать вывод, что разработанная в рамках диссертационных исследований «ИС ИДГ» имеет ряд преимуществ над системами аналогами.

Таблица 2 – Сравнение показателей программно-аппаратных комплексов идентификации по голосу российских производителей

Система идентификации / **Параметры проверки системы**	**«ИКАР Лаб»**	**«ИС ИДГ»**	**«Voice Key»**	**«GTS-ID»**
1	2	3	4	5
EER (Equal Error Rate) - вероятность ошибок биометрической системы доступа, при котором FAR и FRR равны	1-3%	1%	2–3%	4%
Отказ в регистрации	1%	2%	4%	~0%

Продолжение таблицы 2

1	2	3	4	5
FRR (False Rejection Rate) – вероятность ложного отклонения. Вероятность отклонения «своего» пользователя, приняв его за «чужого» (ошибка первого рода)	0,001	0,025	0,01	0,04
FAR (False Acceptance Rate) – вероятность ложного принятия. Вероятность принятия «чужого» пользователя за «своего» (ошибка второго рода)	0,012	0,005	0,025	0,004
Стоимость системы	Очень высо-кая	Низкая	Высо-кая	Высо-кая

В **заключении** сформулированы основные результаты работы и подведены итоги.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ

1. Проведен анализ существующих методов, алгоритмов, систем идентификации пользователей по голосу, который выявил их уязвимость к различным способам фальсификации индивидуальных голосовых характеристик и возможность проведения атак.

2. Разработан метод формирования голосовых эталонов пользователя, включающий построение информационного центра эталона голоса пользователя с последующей кластеризацией голосовых эталонов, позволивший уменьшить количество ошибок при идентификации пользователей информационных систем по голосу в среднем в 1,5 раза.

3. Разработан метод статистического анализа фонем и принцип накопления информации, основанный на цифровом программном обнаружителе и критерии Неймана-Пирсона, который позволяет снизить количество ошибок идентификации пользователей по голосу не менее чем в 4,5 раза по сравнению с существующими методами. Вероятности ошибок 1-ого и 2-ого рода составили 0,025 и 0,005 при наличии в базе 100 эталонов, что позволяет эффективно противодействовать различным атакам на процесс идентификации.

4. Разработаны алгоритмы идентификации пользователей информационных систем по индивидуальным характеристикам голоса в пространстве малоинформативных признаков, основанный на совместном использовании модернизированного метода статистического анализа фонем и кластерной модели элементарных речевых единиц в метрике Кульбака-Лейблера, позволяющий снизить вероятность ошибки 2-ого рода до 0,005, при вероятности ошибки 1-ого рода на прежнем уровне (0,025).

5. Разработан программный комплекс для идентификации пользователей информационных систем по голосу, в основе которого лежат

предложенные методы и алгоритм идентификации. Результаты работы позволяют повысить надежность процесса идентификации от неавторизованного доступа и предотвратить атаки на систему биометрической идентификации.

Перспективы дальнейшей разработки темы. Дальнейшие исследования связаны с расширением экспериментальной базы в области идентификации по голосу, планируются более глубокие исследования по части применения разработанного метода и алгоритма при криминалистической (фоноскопической) экспертизе.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

В рецензируемых журналах из списка ВАК

1. Васильев Р.А. Исследование особенностей фонетического строя речи и текстонезависимая идентификация дикторов по непрерывной речи // Информационная безопасность регионов. 2012. № 2 (11). С. 57-63.

2. Васильев Р.А. Исследование фонетического строя речи и идентификация дикторов по голосу // Вопросы защиты информации. 2013. № 1 (100). С. 43-51.

3. Васильев Р.А. Исследование особенностей фонетического строя речи и определение национальности дикторов при проведении идентификации по голосу // Информация и безопасность.2012.Т.15.№ 4. С. 487-494.

4. Васильев Р.А. Исследование особенностей идентификации пользователей по голосу при использовании технологий клонирования и пародирования речи для модификации голоса дикторов // Известия Тульского государственного университета. Технические науки. 2013. № 3. С. 246-252.

5. Васильев Р.А. Исследование особенностей идентификации дикторов по голосу при различиях в произношении // Безопасность информационных технологий. 2013. № 1. С. 85-86.

6. Савченко В.В., Васильев Р.А. Анализ эмоционального состояния дикторов по голосу на основе фонетического детектора лжи// Научные ведомости Белгородского государственного университета.2014.Вып. № 21(192)32\1. С186-195.

В трудах международных и всероссийских конференций

7. Николаев Д.Б., Васильев Р.А. Анализ возможности применения голосовой идентификации в системах разграничения доступа к информации // Научный результат // Серия: Информационные технологии. Белгородский государственный университет. 2016. Вып. 1. С 30-38.

8. Васильев Р.А. Выявление различия фонем пользователей для процедуры идентификации дикторов по голосу // Материалы V Международной научно-практической конференции Информационные технологии в науке, бизнесе и образовании // Современная наука: актуальные проблемы теории и практики. Серия: Естественные и технические науки. Финансовый Университет при Правительстве РФ, Москва, 2012. № 8-9. С. 19-22.

9. Васильев Р.А. Выявление различия в произношении дикторов для проведения процедуры идентификации по отдельным фразам // Материалы Международной научно-технической конференции Информационные системы и технологии. НГТУ им. Р.Е.Алексеева, Н.Новгород, 2013. С 10-14.

10. Васильев Р.А. Проведение текстонезависимой идентификации дикторов по непрерывной речи // Материалы II Международной научно-практической конференции Технические науки – основа современной инновационной системы. Центр «Коллоквиум», Йошкар-Ола, 2013. С 91-95.

11. Савченко. В. В., Васильев Р. А. Автоматическая оценка качества речи по критерию минимума требуемой избыточности речевого сигнала // Материалы XI Международной научно-технической конференции посвященной памяти Б.И. Рамеева. Новые информационные технологии и системы. Пензенский государственный университет. 2014. С 15-19.

12. Васильев Р.А. Определение национальности дикторов при проведении процедуры идентификации // Материалы Всероссийской конференции студентов и аспирантов. Информационные и инфокоммуникационные технологии – реалии, возможности, перспективы. НГИЭИ. Княгинино 2013. С 22-26.

13. Васильев Р.А. Определение влияния физического и эмоционального состояния дикторов на процесс идентификации // Материалы Всероссийской научно-технической конференции. Исследования в области полиграфии и защиты информации. ТулГУ совместно с ФГУП НИИ Репрографии, Тула. 2013. С 19-24

14. Васильев Р.А. Исследование особенностей идентификации дикторов по голосу // Материалы Семнадцатой научной конференции по радиофизике, посвященной 100-летию со дня рождения В.С. Троицкого. ННГУ им. Н.И. Лобачевского, Н.Новгород, 2013. С 45-48.

15. Васильев Р.А. Разработка новой технологии автоматической идентификации дикторов по голосу // Всероссийский конкурс молодежных инновационных команд. Россия – Ответственность – Стратегия – Технологии. ГБОУ ДПО Нижегородский научно-информационный центр. 2013. С 19-24.

16. Васильев Р.А. Исследование вероятности правильной идентификации при использовании технологий клонирования речи для модификации голоса дикторов. // Материалы 19-й Нижегородской сессии молодых ученых – Технические науки, Н.Новгород, 2014. С 23-28.

17. Савченко. В. В., Васильев Р. А. Результаты экспериментального исследования фонетического детектора лжи // Материалы IV Всероссийской научно-технической конференции Информационно-измерительные и управляющие системы военной техники. Владимирский государственный университет, Владимир. 2014. С 22-26.

18. Николаев Д.Б., Васильев Р.А. Исследование вопросов идентификации на основе метода выделения речевого базиса// Сборник материалов X Всероссийской научно инновационной-школы. СарФТИ НИЯУ МИФИ. Саров. 2016. С 7-9.

Объекты интеллектуальной собственности

19. Свид. о гос. регистрации программы для ЭВМ №2015663306 Программа идентификации дикторов по голосу / Васильев Р.А. Зарег. 15.12.2015г. – М.: Роспатент, 2015.

Диссертант  Васильев Р.А.

ВАСИЛЬЕВ Роман Александрович

БИОМЕТРИЧЕСКАЯ ИДЕНТИФИКАЦИЯ ПОЛЬЗОВАТЕЛЕЙ ИНФОРМАЦИОННЫХ СИСТЕМ НА ОСНОВЕ КЛАСТЕРНОЙ МОДЕЛИ ЭЛЕМЕНТАРНЫХ РЕЧЕВЫХ ЕДИНИЦ

Специальность: 05.13.19 – Методы и системы защиты информации, информационная безопасность

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Подписано в печать 21.11.2016. Формат 60×84 1/16
Бумага офсетная. Печать плоская. Гарнитура Times New Roman.
Усл. печ. л. 1,0. Уч.-изд. л. 0,9.
Тираж 100 экз. Заказ № 13.

Саровский физико-технический институт - филиал ФГАОУ ВО
«Национальный исследовательский ядерный университет «МИФИ»
Центр оперативной полиграфии
607186, Нижегородская обл., г. Саров, ул. Духова, д.6