МИНИСТЕРСТВО ЦИФРОВОГО РАЗВИТИЯ,
СВЯЗИ И МАССОВЫХ КОММУНИКАЦИЙ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Санкт-Петербургский государственный университет телекоммуникаций
им. проф. М.А. Бонч-Бруевича»

На правах рукописи

Альотум Юсеф Мохаммед Абд Аллх

РАЗРАБОТКА МЕТОДИКИ И АЛГОРИТМОВ ЗАЩИТЫ АУТЕНТИФИКАЦИОННЫХ ДАННЫХ ПОЛЬЗОВАТЕЛЕЙ В WEB - ПРИЛОЖЕНИЯХ

2.3.6. Методы и системы защиты информации, информационная безопасность

Диссертация на соискание ученой степени
кандидата технических наук

Научный руководитель
кандидат технических наук, доцент
Красов Андрей Владимирович

Санкт-Петербург – 2025

Оглавление

ВВЕДЕНИЕ

Актуальность темы исследования. Сегодня мировой рынок поведенческой биометрии находится в фазе активного роста (по данным ResearchAndMarkets вырос с 2,14 млрд долларов в 2023 году до 2,57 млрд долларов в 2024 году. Ожидается, что он продолжит расти со среднегодовым темпом роста 20,64% и достигнет 7,97 млрд долларов к 2030 году).

Поведенческие биометрические системы внедряются повсеместно, например: в онлайн-банкинге, электронной коммерции, платежах и на рынках аутентификации с высоким уровнем безопасности.

Рост рынка биометрических поведенческих систем обусловлен новыми тенденциями и вызовами, с которыми столкнулось общество и государство. К таким вызовам относятся: увеличение объемов данных о действиях пользователей в сети Интернет, которые могут быть использованы в злоумышленных целях (обострились проблемы приватности, анонимности пользователей и защищенности биометрических шаблонов от компрометации); использование методов искусственного интеллекта для проведения атак с использованием затенения и мошенничества, а также фальсификации биометрических изображений человека на основе физиологических биометрических измерений, таких как отпечатки глаз, отпечатки лиц и отпечатки пальцев.

В настоящее время поведенческий биометрический анализ используется в качестве метода аутентификации, потому что технология поведенческой биометрии предлагает надежную, соответствующую риску аутентификацию личности и меры по борьбе с мошенничеством, которые не требуют усилий со стороны пользователей и не требуют специального оборудования или дополнительных мер безопасности.

Данный метод обеспечивает: Гибкость — практически безграничный набор поведенческих биометрических характеристик доступен для анализа, а выбранные функции можно легко настроить в соответствии с конкретными потребностями использования; Удобство — поведенческая биометрия анализирует характерное

поведение пользователя устройства, не нарушая пользовательский опыт; Эффективность — для аутентификации личности поведенческая биометрия применяется в режиме реального времени и работает одновременно с устаревшими механизмами аутентификации, такими как ввод пароля. Для обнаружения мошенничества биометрический поведенческий анализ значительно сокращает время, необходимое для выявления и дифференциации мошенничества от законного поведения пользователя; Безопасность — поведенческая биометрия — это внутренние характеристики, которые людям чрезвычайно трудно распознать и практически невозможно воспроизвести, особенно когда одновременно исследуются несколько поведенческих характеристик.

Поскольку пользователи могут быть зарегистрированы в фоновом режиме во время нескольких обычных взаимодействий, поведенческая биометрия абсолютно беспроблемна и не замедляет, не прерывает и иным образом не мешает сеансу пользователя.

Однако для веб-приложений метод поведенческой биометрии не рассматривается в качестве основной системы аутентификации из-за изменяющейся баллистической природы этого метода и опасений по поводу неверного чтения шаблонов поведенческой биометрии пользователя, путем принятия недействительного пользователя в качестве действительного и наоборот.

Так же рассматривается метод статической аутентификации - тип управления доступом, обычно используется в качестве одноразовой проверки личности во время первого процесса входа в систему. Для всего сеанса предполагается, что пользователь является законным. При создании любого веб-приложения необходимо учитывать угрозы безопасности и уязвимости, которым может подвергнуться пользователь во время входа в систему. Необходимо создать интегрированную систему аутентификации, основанную на проверке личности пользователя с самого начала процесса входа в систему до момента завершения сеанса. До этого момента не существует системы биометрической поведенческой аутентификации, основанной на аутентификации пользователя на всех этапах использования веб-приложения и с наименьшими затратами.

В связи с этим современная высоконадежная поведенческая биометрическая система должна быть статической и непрерывной, и строиться на основе алгоритма динамики нажатия клавиш клавиатуры и мыши; с использованием алгоритма и метода непрерывной аутентификацией для классификации поведения пользователя и принятия решений на их основе для снижения риска принятия ложного пользователя за действительного.

Настоящее диссертационное исследование посвящено решению **научно-технической проблемы**, которая заключается в повышении надежности многофакторной поведенческо-биометрической аутентификации (статической и непрерывной) и защищенности поведенческо-биометрических систем от хакерских атак на основе технологии исполнения алгоритмов динамика нажатия клавиш и мыши.

Степень разработанности темы. Динамическая аутентификация пользователей, основанная на использовании рукописного ввода на клавиатуре и динамики движения мыши, является достаточно перспективным направлением исследований и широко применяется для обеспечения безопасности несанкционированного доступа злоумышленников и защиты данных пользователей.

Первые исследования в области анализа динамики нажатия клавиш были проведены в 1980-х годах Национальным научным фондом и Национальным бюро стандартов и биометрии, а анализ на основе динамики мыши был впервые предложен Ахмедом и соавторами в 2007 году. В качестве анализа привычек использования мыши человеком был сделан вывод о том, что шаблоны набора текста и стиль жестов мыши имеют уникальные характеристики, которые можно идентифицировать и использовать в качестве критериев идентификации и проверки. Значимые результаты в области анализа динамики нажатия клавиш и мыши на основе аутентификации пользователя были получены в работах

российских и зарубежных ученых, таких как: Р.В. Киричк, В.И. Коржик, Q. Zhou, Y. Yang, F. Hong, Y. Feng, Z. Guo, R. Maxion, N. Zheng, A. Paloski, H. Wang, S. J. Quraishi, S. S Bedi, C. Shen, Z. Cai, X. Guan, P. Kasprowski, Z. Borowska,

K. Harezlak, C. Shen, Z. Cai, X. Guan, Y. Deng, Y. Zhong, J. Gaikwad, B. Kulkarni, N. Phadol, S. Sarukte, / M. Seeger, B. Bours, E.L. Gaines, E. Rybnik, S.H. Pin, S. Deian, Y. Zhong, I.H. Shimaa, H.Z. Hala, M.S. Mazen, G. Jyotsna, Bryan, J.V. Harter, Monaco, N. Benkelman, P. Bours, S. Mondal, Y. Deng. А.Р. Абзалов, И.И. Кашапов, А.Ю.Орлов, И.Р. Мамлеев, Е.А. Кочегурова, Ю.А. Мартынова, А.А. Стрельников, М.В. Тумбинская, М.А. Казачук, N. Altwaijry, O.A. Salman, S.M. Hameed, J. Kim, P. Kang, H. Kim, В этих работах уделяется внимание методам извлечения биометрических характеристик нажатий клавиш и мыши, а также представлению различных методов проверки и аутентификации пользователей с использованием фиксированного текста данных о нажатиях клавиш и нейронных классификационных сетей. Проанализирована точность и эффективность используемой рабочая характеристика приёмника (ROC, receiver operating characteristic) на основе проведения эксперимента на группе образцов и расчета доли ложных отклонений и доли ложных приемок. Продление сеанса пользователя обеспечено созданием системы непрерывной аутентификации на основе динамических нажатий клавиш и систем нейронной классификации как (SVM, Support Vector Machine).

Недостатком этого подхода является то, что из-за различий в формате или размере обучающих данных разные алгоритмы классификации могут давать разную частоту ошибок, и, кроме того, показатели точности могут сильно различаться между отдельными пользователями и зависеть от общего числа пользователей в базе данных. Выбор способа анализа набора текста с клавиатуры и жестов мыши может увеличить время, необходимое для принятия правильного решения о проверке, а также повысить риск кражи, компрометации или несанкционированного доступа к конфиденциальной информации.

Объект и предмет исследования. Объектом исследования является система статической и непрерывной многофакторной аутентификации на основе поведенческо-биометрического подчерка.

Предметом исследования является поведенческо-биометрическая аутентификация.

Цель диссертационной работы: *Целью* работы является повышение точности многофакторной и непрерывной поведенческо-биометрической аутентификации на основе динамики нажатия клавиш клавиатуры и мыши.

Для достижения цели исследования в работе решена *научная задача*: Разработка системы многофакторной аутентификации на основе биометрических измерений динамики нажатий клавиш и мыши и мониторинга поведения пользователя во время сеанса.

Данная научная задача подразделяется на следующие частные *задачи*:

- Создать модель, извлекающую все биометрические характеристики нажатий клавиш и движений мыши и разработать модели идентификации руки на основе динамики нажатия клавиш;
- Создать модель для генерации случайного одноразового пароля на основе динамики нажатий клавиш и создать трехфакторную технологию аутентификации пользователей и субъектов доступа для веб-приложения;
- Создать непрерывную систему аутентификации на основе динамики мыши при использовании веб-приложений, с помощью кинематики и расстояния Левенштейна;

Научная новизна результатов исследования. Научная новизна полученных результатов состоит в следующем:

- Создана модель двухфакторной аутентификации веб-приложений, которая способна идентифицировать пользователя с высокой точностью, в отличие от известных систем аутентификации. Предлагаемая модель аутентификации построена на основе поведенческих и мягких биометрических измерений нажатий клавиш и мыши. Чтобы найти отдельное пороговое значение для каждого пользователя, расстояние, полученное от клавиатуры, было найдено путем объединения трех расстояний: Манхэттенского расстояния, Евклидова расстояния и расстояния Чебышева, чтобы найти прямоугольный треугольник и вычислить теорему Пифагора для нахождения угла, прилежащего к гипотенузе, как отдельного порогового значения для каждого

пользователя, чтобы уменьшить значение частоты ложного отклонения и ложного принятия. Для нахождения порогового значения через данные, полученные от мыши, используются расстояние Минковского, которое рассчитывается через кривую четверти круга, и Манхэттенское расстояние, которое находится через площадь четверти круга и длину дуги четверти круга. Извлекаются все биометрические характеристики нажатий клавиш и движений мыши через значение временной метки каждого нажатия кнопки на клавиатуре и каждого движения мыши, совершаемого пользователем. Разработана модель для идентификации пишущей руки, чтобы добавить степень безопасности, позволяющую идентифицировать пользователя на основе динамики нажатий клавиш с использованием законов движения кинематики. В результате повышается количество использованных биометрических систем до 3; количество извлечённых поведенческо-биометрических характеристик до 21; скорость обработки данных ~ 0.37 С; снижается уязвимость от брутфорс атак до ~ 8%. Степень точности системы по разработанной методики составляет 97.9%. Эффективность динамики нажатия клавиш повышается на 4%, динамики мыши на 2%, определения рук на 10%.

- Создана Создана методика многофакторной аутентификации пользователей веб-приложения на основе генерации случайного пароля с учетом модели биометрического клавиатурного подчерка пользователя, которая способна идентифицировать пользователя с низкими затратами и высокой скоростью, В отличии от известных предложенная методика многофакторной аутентификации основана на использовании множества способов измерения расстоянию Джакара для принятия решения будут проходить измерения тестирования через Манхэттенское или Евклидово расстояние. При наименьших затратах и скорости реализации он превосходит другие методы аутентификации из-за отсутствия зависимости от внешних устройств. В результате, количество факторов аутентификации повышается до 3; снижается уязвимость от брутфорс

атак до ~10%; снижается уязвимость связи с фишинговыми атаками до ~ 5%; скорость обработки данных ~ 0.12%; уменьшаются затраты на внедрение системы на ~ 85%. Степень точности системы по разработанной методики составляет 93%.

- Модель Созданная система непрерывной аутентификации пользователей на основе разделения пространства web-страниц на сектора с четырьмя особыми типами динамики мыши. Каждое из движений представляют соответствующие метрики, с использованием расстояния Левенштейна, которое рассчитывает отличия от обучающей выборки. В отличии от известных, предложенная непрерывная динамическая аутентификация позволяет проверять аутентификацию на всем времени работы с приложением, учитывает не только клавиатурный подчерк пользователя, но и динамику движений мыши с использованием расстояний Левенштейна, Манхэттенского, Евклидова, векторного и Минковского. За счет этого удалось снизить число ложно положительных решений на 3.4%, ложно отрицательных на 1.8%, и сократить время выявления аномалий в поведении пользователя на 4%. Благодаря этому удалось повысить точность аутентификации до 97.2%, по сравнению с предыдущими результатами. Эффективность повышается на 2%.

Теоретическая и практическая значимость работы

Теоретическая значимость работы заключается в следующем:

1. Заключается в построении модели учитывающий большее число факторов биометрического подчерка пользователя по нажатию клавиш клавиатуры и мыши, использование которых позволяет создать более эффективные алгоритмы аутентификации.
2. Заключается в сочетании в методике различных методов изменения расстояния, осуществления процедуры аутентификации на всех этапах работы пользователя с web-приложением, учете особенностей клавиатурного подчерка в процессе генерации одноразовых паролей, что позволяет создать более надежные системы аутентификации

пользователей.

3. Заключается в создании непрерывной аутентификации на всем этапе работы web приложений используя динамику движения мыши, т.е. без привлечения дополнительного оборудования. Применяются методы определения расстояния Евклидова, Манхэттенского, векторного расстояния и расстояния Минковского. Использование всех перечисленных методов позволит разрабатывать программное обеспечение, повышающее точность аутентификации web-приложений.

Практическая значимость диссертации заключается в том, что:

1. Использование предложенной модели позволят более эффективно решать задачи построения программных систем идентификации пользователей web-приложений не только на этапе запуска, но и на всем протяжении его работы без использования дополнительного оборудования. Модель рассматривается, как безопасная, экономичная и надежная система с классификацией степени точности результатов и возможностью сокращения времени аутентификации за счет отсутствия прямого контакта с пользователем.
2. Заключается в возможности создания систем аутентификации пользователей web-приложений, усиленных одноразовыми паролями, дополнительно проверяемыми по уникальному клавиатурному подчерку пользователя, что особенно актуально для систем онлайн-платежей и подтверждения покупок в интернет-магазинах.
3. Заключается в том, что за счет использования предложенных непрерывная аутентификации на основе динамического динамики движение мыши при работе с web приложениями, удаётся создать более точную систему аутентификации. Предлагаемая система непрерывной аутентификации может применяться в банковских системах, интернет-магазинах и других ресурсах доступ к которым осуществляется с помощью web-приложений.

Методология и методы исследования. Для решения задач, представленных в диссертации, были использованы поведенческие биометрические измерения, основанные на динамике нажатия клавиш, динамике мыши и мягких биометрических измерениях для определения рукописного текста на клавиатуре, реализации Евклидова, Манхэттенского, векторного расстояния, расстояния Минковского, Чебышева для определения порогового значения и аутентификации пользователя, разрабатывающий одноразовый пароль или метод OTP для генерации случайного пароля. обеспечения дифференциальной конфиденциальности данных и знаний, идентификации и аутентификации. Модель непрерывной аутентификации, основанная на биометрических измерениях динамики движения мыши, реализации законов движения кинематика, моделировании предлагаемого метода многофакторной и непрерывной аутентификации. Аутентификация реализована на основе веб-приложения, разработанного на языках программирования PHP, JavaScript, HTML, CSS, JQuery, и с использованием базы данных PHPMyAdmin.

Положения, выносимые на защиту:

1. Биометрическая модель аутентификации пользователя на основе динамики нажатия клавиш и мыши.
2. Методика трех факторной аутентификации пользователей для веб-приложения.
3. Динамическое Непрерывная аутентификации пользователей и субъектов доступа для веб-приложения в процессе работы Степень достоверности и апробация результатов.

Степень достоверности и апробация результатов.

Достоверность результатов, обоснованность положений и выводов, сформулированных в диссертации, обеспечивается учетом большого количества факторов, влияющих на решение поставленной научной задачи; обоснованным выбором основных допущений и ограничений, принятых в качестве исходных данных при ее постановке; использованием современного математического аппарата; обсуждением результатов диссертационной работы на конференциях;

публикацией основных результатов диссертации в ведущих рецензируемых журналах.

Апробация результатов. Основные результаты диссертации докладывались и обсуждались на конференциях

«Актуальные проблемы инфотелекоммуникаций в науке и образовании» (Санкт- Петербург, 2023–2024); «Региональная информатика (Ри-2024)» (Санкт-Петербург, 2024); «подготовка профессиональных кадров в магистратуре в эпоху цифровой трансформации (пкм-2024) » (Санкт- Петербург, 2024); «научно-техническая конференция профессорско-преподавательского состава, научных работников и аспирантов (НТК ППС 2025)» (Санкт- Петербург, 2025); «Международной научно-практической конференции, Астрахань» (Астрахань, 2021).

Публикации по теме диссертации. Всего по теме диссертации опубликовано 12 работ, из них 4 статьи в рецензируемых научных журналах, входящих в перечень изданий, рекомендуемых ВАК Минобрнауки России, 1 Регистрация программы для ЭВМ, 7 статей в журналах и сборниках конференций, включенных в РИНЦ.

Соответствие паспорту специальности. Содержание диссертации соответствует следующим пунктам паспорта специальности 2.3.6 Методы и системы защиты информации, информационная безопасность: п.12. Технологии идентификации и аутентификации пользователей и субъектов информационных процессов. Системы разграничения доступа.

Личный вклад автора. Все научные результаты получены автором лично, что подтверждается наличием личных публикаций. Личный вклад автора заключается в анализе систем и факторов аутентификации, а также принципов построения системы многофакторной аутентификации и непрерывной аутентификации на основе поведенческой и мягкой биометрии. Результаты теоретических и экспериментальных исследований получены автором самостоятельно.

Структура и объем диссертации. Диссертации состоит из введения, четырех глав с выводами по каждой из них, заключения, списка литературы.

Общий объем работы – (176) страницы, из них основного текста (153) страниц. Работа содержит (47) рисунок и (8) таблицы. Список литературы включает 198 источников.

ГЛАВА 1. АНАЛИТИЧЕСКИЙ ОБЗОР РЕШЕНИЙ И АЛГОРИТМОВ ПО ПОВЕДЕНЧЕСКОЙ БИОМЕТРИЧЕСКОЙ АУТЕНТИФИКАЦИИ ПОЛЬЗОВАТЕЛЕЙ

1.1 Значимость веб-приложений и их проблематика

Более трех миллиардов человек по всему миру пользуются Интернетом, а также веб-приложениями на различных устройствах благодаря простоте использования и легкому доступу в любом месте и в любое время [1]. В данный момент веб-приложение является первым шагом на пути автоматизации основных действий повседневной жизни, путем модернизации существующих решений. По вышеупомянутой причине, большинство организаций или поставщиков услуг, таких как промышленность, банковское дело, правительство, образование, медицина и другие секторы, хотят предоставлять свои услуги клиентам через веб-приложения.

Веб-приложения являются одной из наиболее распространенных целей для взлома, поскольку они обеспечивают легкий доступ к более широкой аудитории, позволяя вредоносному коду распространяться быстрее. Но, увы, многие компании серьезно задумываются о том, как защитить компанию от уязвимостей веб-приложений только после того, как инцидент уже произошел [1].

Веб-приложения могут быть атакованы по разным причинам, включая системные недостатки, вызванные неправильным кодированием, некорректно настроенными веб-серверами, недостатками дизайна приложения или неспособностью проверить формы. Любое веб-приложение имеет по крайней мере одну уязвимость, которую хакеры могут использовать [2].

Такие уязвимости веб-приложений позволяют злоумышленникам получать прямой и публичный доступ к базам данных, содержащим ценную информацию (например, финансовые данные или персональные данные), что делает их частой целью атак.

Одной из основных проблем, с которыми сталкиваются веб-приложения, являются ошибки идентификации и аутентификации [3]. В цифровом мире,

который становится все сложнее, ошибки аутентификации являются относительно распространенными уязвимостями безопасности в веб-приложениях. Если функции идентификации пользователя, аутентификации или управления сеансами вашего веб-приложения не реализованы точно или не защищены должным образом, это может привести к уязвимости.

Сама суть этих уязвимостей веб-приложений заключается в тонком балансе предоставления доступа законным пользователям и запрета входа неавторизованным лицам. Любая ошибка в поддержании этого баланса, например, использование слабых методов аутентификации, неправильное обращение с токенами сеанса или пренебрежение мерами безопасности для процессов восстановления пароля, может послужить приглашением для злоумышленников.

Уязвимости аутентификации в веб-приложении могут включать атаки методом подбора, когда хакеры пытаются использовать многочисленные комбинации паролей, пока не попадут в нужную [4]. Или это могут быть не хешированные и недостойно соленые пароли, которые гораздо легче взломать. Утечки данных, связанных с учетными записями пользователей, неправильно установленное время остановки, из-за которых сеансы остаются открытыми дольше, чем необходимо, или даже что-то, казалось бы, безобидное, например, слабые пароли, такие как «password1» или «admin1234», все это уязвимости веб-приложений, которые можно эксплуатировать [4]. Из-за тревожных угроз и компрометации учетных данных аутентификации пользователей организации переходят на строгую аутентификацию, чтобы защитить конфиденциальную информацию пользователей от злоумышленников.

1.2 Анализ основ системы аутентификации

Аутентификация — это обязательный процесс проверки личности пользователя и ограничения доступа неавторизованных пользователей к системе [5, 6].

Идентификация — первоначальный процесс подтверждения личности пользователей путем запроса их учетных данных. Например, пользователь вводит

идентификатор (имя пользователя) для входа в систему или сеть, где имя пользователя является уникальным идентификатором человека в этой системе [5, 94].

Авторизация — это второй шаг в процессе сквозной аутентификации, который устанавливает учетные данные пользователя в вычислительной системе на определенный период времени на основе определенной политики использования [5, 6].

Основные факторы аутентификации делится на четыре типа аутентификации:

1. Фактор знания: является наиболее распространенным фактором, который может быть паролем или простым персональным идентификационным номером (ПИН-кодом). Однако его также легче всего взломать. При использовании паролей важно использовать надежные пароли. Надежный пароль представляет собой смесь заглавных и строчных букв, цифр и специальных символов. В прошлом специалисты по безопасности рекомендовали, чтобы пароли были длиной не менее восьми символов. Однако с ростом надежности взломщиков паролей все чаще можно услышать, как специалисты рекомендуют более длинные пароли [5, 6].
2. Фактор владения: поскольку люди забывают вещи и теряют их, можно было бы подумать о том, чтобы основать схему аутентификации для людей на чем-то, чем является человек. В конце концов, мы узнаем людей, с которыми взаимодействуем, не из-за какого-то протокола паролей, а из-за того, как они выглядят или как они звучат — «что-то, чем они являются». Аутентификация, основанная на «чем-то, чем вы являетесь», будет использовать поведенческие и физиологические характеристики принципала. Эти характеристики должны легко и точно измеряться и, желательно, быть вещами, которые трудно угадать [5, 6].
3. Фактор свойства: включает в себя все предметы, которые являются физическими объектами, такими как ключи, смартфоны, смарт-карты, USB-накопители и устройства-токены [5, 6] (Устройство-токен выдает

PIN-код с ограниченным сроком действия или может вычислять ответ на основе номера запроса, выданного сервером.)

4. Фактор местоположения: этот тип процесса аутентификации использует местоположение пользователя для определения его/ее личности. Как правило, для этого типа аутентификации используются глобальная система позиционирования (GPS), IP-адрес, идентификатор вышки сотовой связи и т. д. Как правило, он используется в сочетании с другой аутентификацией для проверки личности пользователя. Если у пользователя есть правильные учетные данные, то система будет проверять, где он/она находится, чтобы проверить, авторизован ли пользователь для доступа к системе из этого конкретного местоположения [5, 6].

Организации и предприятия выбирают различные типы методов аутентификации в зависимости от желаемой цели безопасности, стоимости и предпочтений пользователей.

Методы аутентификации

1. **Однофакторная аутентификация (SFA):** метод аутентификации, использующий комбинацию имени пользователя и пароля, является наиболее популярным, как показано на рисунке 1.1 [6, 7]

   Благодаря своей простоте и легкости в использовании SFA широко использовался, например, с использованием пароля (или PIN-кода) для проверки личности пользователя. Пароли состоят из комбинации букв, цифр и специальных символов. Чем сложнее комбинация вышеперечисленного, тем надежнее пароль и, следовательно, тем труднее его обнаружить злоумышленнику.

Каждая существующая система имеет свои преимущества и недостатки, и одним из ее недостатков является то, что многие люди выбирают простые пароли, а не надежные. Очень простые пароли, которые могут включать имя пользователя, дату рождения и т. д., уязвимы для фишинговых атак. Пароли имеют множество недостатков и сами по себе больше не эффективны для защиты данных, доступ и

передача к которым происходит через Интернет. Безопасность учетной записи пользователя может быть скомпрометирована, если пароль будет раскрыт или обнаружен. Неавторизованный пользователь может использовать грубую силу в виде атак по словарю или методов социальной инженерии для получения доступа. Хакеры могут просто использовать свободно доступные инструменты, которые могут быть автоматизированы для подбора пароля пользователя путем перебора всех возможных комбинаций, пока не будет найдено совпадение [5].

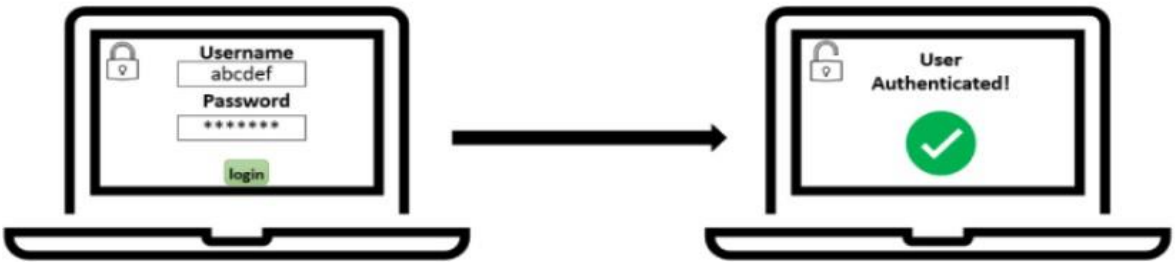


Рис. 1.1. Однофакторная аутентификация (SFA).

2. Двухфакторная аутентификация (2FA): комбинируя репрезентативные данные (комбинацию имени пользователя и пароля) с дополнительной формой идентификации, такой как фактор личной собственности, который может включать безопасный токен с использованием одноразового пароля, двухфакторная аутентификация (2FA) усиливает безопасность (OTP) [6, 7, 94]. Использование этой формы идентификации требует использования дополнительного механизма, которым может быть физический компонент или электронное устройство, такое как мобильный телефон, планшет или компьютер как показано на рисунке 1.2. На начальном уровне аутентификации пользователь должен предоставить физическое устройство или OTP, который был отправлен по электронной почте, SMS или другому устройству [8].

В результате, если пароль пользователя украден, злоумышленник больше не имеет доступа ко второму механизму аутентификации, что еще больше повышает безопасность личной информации пользователя.

Один из её недостатков - без обоих механизмов аутентификации даже авторизованный пользователь не может получить доступ [9].

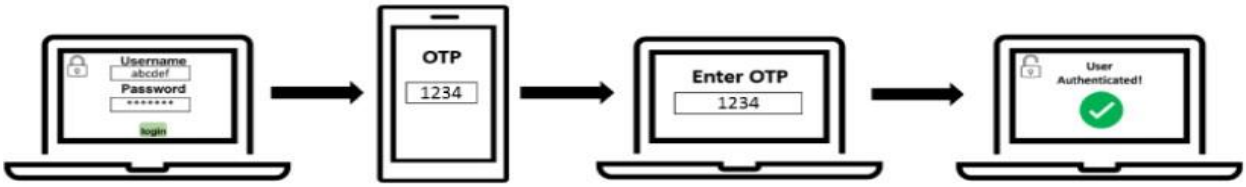


Рис. 1.2. Двухфакторная аутентификация (2FA).

3. В настоящее время требуется больше степеней безопасности, поскольку атаки становятся все более целенаправленными, а несанкционированный доступ имеет катастрофические последствия. Это особенно характерно для платформ, которые обрабатывают личные данные, такие как банковские операции. Теперь необходим больший контроль над проверкой личности человека, пытающегося получить доступ к этим услугам. Нет никаких сомнений в том, что при этих дополнительных требованиях обеспечиваемая защита намного выше, но в некоторых случаях она все еще недостаточна. В результате для повышения безопасности требуются дополнительные этапы аутентификации. Чтобы снять наличные в банкомате (ATM), пользователь должен предъявить физический токен (банковскую карту), представляющий фактор владения, и PIN-код, представляющий фактор знаний, чтобы получить доступ к личному счету и вывести средства. Эту систему легко можно сделать более безопасной, включив второй биометрический механизм. Пример показано на рисунке 1.3.

Многофакторная аутентификация (MFA) — это безопасный процесс аутентификации, который требует более одного метода аутентификации, выбранного из независимых категорий учетных данных. Многофакторная аутентификация все чаще используется для проверки личности пользователей при доступе к киберсистеме и информации. MFA объединяет два или более типов аутентификации для обеспечения лучшего и безопасного способа аутентификации пользователей [6,10].

Благодаря интеграции критериев знания и владения с биометрией для улучшения проверки личности, биометрия поддерживает многофакторную

аутентификацию, что затрудняет злоумышленникам обман системы путем выдачи себя за другое лицо.

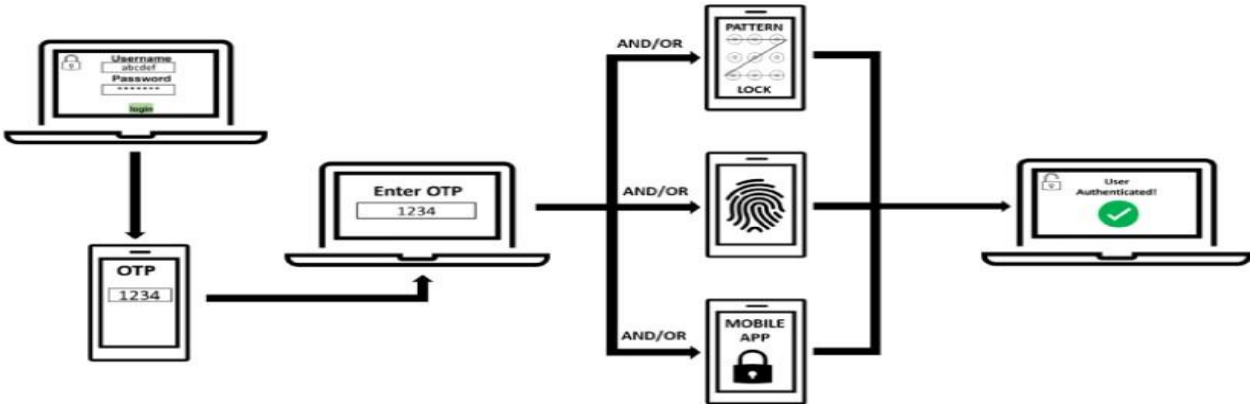


Рис. 1.3. Многофакторная аутентификация (MFA).

В течение очень долгого времени предприятия защищали свои данные и активы с помощью паролей. Но по мере увеличения скорости обработки и усложнения методов криптоанализа защита паролем стала менее эффективной. В результате требуется более сложный метод аутентификации. Одним из таких методов является биометрия.

Биометрическая аутентификация — это процесс распознавания людей по их физиологическим и поведенческим чертам или атрибутам. Термин «биометрический» представляет собой комбинацию двух греческих слов, а именно Bio (жизнь) и Metric (измерять), и относится к типу аутентификации «то, чем вы являетесь». Таким образом, биометрическая аутентификация фактически измеряет, а также анализирует биологические и поведенческие атрибуты человека. Во многих ситуациях биометрическая аутентификация способна обеспечить более высокую точность идентификации пользователя, чем типичные системы на основе паролей [9, 11].

Международное биометрическое общество и другие группы постоянно изучают новые биометрические признаки, которые просты в использовании, ненавязчивы и обеспечивают наиболее точные результаты при идентификации. В статье Речника [12], опубликованной в выпуске Computer Magazine за сентябрь 2014 года под названием «За пределами узнавания: обещание биометрической аналитики», упоминается, что вскоре биометрическая аналитика будет

использоваться для обнаружения потенциально интересной информации о человеке, помимо проверки личности с использованием шаблонов биометрических сигналов.

на рисунке 1.4. показано области аутентификации биометрические измерения подразделяются на два основных типа:

– физиологические биометрические данные, которые связаны с формой человеческого тела и его особенностями, так что по мере изменения формы и геометрии тела эти биометрические данные необходимо обновлять, чтобы избежать сбоя аутентификации. Некоторые примеры этого типа - отпечаток пальца, лицо и радужная оболочка глаза [5];

– поведенческая биометрия связана с конкретным типом поведения человека, поэтому аутентификация может не удаться, если текущий шаблон поведения отличается от сохраненного шаблона. При практике примеры этого типа включают в себя динамику нажатий клавиш, мыши, подпись и анализ звука [5].

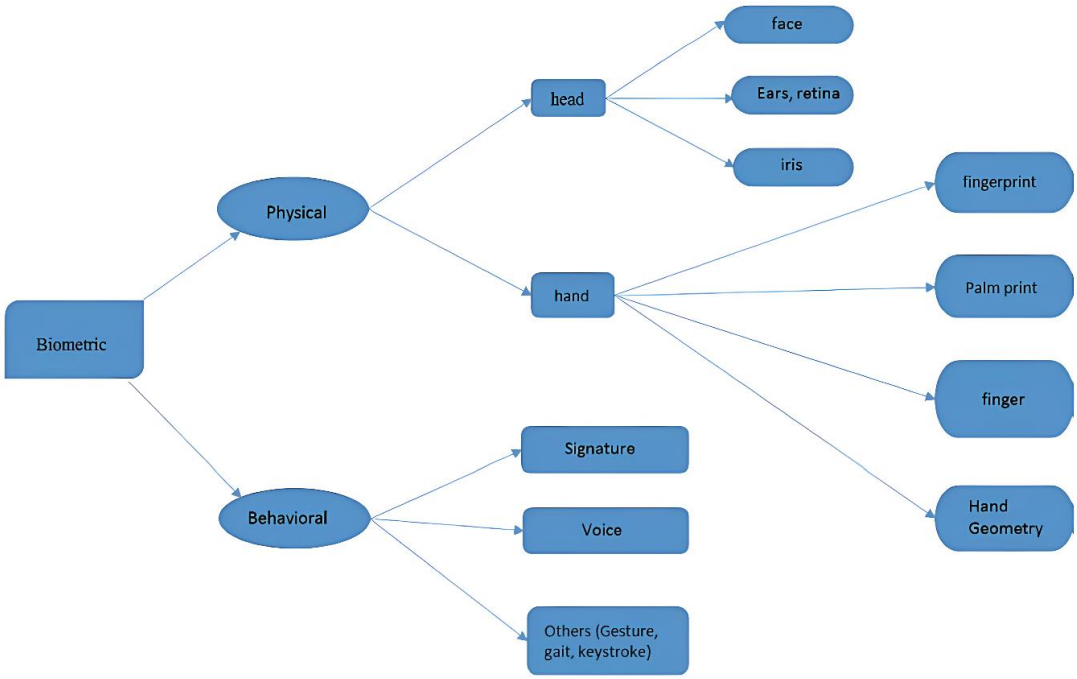


Рис. 1.4. Физиологические и поведенческие черты, используемые в различных биометрических методах аутентификации

1.3 Анализ уязвимостей безопасности в системе аутентификации веб-приложений

По данным Statista, в 2020 году более 1000 утечек данных раскрыли более 155 миллионов записей. Средняя стоимость этих вторжений составила три с половиной миллиона долларов [13]. Еще более тревожным является тот факт, что причиной почти 82% утечек стали слабые или украденные учетные данные.

Годовой размер денежного ущерба, причиненного зарегистрированными киберпреступлениями в США с 2001 по 2023 гг. По данным Центра жалоб на интернет-преступления США (IC3), в 2023 году сумма денежных потерь, связанных с киберпреступностью, выросла почти на 21%, достигнув рекордного уровня в 12,5 млрд долларов как показано на рисунке 1.5.

Одной из крупнейших проблем, с которой сталкиваются правительства по всему миру, по-прежнему является киберпреступность. Более 300 000 человек стали жертвами фишинговых атак в 2022 году, что делает фишинг и утечки персональных данных одними из наиболее часто регистрируемых видов киберпреступности в США. Кроме того, по состоянию на январь 2023 года средняя стоимость утечек данных для американских корпораций составила более девяти миллионов долларов.

Миллионы американцев пострадали от кражи личных данных, что было распространенной проблемой среди других зарегистрированных онлайн-преступлений. По оценкам, 13,5 миллионов американцев стали жертвами кражи личных данных, что делает страну второй по количеству зарегистрированных жертв в мире. Тридцать три процента жертв кражи личных данных не имели другого выбора, кроме как заморозить свои кредитные карты в январе 2023 года, а сорок три процента заявили, что потратили много времени, пытаясь решить эту проблему. Злоумышленники также иногда нападают на пожилых людей; в 2022 году более 4800 человек старше 60 лет сообщили о краже своих личных данных.

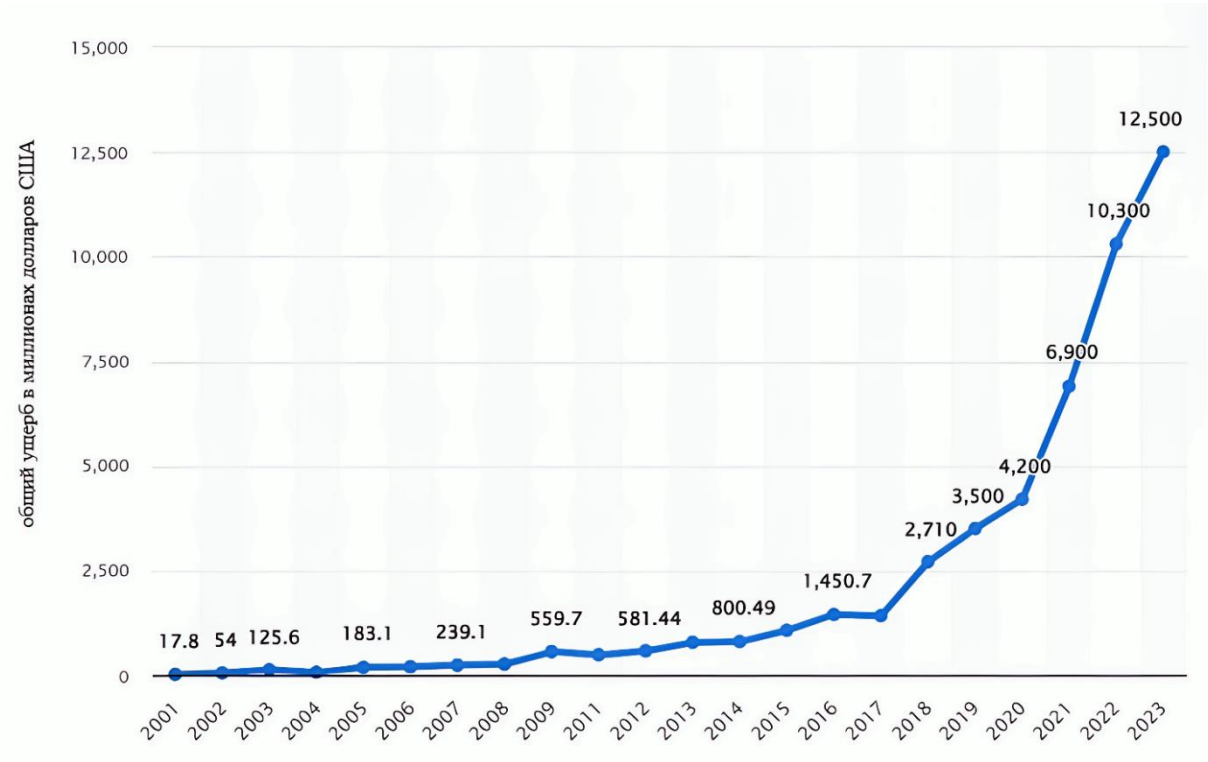


Рис. 1.5. Годовой отчет по данным Центра рассмотрения жалоб на интернет-преступления США (IC3)

Уязвимости аутентификации: проблемы, которые влияют на процедуры аутентификации и подвергают веб-сайты и приложения уязвимостям, в которых хакер может выдавать себя за настоящего пользователя.

Поскольку аутентификация — это всего лишь процесс идентификации личности пользователя, она является важным компонентом каждого веб-сайта или программы.

Уязвимости в безопасности сайта, учетных данных для входа или даже базовом коде могут привести к ряду типичных проблем аутентификации [14].

Уязвимости механизмов аутентификации [15] чаще всего возникают одним из двух способов:

1. Поскольку механизмы аутентификации не обеспечивают достаточную защиту от атак методом подбора, они неэффективны;
2. Злоумышленник может полностью обойти механизмы аутентификации с помощью логических ошибок или плохого реализационного кодирования. Это также называется «сломанной аутентификацией».

В некоторых случаях логические ошибки вызывают непредсказуемость веб-сайта, что может означать проблему безопасности. Однако, поскольку аутентификация так важна для безопасности, неправильная логика аутентификации может вызвать проблемы с безопасностью веб-сайта. Человеческий фактор является слабым звеном в аутентификации, поскольку человеческое поведение играет важную роль в возникновении уязвимостей аутентификации. Для удобства пользователи часто выбирают слабые пароли или повторно используют их на разных платформах, ставя удобство выше безопасности [16].

Фишинговые атаки используют человеческое доверие и доверчивость, обманывая неосторожных людей, заставляя их раскрывать свои пароли. Хакерам может быть проще проникнуть в учетные записи, когда люди не знают о безопасных процедурах аутентификации и принимают неверные решения [17].

Уязвимости аутентификации имеют серьезные последствия – будь то из-за слабых паролей или плохой разработки и реализации аутентификации.

В статье John Martinez [18] (11 распространенных уязвимостей аутентификации, о которых вам нужно знать», опубликована в сентября 2024) показал, что, злоумышленники могут использовать эти уязвимости для получения доступа к системам и учетным записям пользователей выполнения, следующих действия:

1. Кража конфиденциальной информации;
2. Имитация законного пользователя;
3. Управление приложением;
4. Полное уничтожение системы.

Если злоумышленники могут незаконно пройти процесс аутентификации и получить доступ к учетной записи пользователя, они могут украсть важные данные, такие как имена, данные кредитной карты, номера социального страхования, медицинские записи и налоговые идентификаторы. Они также могут предпринимать действия от имени пользователя, например инициировать

финансовые транзакции, удалять данные или передавать право собственности на ресурсы.

Одна из серьезных проблем - если злоумышленнику удастся получить доступ к высокоуровневым учетным записям, таким как профили администраторов, он может полностью захватить вашу систему или даже отключить ее. После обнаружения эти недостатки аутентификации могут серьезно навредить легитимности и устойчивости вашего бизнеса.

1.3.1 Изъяны безопасности в системах аутентификации веб-приложений

1. Слабые учетные данные для входа: Имя пользователя и пароль должны быть созданы, когда пользователи регистрируются для учетной записи на веб-сайте или в приложении, которое использует входы на основе пароля. Однако процедура аутентификации может стать уязвимой, если пароль легко угадать. Злоумышленникам может быть проще нацеливаться на конкретных людей с предсказуемыми именами пользователей. Злоумышленники будут искать учетные записи с часто используемыми, простыми паролями, а не использовать полномасштабную атаку методом подбора. Они попытаются использовать стандартные учетные данные для входа, такие как «admin», «admin1» и «password1». Если слабые пароли не ограничены, даже веб-сайты, защищенные от атак методом подбора, могут стать уязвимыми [19, 20];
2. Неисправная двухфакторная аутентификация: хотя двухфакторная аутентификация (2FA) хорошо подходит для безопасной аутентификации, если она не реализована должным образом, она может привести к серьезным проблемам безопасности. Если она предоставляется через SMS, злоумышленники могут использовать атаки с подменой SIM-карты для расшифровки четырех- и шестизначных номеров проверки 2FA. Кроме того, некоторая двухфакторная аутентификация на самом деле не является двухфакторной; например, «второй фактор», который отправляет сообщение на тот же телефон, не обеспечивает никакой дополнительной

безопасности, если пользователь пытается получить доступ к личным данным на украденном телефоне, используя учетные данные, которые были кэшированы. Если нет защиты от перебора, чтобы заблокировать учетную запись после заранее определенного количества попыток входа, могут также возникнуть уязвимости двухфакторной аутентификации [19, 20];

3. Отсутствие многофакторной аутентификации (MFA): одним серьезным недостатком, который многие игнорируют, является отсутствие MFA. Прося пользователей заполнить несколько форм проверки, прежде чем они смогут получить доступ к своим аккаунтам, MFA добавляет дополнительную степень защиты. Вы можете значительно повысить безопасность своего аккаунта от нежелательного доступа, включив MFA [5];
4. Недостатки защиты от подбора паролей: цель атаки методом подбора паролей, как и атаки по словарю, — получить несанкционированный доступ к системе или учетной записи пользователя путем ввода большого количества заранее подготовленных или случайно сгенерированных комбинаций имени пользователя и пароля, пока не будет найдена работающая. Безопасность учетных данных пользователя может быть поставлена под угрозу, если механизм защиты методом подбора паролей неисправен, например, брандмауэр, логика аутентификации или ошибка протокола SSH. Это позволяет злоумышленникам получить контроль над учетными данными и процессами входа [20];
5. Перечисление имен пользователей: перечисление имен пользователей технически не является недостатком аутентификации. Однако, снижая сложность других атак, таких как атаки методом подбора или плохие проверки учетных данных, оно может облегчить жизнь злоумышленника. Злоумышленники могут определить, какие имена пользователей являются действительными, что является проблемой перечисления имен пользователей. Не тратя время и деньги на тестирование большого

количества поддельных имен учетных записей, они могут попытаться захватить учетные записи пользователей, используя методы подбора [19];

6. Уязвимая логика аутентификации: программные приложения часто имеют логические недостатки. Это происходит из-за некачественного кодирования или дизайна, которые ставят под угрозу функционирование приложения, доступ к авторизации и аутентификации. Слабая защита безопасности, неправильное использование функциональности или пропуск этапа в процессе проверки могут привести к ошибочной логике приложения. Приложение может попросить пользователя ответить на контрольный вопрос, который, согласно логике, «знает только нужный человек». Однако ответы на такие запросы, как девичья фамилия матери пользователя или день рождения, часто легко найти. Из-за этого недостатка киберпреступники могут легко обойти аутентификацию и получить доступ к учетным записям без авторизации [21];
7. Человеческая халатность: по данным исследования (Shred-it) 2020, безответственность сотрудников была названа 31% руководителей высшего звена второй основной причиной утечек данных. В отличие от атак методом подбора, SQL-инъекций и обходов аутентификации, человеческая ошибка может привести к серьезным ошибкам аутентификации, которые гораздо проще использовать [21]. Подобная халатность охватывает такие действия, как:
   - Оставление компьютера разблокированным и работающим в общественном месте;
   - Потеря гаджетов из-за кражи;
   - Передача личной информации неустановленным лицам;
   - Написание плохо написанного кода.
8. Перехват сеанса: когда злоумышленник перехватывает и берет идентификатор сеанса пользователя, это называется перехватом сеанса или кражей сеанса. Веб-сайты должны использовать безопасные, случайно сгенерированные токены сеанса, которые трудно предсказать, и

защищенные маршруты связи, такие как HTTPS, чтобы избежать этого [21].

1.4 Обзор современных индустриальных решений по аутентификации пользователей на основе поведенческой биометрии

Сегодня на рынке постоянно растет количество промышленных решений для аутентификации пользователей и безопасности систем на основе поведенческой биометрии, такой как динамика нажатия клавиш, движение мыши, анализ подписи и походки. В то же время эти системы в прошлом анализировали только имя пользователя или пароль и ограничивались статической аутентификацией, однако сейчас активно развиваются системы, способные анализировать поведение пользователя непрерывно. Степень эффективности глобальных решений биометрической верификации измеряется ведущими институтами и компаниями, которые анализируют поведенческую биометрическую систему и выдают сертификат, который подтверждает, что компания, использующая поведенческие биометрические решения для аутентификации, является высокоэффективной и имеет высокую способность верификации пользователей, например, компания " Frost & Sullivan"

В данный момент мировой рынок поведенческой биометрии находится в фазе активного роста (как показано на рисунке 1.6 по данным ResearchAndMarkets вырос с 2,14 млрд долларов в 2023 году до 2,57 млрд долларов в 2024 году. Ожидается, что он продолжит расти со среднегодовым темпом роста 20,64% и достигнет 7,97 млрд долларов к 2030 году) [22].

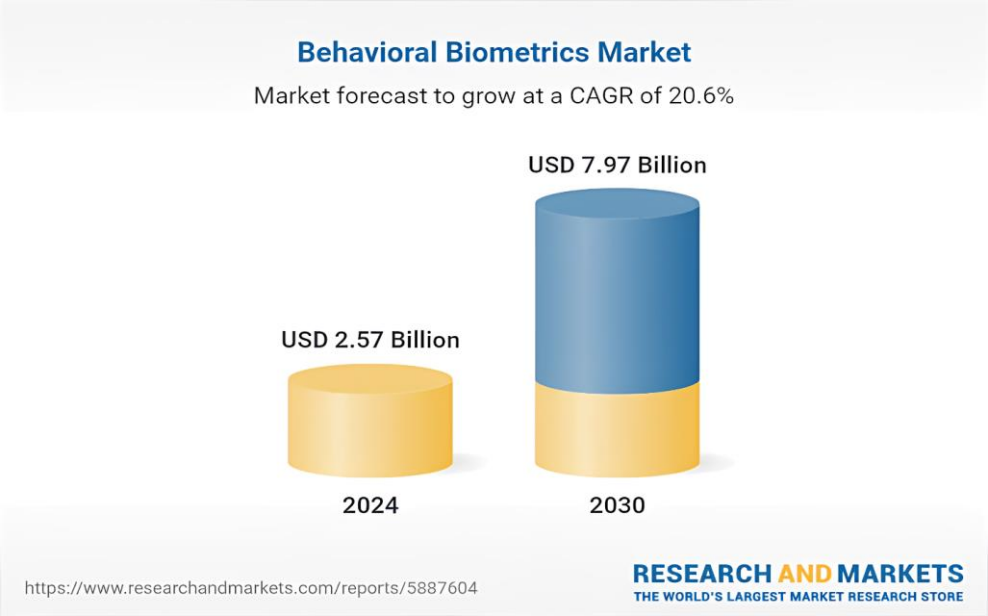


Рис. 1.6. Рост рынка поведенческой биометрии

Северная Америка управляла общей долей рынка поведенческой биометрии с точки зрения выручки в 2017 году благодаря значительному внедрению решений поведенческой биометрии, поддерживаемых хорошо зарекомендовавшей себя отраслью BFSI, и высоким расходам на ИКТ среди организаций на решения по безопасности данных по всему региону. Однако ожидается, что Азиатско-Тихоокеанский регион станет свидетелем самых высоких темпов роста в течение прогнозируемого периода из-за таких факторов, как рост расходов на безопасность среди государственных организаций и развитие рынка смартфонов в странах с развивающейся экономикой, таких как Китай, Япония, Индия, Австралия и Южная Корея [23].

Глобальные компании-разработчики программного обеспечения, основанные на программах аутентификации на основе поведенческой биометрии:

– **Typingdna:** это одно из самых популярных коммерческих решений в области аутентификации пользователей на основе динамики нажатия клавиш, как сервис с двумя основными продуктами безопасности: двухфакторной аутентификацией и постоянной аутентификацией. Компания была основана в 2016 году в Соединенных Штатах Америки [24].

Принцип работы основан на аутентификации по короткой фразе Для

анализа поведения пользователя в нем используются ритм и скорость набора текста, а также сила нажатия на клавиши.

аутентифицировать пользователей легко, попросив их ввести несколько слов. Вместо того, чтобы вводить коды TOTP или запоминать сложные фразы, пользователи вводят короткую фразу, выбранную им.

Аутентификация с помощью биометрии ввода текста работает на Android, iOS и React Native.

TypingDNA признана Frost & Sullivan лидером в категории решений для биометрической верификации 2023 г. TypingDNA была проанализирована по десяти критериям роста и инноваций в области биометрической верификации, сравнившись с 11 лидерами в этой области, включая Samsung, Panasonic, Hitachi ID, Fujitsu, DERMALOG, Toshiba, M2SYS, Suprema, Kaizen Voiz, Aratek и Iris ID.

– **BioCatch:** Одно из решений, реализующее аутентификацию на основе динамики и поведенческой биометрии пользователей, которое анализирует динамику работы пользователей с использованием клавиатуры и мыши и извлекает их характеристики для определения легитимности пользователя, но не уточняется, какие именно функции извлекаются и какие алгоритмы используются [25]. Разработчики хранят молчание из соображений защиты и безопасности. Данный продукт предлагается использовать для защиты пользователей онлайн банковских систем от доступа злоумышленников к их счетам. Аналитикам предоставляется широкий инструмент мониторинга, отображающий на различных графиках возможные риски и аномалии, а также формирующий высоко информативные отчеты о работе системы.
– **BehavioSec:** Шведская компания, основанная в 2007 году, которая специализируется на системах непрерывной аутентификации — программном обеспечении, которое отслеживает активность на компьютере, чтобы гарантировать, что компьютером пользуется именно настоящий владелец учетной записи. BehavioSec учитывает не только динамику нажатия клавиш, но и динамику мыши, а также общий способ

взаимодействия пользователя с компьютером [26].

Для анализа поведения пользователя в нем используются ритм и скорость набора текста, а также сила нажатия на клавиши. Однако, используемые для этого алгоритмы не называются.

- **SECUVE:** Ведущая компания в области мягкой поведенческой биометрии, основанная в Южной Корее в 2000 году. В ноябре 2017 года компания запустила мобильную платежную систему O2O или P2P для финансовых услуг в Соединенных Штатах Америки, основанную на мягких поведенческих измерениях подписи пользователя, где извлекаются все особенности движения пальцев во время процесса подписания, обучения и аутентификации во время процесса покупки. Используемые алгоритмы и степень точности не были упомянуты [27].
- **Plurilock:** Ведущий поставщик непрерывной аутентификации на основе поведенческой биометрии нажатий клавиш и мыши для строго регулируемых сред, таких как правительство, критическая инфраструктура, финансовые услуги и здравоохранение. Основан в 2016 году и проверяет личность пользователя каждые 3–5 секунд [28].
- **NoPassword:** недавно выпущенное коммерческое решение NoPassword, которое предлагает непрерывную многофакторную аутентификацию пользователей на основе их шаблонов использования клавиатуры и мыши, а также файловой системы их компьютера или планшета, также становится все более популярным. Даже при использовании сеансов удаленного рабочего стола для аутентификации ПК и планшета используются одни и те же методы, что демонстрирует высокое качество системы. Это решение определяет, сколько времени занимают нажатия клавиш и переходы при использовании клавиш клавиатуры, чтобы оценить почерк пользователя. Создатели рекомендуют это программное обеспечение для защиты ПК клиентов дома и на работе от нежелательного доступа хакеров. Пользователям предоставляется удобный для пользователя SDK и API NoPassword для интеграции этого

решения. Системным администраторам также предоставляется возможность динамически устанавливать элементы управления доступом и надежный инструмент для отслеживания активности пользователей в сети. Все наиболее широко используемые операционные системы и браузеры, доступные сегодня, поддерживают это решение. Алгоритмы машинного обучения этого продукта не идентифицированы [29].

1.4.1 Анализ недостатков существующей системы аутентификации пользователей на основе поведенческой биометрии

1. Используются дополнительные внешние устройства: некоторые из хорошо известных методов используют внешние устройства для извлечения и отслеживания динамических характеристик нажатия клавиш и мыши, такие как KeyTracker, микрофон для записи звука нажатий клавиш и другие устройства, которые могут быть дорогими и недоступными для выполнения всеми. процесс аутентификации;
2. Извлеченных биометрических поведенческих характеристик немного: чем больше биометрических поведенческих характеристик извлекается для конкретного метода, тем выше точность расчета и тем ниже уровень ошибок;
3. Отсутствует система многофакторной биометрической аутентификации: когда была придумана и создана идея двухфакторной аутентификации, целью было повышение безопасности. Уже была однофакторная аутентификация, но с применением множества методов аутентификации к разным системам, от веб-приложений до приложений для мобильных телефонов, стало ясно, что ее недостаточно для защиты пользовательских данных и системы в целом. Таким образом появилась концепция многофакторной аутентификации, подразумевающая объединение фактора того, что вы знаете, с тем, чем вы являетесь. Здесь действует правило: чем больше факторов, тем выше степень точности и защиты, и чем меньше факторов аутентификации, тем больше атак и взломов;

4. Слабый фактор аутентификации для защиты от грубой силы и фишинговых атак;
5. Небольшое количество биометрических факторов, используемых в процессах аутентификации, влияют на степень защиты: в известных методах процесс аутентификации осуществляется на основе поведенческой биометрической системы либо с использованием нажатий клавиш, либо с помощью динамики щелчков мыши, либо с помощью того и другого вместе. Здесь желаемые цели достигнуты, но точность недостаточна, чтобы судить о том, что результаты полностью удовлетворительны. Система имеет высокую степень безопасности и доверия. В дополнение к клавишам и мыши можно использовать и добавлять дополнительные поведенческие биометрические системы, поскольку они невидимы, мало контактируют с пользователем и не влияют на ход процесса аутентификации;
6. Отсутствует интегрированная система аутентификации для веб-приложения, которая связывает статическую аутентификацию с непрерывной аутентификацией, основанной на поведенческих биометрических измерениях движений мыши и нажатие клавши;
7. Этап обучения сложен или содержит много этапов: в известных методах этап обучения включает либо использование фиксированного текста, который пользователь должен повторить более одного раза, либо переменного текста, который пользователь должен повторить до восьми раз, в зависимости от системы. Некоторые из них требуют ввода сложного пароля — это заставляет пользователя использовать символы, буквы и слова, которые могут быть забыты в будущем;
8. Этап аутентификации включает в себя условия, находящиеся вне контроля пользователя: в процессе ввода учетных данных пользователя ожидается, что пользователь допустит ошибку, введя букву из пароля, и удалит неправильную букву, но в известных способах при нажатии клавиши (Backspace) пользователь вынужден повторно ввести свои

учетные данные;

9. Проблема изменения поведенческих биометрических данных из-за баллистической природы: нет сомнений в том, что одним из недостатков поведенческой биометрии является то, что баллистическая природа биометрической модели пользователя меняется со временем, так как со временем меняется стиль письма пользователя или стиль использования мыши, или даже мягкие биометрические методы, такие как указание пишущей руки;

   В известных методах баллистическая природа не учитывается, и поэтому эти системы аутентификации хороши вначале, и при экспериментировании можно получить точные результаты, но со временем точность этих систем снижается, и в результате возникает множество проблем, связанных с высокий уровень ложного отклонения и высокий уровень ложного принятия.

10. Результаты точности на этапе тестирования аутентификации доходят до 94.8%: это хороший процент, но системы аутентификации всегда стремятся к более высокому проценту и ближе к 100%. Это невозможно, потому что в настоящее время не существует системы аутентификации, которая могла бы получить полный процент, в связи с внешними условиями и естественными изменениями. С течением времени могут произойти некоторые дефекты, которые приведут к уязвимости.

1.5 Биометрические системы идентификации: законодательное регулирование в Российской Федерации

В сфере защиты прав субъектов персональных данных особое внимание уделяется биометрическим данным. Статья 11 Федерального закона «О персональных данных» дает следующее определение биометрических данных — «данные, которые характеризуют физиологические и биологические особенности человека, по которым можно установить его личность (биометрические персональные данные), используются оператором для идентификации субъекта

персональных данных. На обработку биометрических персональных данных необходимо согласие в письменной форме субъекта персональных данных, за исключением случаев, предусмотренных пунктом 2 настоящей статьи» [30]. Часть 2 статьи 11 Федерального закона «О персональных данных» устанавливает, что обработка биометрических персональных данных может осуществляться без согласия субъекта персональных данных в связи с реализацией международных договоров Российской Федерации о реадмиссии, в связи с осуществлением правосудия и всем, что с этим связано. Кроме того, законом предусмотрены особые требования к материальным носителям биометрических персональных данных и технологиям хранения таких данных, если персональные данные находятся вне информационных систем.

Материальный носитель должен обеспечивать:

- защиту от несанкционированного добавления и перезаписи информации после ее удаления из информационной системы персональных данных;
- доступ к биометрическим персональным данным, зафиксированным на материальном носителе, осуществляемый оператором и лицами, уполномоченными на работу с биометрическими персональными данными в соответствии с законодательством Российской Федерации;
- возможность идентификации информационной системы персональных данных, в которой осуществлена биометрическая запись персональных данных, а также оператора, осуществившего такую запись;
- предотвращение несанкционированного доступа к биометрическим персональным данным, содержащимся на материальном носителе.

Оператор биометрических данных, используемых вне информационных систем персональных данных, обязан:

- вести учет тиража материальных носителей;
- предоставить соответствующий материальный носитель с уникальным идентификационным номером, позволяющим точно определить

оператора.

1.6 Критерии и требования к поведенческой биометрической аутентификации

У каждой биометрической модальности есть свои плюсы и минусы. Более того, даже если некоторые из недостатков можно преодолеть, сама модальность может иметь присущие ей недостатки. Поэтому выбор биометрической черты для конкретного приложения зависит от проблем, помимо производительности сопоставления.

Рафаэль и Янг определили ряд факторов, которые делают физическую или поведенческую черту пригодной для биометрического приложения. Следующие семь факторов взяты из статьи Джейна 1998 [31]:

1- Универсальность: каждый человек, получающий доступ к приложению, должен обладать определенной чертой.

2- Уникальность: данная черта должна достаточно отличаться у разных членов популяции.

3- Постоянство: биометрическая черта человека должна быть достаточно инвариантной с течением времени относительно данного алгоритма сопоставления. Черта, которая существенно меняется, не является полезной биометрией.

4- Измеримость: должна быть возможность получения и оцифровки биометрической черты с помощью подходящих устройств, которые не доставляют человеку неоправданных неудобств. Кроме того, полученные необработанные данные должны поддаваться обработке для извлечения репрезентативных признаков.

5- Производительность: точность распознавания и ресурсы, необходимые для достижения этой точности, должны соответствовать требованиям приложения.

6- Приемлемость: лица из целевой группы, которые будут использовать приложение, должны быть готовы предоставить свои биометрические признаки

системе.

7- Обход: легкость, с которой биометрические признаки могут быть имитированы с использованием артефактов, например, поддельных пальцев в случае физических признаков и мимикрии в случае поведенческих признаков, должна соответствовать требованиям безопасности приложения.

Уникальность и постоянство — два элемента, которые имеют решающее значение для оценки эффективности. В то время как постоянство связано с необходимостью иметь возможность идентифицировать человека в течение более длительного периода времени, уникальность относится к необходимости различать двух разных людей.

1.7 Структурный анализ известных методов поведенческой биометрической аутентификации, на основе нажатия клавиш и мышь

1.7.1 Обзор подходов известных методов

– Модель смеси гауссовых распределений

Это вероятностная модель, которая предполагает, что все точки данных генерируются из смеси конечного числа гауссовых распределений с неизвестными параметрами. В работе указано [32], что распределение значений динамических данных о нажатии клавиш, извлеченных для всех пользователей, является нормальным, но значение распределения варьируется для каждого пользователя. Для каждого протестированного набора данных его распределение сравнивается с нормальными компонентами распространения данных действительного пользователя. На рисунке 1.7 показаны гауссовы смеси и ряд точек, разбросанных в широком пространстве, где эти точки поляризованы в соответствии с каждой смесью, к которой они принадлежат.

Плотности вероятности гауссовского распределения вычисляется по формуле:

$$G(X|\mu,\sigma)^n = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (1.1)$$

где X–представляет собой точки данных; σ – дисперсия случайной

величины; μ – математическое ожидание случайной величины.

Иногда логарифм используется в модели для нахождения оптимальных значений параметров и вычисляется по формуле:

$$\ln N\,(x|\mu,\Sigma\,)^n = -\frac{D}{2}\,\ln 2\pi - \frac{1}{2}\ln\Sigma - \frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu), \qquad (1.2)$$

где D – количество измерений каждой точки данных; μ и Σ — среднее значение и ковариация соответственно.

Однако, поскольку эта модель гауссова имеет дело не с одним гессианам, а со многими гессианами, все станет немного сложнее, когда придет время находить параметры для всей смеси.

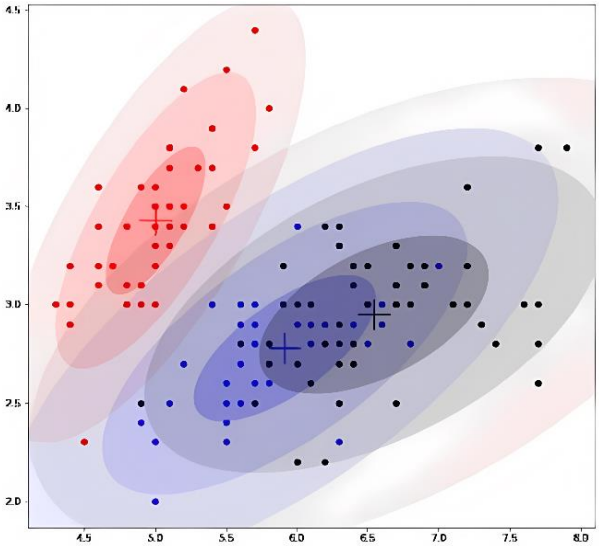


Рис. 1.7. Данные модели гауссовой смеси для кластеров

Стоит отметить, что чем больше смесей используется в модели, тем больше это может повлиять на результат алгоритма, например, точки в группе, которые ей не принадлежат, будут приняты в эту группу.

– Наивный гауссовский байесовский метод — это тип наивного байесовского метода, в котором рассматриваются непрерывные атрибуты, а характеристики данных следуют гауссовскому распределению по всему набору данных [33].

Наивная байесовская классификация основана на байесовской теории вероятностей и вычисляется по формуле:

$$P(\mathrm{c}|x) = \frac{\mathrm{P}(x|c)P(c)}{P(x)}, \tag{1.3}$$

где $P(\mathrm{c}\,|x)$– апостериорная вероятность; $\mathrm{P}(x|c)$ – вероятность; $\mathrm{P}(\mathrm{c})$ – класс априорная вероятность; $\mathrm{P}(\mathrm{x})$ – предиктор априорная вероятность.

Алгоритм классификации обладает прекрасной способностью предсказывать правильный класс, к которому относятся извлеченные динамические признаки и вычисляется по формуле:

$$P(x_i|\,y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x_i - \mu_y)^2}{2\sigma_y^2}\right), \tag{1.4}$$

где x–представляет собой точки данных; σ – дисперсия случайной величины; μ – математическое ожидание случайной величины.

На рисунке 1.8 показаны образцы классифицированные по байесовской, и набор точек, распределенных по большой площади. Эти точки считаются динамическими данными, извлеченными как для нажатий клавиш, так и для щелчков мыши. Эти поляризованные точки принадлежат определенной группе и разделены границами, чтобы различать эти данные, чтобы они были независимы от других.

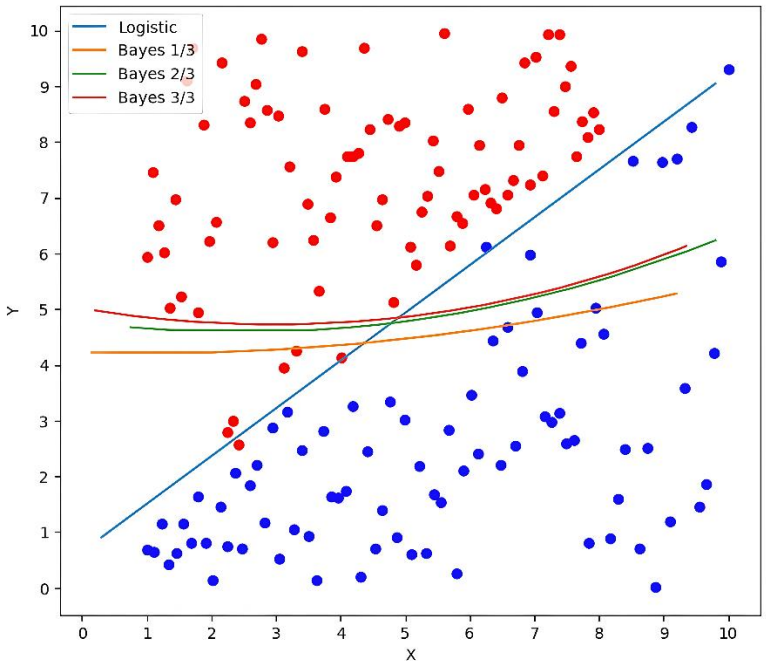


Рис. 1.8. Классификация гауссовский Наивный Байес

Если признак отсутствует в обучающих данных для определенного класса, расчет вероятности может стать нулевым.

Производительность может ухудшиться, если фактическое распределение признаков значительно отклоняется от предполагаемого распределения.

Несмотря на упрощающие предположения, наивный байесовский алгоритм может быть мощным и эффективным алгоритмом для определенных типов задач классификации, особенно при работе с данными высокой размерности и относительно небольшими наборами данных.

– decision tree classifier

Дерево решений — это структура, похожая на блок-схему, используемая для принятия решений или прогнозов. Оно состоит из узлов, представляющих решения или тесты по атрибутам, ветвей, представляющих результаты этих решений, и конечных узлов, представляющих окончательные результаты или прогнозы. Каждый внутренний узел соответствует тесту по атрибуту, каждая ветвь соответствует результату теста, а каждый конечный узел соответствует метке класса или непрерывному значению [34, 35].

Одним из методов, используемых для классификации дерева решений, является: ID3 Матриц. Она вычисляется по формуле:

$$H(T \mid A) = \sum_{a} p(a) \sum_{i=1}^{J} -\Pr(i \mid a)\, log_2\, Pr(i \mid a)\,, \qquad (1.5)$$

где $H(T \mid A)$ – сумма энтропий (детей); S– текущий набор данных, для которого вычисляется энтропия; a – Множество классов в T; p(a) – Отношение числа элементов в классе (a) к числу элементов в наборе T.

Дерево решений содержит много слоев, что делает его сложным. Обучающиеся на основе дерева решений могут создавать слишком сложные системы, которые не обобщают данные должным образом. Это называется переобучением. Для большего количества меток классов вычислительная сложность дерева решений может возрасти.

– KNN: k-Nearest Neighbor — это непараметрическая модель, которая

использует функцию расстояния для оценки метки новой контрольной точки. Она включает в себя взятие среднего значения прогнозов k ближайших точек к заданной контрольной точке. Она часто служит базовой моделью для многих задач прогнозирования и часто ее трудно превзойти.

Поскольку KNN является ленивым алгоритмом, он занимает больше памяти и данных, чем другие классификаторы, и это может быть дорогостоящим с точки зрения времени и денег. Он также не очень хорошо работает с вводом многомерных данных из-за «проклятия размерности» KNN также более склонен к переобучению. Хотя методы отбора признаков и снижения размерности используются для предотвращения этого, значение k также может влиять на поведение модели. Более низкие значения k могут переобучить данные, тогда как более высокие значения k имеют тенденцию «сглаживать» значения прогноза, поскольку они усредняют значения по большей области или соседству. Однако, если значение k слишком велико, оно может не подготовить данные [36].

– Метод опорных векторов (SVM)

SVM — это мощный контролируемый алгоритм, который лучше всего работает с небольшими наборами данных, но на сложных часто реализуется с помощью модели SVM. Метод опорных векторов, сокращенно SVM, можно использовать как для задач регрессии, так и для задач классификации [37]. На рисунке 1.9 показано этап классификации и исследования во время извлечения данных биометрической динамики, так что каждый образец данных отделен от другого образца, а каждый образец нажатия клавиш рассматривается как одна нажатая буква или одно движение мыши, где выборки данных разделены сплошным или гибким полем. Сплошное поле не позволяет выполнять выборку данных вместе, но гибкое поле позволяет данным различаться между ними, что называется пороговым значением [38,39].

В качестве алгоритма классификации можно использовать линейный пороговый классификатор:

$$a(\overrightarrow{x}) = sign\big((\vec{w}, \vec{x}) - b\big) = sign\left(\sum_{i=1}^{l} w_i x_i - b\right), \quad (1.6)$$

где $\vec{x}$ — вектор значений признаков объекта; $\vec{w} \in R^n$ и b ∈ R— параметры гиперплоскости.

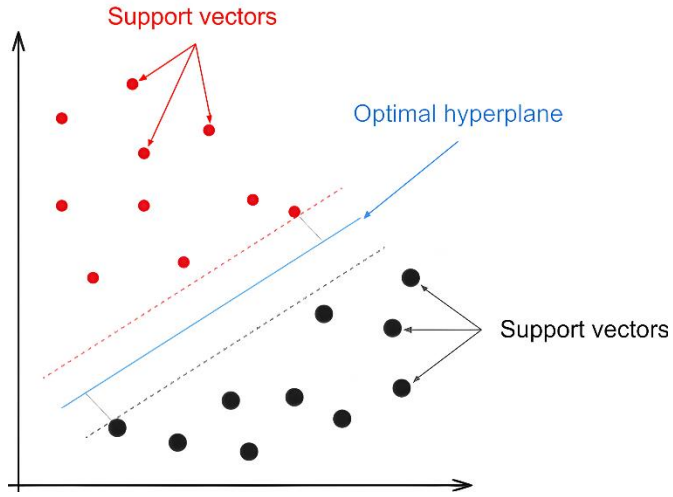


Рис. 1.9. Механизм метода опорных векторов (SVM)

Недостатки метода опорных векторов (SVM): SVM может быть медленным для больших наборов данных, что влияет на производительность SVM в задачах интеллектуального анализа данных; Выбор правильного ядра и параметров, таких как C, требуют тщательной настройки, что влияет на алгоритмы SVM; SVM испытывает трудности с зашумленными наборами данных и перекрывающимися классами, что ограничивает эффективность в реальных сценариях; Сложность гиперплоскости в более высоких измерениях делает SVM менее интерпретируемой, чем другие модели; Правильное масштабирование признаков имеет важное значение, в противном случае модели SVM могут работать плохо [40, 41].

– Нейронные сети

Нейронная сеть — это вычислительная модель, вдохновленная нейронной системой человеческого мозга. Она считается важной частью области искусственного интеллекта и призвана представлять и моделировать процесс обучения и обработки информации аналогично человеческому мозгу [42,43]. Нейронная сеть состоит из группы элементов, известных как узлы (нейроны),

которые соединены друг с другом связями (синапсами). Нейронная сеть состоит из “input layer” для получения информации, нескольких “hidden layers”, которые обрабатывают данные и наконец, “output layer”, который выдает окончательные результаты как показное на рисунке 1.10.

Нейронная сеть работает путем преобразования входящей информации в узлы первого слоя, а затем определенными способами передает эту информацию через узлы внутренних слоев. Сила связей между узлами регулируется на основе опыта и обучения на основе доступных данных, что известно как обучение нейронной сети, при котором нейронная сеть постепенно адаптируется для выполнения конкретной задачи путем корректировки весов и значений, связанных с узлами [32].

Существует множество нейронных сетей, которые использовались в области поведенческих биометрических измерений щелчков клавиш и мыши с целью классификации, включая то, что было упомянуто: Recurrent Neural Networks (RNNs); convolutional neural network (cnn); Feedforward Multi-Layer Perceptron (FF-MLP);

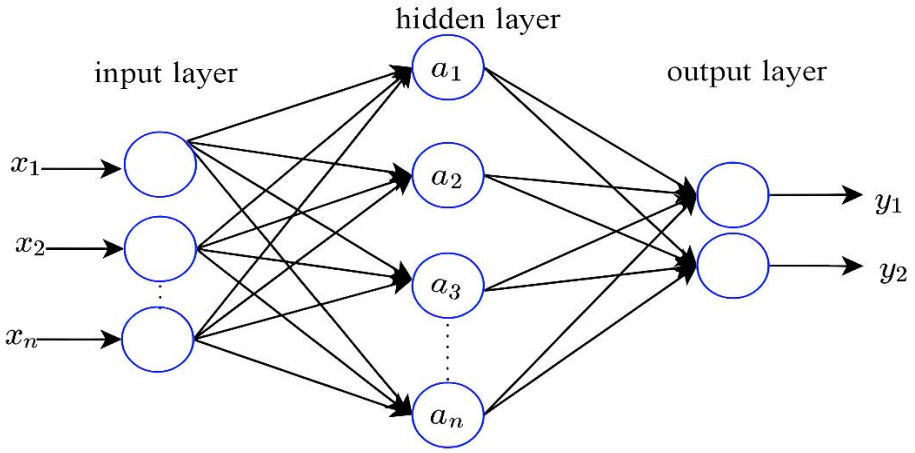


Рис. 1.10. Слои нейронной сети

1.7.2 Сравнение эффективности работы используемых подходов

Биометрические системы аутентификации предлагают сочетание передовых технологий и удобства для пользователя. Для измерения эффективности и результативности поведенческих биометрических систем,

основанных на динамике нажатий клавиш и движений мыши, требуется тонкий баланс между безопасностью и доступностью [44]. В существующих работах используются следующие метрики:

- Точность классификации (Accuracy) — это доля правильных классификаций либо для независимого набора тестов, либо с использованием некоторого варианта идеи перекрестной проверки.
- FAR (False Acceptance Rate) — это статистическая метрика, которая измеряет вероятность того, что система биометрической аутентификации или механизм контроля доступа неправильно предоставит доступ неавторизованному лицу. Проще говоря, она количественно определяет скорость, с которой система ложно принимает самозванцев за законных пользователей (величина ошибки второго рода).
- FRR (False Rejection Rate) — доля случаев, когда зарегистрированный в системе действительный пользователь отклоняется и считается злоумышленником (величина ошибки первого рода);
- EER (Equal Error Rate) – это коэффициент, при котором FAR И FRR равны, чем ниже ERR, тем точнее биометрическая система.
- AUC (Area Under Curve) – количественная интерпретация ROC-кривой, обозначающая площадь под ней.
- ROC-кривая (Receiver Operating Characteristic) – показывает истинно положительный показатель (TPR- Процент принятых реальных пользователей, sensitivity) в сравнении с ложноположительным показателем (FPR- Процент злоумышленников, которые были правильно отклонены, specificity) для различных пороговых значений баллов классификации.

Достигнутые в рассмотренных выше научных публикациях результаты в области статической и непрерывной поведенческо-биометрической аутентификации пользователей на основе динамика нажатия клавиш и мыши представлены в таблице 1.

Таблица 1 — Сравнение эффективности работы используемых подходов

Работа	**метод**	**Поведенческая биометрия**	**Тип аутентификации**	**Эффективность аутентификации**
[42]	RNN нейронная сеть	Динамика нажатия клавиш	Статическая	ERR =0.63 Accuracy =91.23%
[36]	One-Class KNN	Динамика нажатия клавиш	Непрерывная	ERR=0.44
[37]	One-Class SVM	нажатие клавиши акустическое	Статическая	Accuracy =92.8% FAR=11% FRR=12%
[38]	One-Class SVM	Динамика нажатия мыши	Непрерывная	ERR=1.3
[34]	decision tree classifier	Динамика нажатия мыши	Непрерывная	FRR=0.43% FAR=1.75% ERR=1.09
[35]	decision tree classifier	Динамика нажатия клавиш и мыши	Непрерывная	Accuracy = 62.2%
[43]	CNN Convolutional Neural Network	Динамика нажатия клавиш	Статическая	Accuracy = 82%
[39]	SVM	Динамика нажатия мыши	Статическая	FRR=5.5% FAR=8.8%
[44]	one-class classification	Динамика нажатия мыши	Непрерывная	FAR=0.37% FRR=1.12%
[40]	One-class SVM	Динамика нажатия мыши	Непрерывная	FAR=10 FRR=16
[41]	One-class SVM	Динамика нажатия мыши	Статическая	FAR=8.74% FRR=7.69%
[33]	Gaussian Naive Bayes classification	Динамика нажатия мыши	Статическая	Accuracy =93.9% FAR= 0.062 FRR=0.042
[46]	GMM	Динамика нажатия клавиш	Непрерывная	EER= 0.09
[32]	FF-MLP	Динамика	Непрерывная	FAR = 0.02 FRR

	нейронная сеть	нажатия клавиш		= 4.82
[47]	Manhattan	Динамика нажатия клавиш	Статическая	EER= 9.6

Выводы по 1-й главе

Из обзора рынка поведенческой биометрии и современных методов анализа динамики щелчков клавиш и мыши пользователя с целью проведения процесса аутентификации, а также обзора современных компаний в этой области. Сделаны следующие выводы:

- Использование поведенческой биометрии в области аутентификации пользователей на основе анализа биометрических данных динамики нажатий клавиш и мыши является перспективным направлением исследований и широко применяется для обеспечения безопасности систем, а также для повышения степени защищенности систем предотвращение несанкционированного доступа хакеров;
- Современные методы не в полном объёме способны решить проблему изменения баллистического характера пользователей путем изменения со временем стиля письма или движения мыши;
- В ходе исследований в области поведенческой биометрии стало ясно, что не существует стабильной и непрерывной интегрированной системы аутентификации, основанной на аутентификации пользователя в процессе входа в систему и мониторинге пользователя во время использования системы до завершения сеанса;
- Целью процесса аутентификации является повышение степени безопасности и предотвращение несанкционированного доступа. В ходе исследований в области и методах аутентификации было обнаружено, что многофакторная аутентификация имеет возможность повысить степень безопасности и проверить подлинность пользователя. Текущие работы не используют многофакторную аутентификацию в поведенческих

биометрических измерениях;

- Чтобы провести процесс аутентификации для поведенческих биометрических измерений, необходимо извлечь динамические характеристики нажатия клавиш и мыши, такие как нажатие или отпускание и тип события, а также извлечь отметку времени для каждого произошедшего события. После фазы обучения, этап проверки пользователя и аутентификации завершен;
- Обзор Российского законодательства в области систем биометрической идентификации, где Статья 11 Федерального закона предусмотрено, что физиологические, биологические и поведенческие данные человека используются оператором для идентификации субъекта персональных данных;
- Анализ уязвимостей безопасности, с которыми сталкиваются системы аутентификации веб-приложений, показывает, что убытки в Соединённых Штатах Америки увеличиваются с каждым годом, а потери слабых систем аутентификации в 2023 году достигли 12 с половиной миллиардов долларов;
- Обзор стандартов и требований к поведенческой биометрической аутентификации. Были проанализированы семь основных стандартов: Универсальность; Уникальность; Постоянство; Измеримость; Производительность; Приемлемость; Обход.

Для дальнейшего совершенствования методов исследования аутентификация пользователя на основе поведенческих биометрических измерений нажатий клавиш и мыши следует признать целесообразным решение следующих задач:

1. Создание модели, извлекающей все биометрические характеристики нажатий клавиш и движений мыши и разработка модели идентификации руки на основе динамики нажатия клавиш;
2. Создание модели для генерации случайного одноразового пароля на основе динамики нажатий клавиш и создание трехфакторной технологии

аутентификации пользователей и субъектов доступа для веб-приложения;

3. Создание непрерывной системы аутентификации на основе нейронной сети и динамики мыши при использовании веб-приложений.

В главе 1 сформулирована цель диссертационного исследования. Определены задачи, которые необходимо решить для достижения поставленных целей. Решение задач позволит эффективность системы аутентификации web-прилежаний, основанной на поведенческой биометрии нажатия клавиш и мыши, обеспечивая очень низкий уровень ложного отклонения и ложного принятия и, следовательно, высокую степень безопасности.

ГЛАВА 2. Биометрическая модель аутентификации пользователя на основе динамики нажатия клавиш и мыши.

2.1 Разработка модели аутентификации на основе поведенческой биометрические

Поведенческая биометрия анализирует то, как пользователь взаимодействует с устройствами, создавая уникальную для пользователя модель поведения.

Исследователь и консультант Дэвид Смит [48] объяснил в журнале Info Security, что область аутентификации на основе поведенческой биометрии приобрела большую и растущую популярность, поскольку обеспечивает механизм пассивной проверки личности людей без их ведома.

Было показано, что существует еще один фактор, доказавший строгость и надежность поведенческих биометрических измерений в области аутентификации, а именно фактор сбора динамических биометрических данных. Другие виды аутентификации, такие как пароль, отпечаток пальца или лицо, содержат фиксированные биометрические данные, хранящиеся в базе данных, и могут быть украдены хакерами в любой момент. Однако динамические биометрические данные постоянно меняются с изменениями баллистической природы человека, в связи с чем, украденные поведенческие биометрические данные бесполезны [48].

Ширли Инскоу, старший аналитик Aite Group [49], считает, что поведенческая биометрия желательна для клиента, поскольку современные методы аутентификации предполагают прямой контакт с клиентом. Но поведенческая биометрия аутентифицирует пользователя, не оказывая негативного влияния и не беспокоя его, и, таким образом, отсутствие трений между системой и потребителем является важным и эффективным элементом процесса аутентификации.

Существуют дополнительные причины, по которым все государственные и коммерческие и, в особенности, банковские учреждения [48, 50, 51] рассматривают поведенческую биометрию для целей аутентификации:

– Сокращение администрирования. При развертывании в онлайн-канале более плавный процесс аутентификации снижает административную

нагрузку, связанную с доступом и обслуживанием базы пользователей.

- Сокращение мошенничества. Поведенческая биометрия играет активную роль в снижении риска мошенничества, а также демонстрирует постоянную приверженность сокращению мошенничества и соблюдению нормативных требований.
- Меньше ложных срабатываний. Поведенческая биометрия также снижает ложные срабатывания и соответствующее воздействие на клиентов и сотрудников бэк-офиса финансовых учреждений.
- Экономия средств. Поведенческая биометрия также имеет преимущества в плане затрат, поскольку не требует дополнительного развертывания оборудования.
- Повышение удовлетворенности клиентов. Благодаря ненавязчивости этого подхода банковские учреждения могут ожидать снижения оттока клиентов.
- Сокращение проблем с конфиденциальностью. Меньше проблем с конфиденциальностью по сравнению с физической биометрией, такой как сканирование отпечатков пальцев или радужной оболочки глаза. Вместо этого поведенческие данные преобразуют поведение пользователя в математическое представление в его профиле, что бессмысленно для потенциальных хакеров.

В результате первого исследования были разработаны три метода поведенческих биометрических измерений: динамика нажатия клавиш, динамика мыши и мягкие биометрические измерения для определения типа руки на клавиатуре. Эти методы взаимосвязаны друг с другом: пользователь не сможет добиться успеха в аутентификации во втором методе, если он не добьется успеха в первом. Аналогично, пользователь не сможет пройти через третий метод, если он не достигнет успеха в аутентификации в первом и втором методах соответственно, как показано на рисунке 2.1.

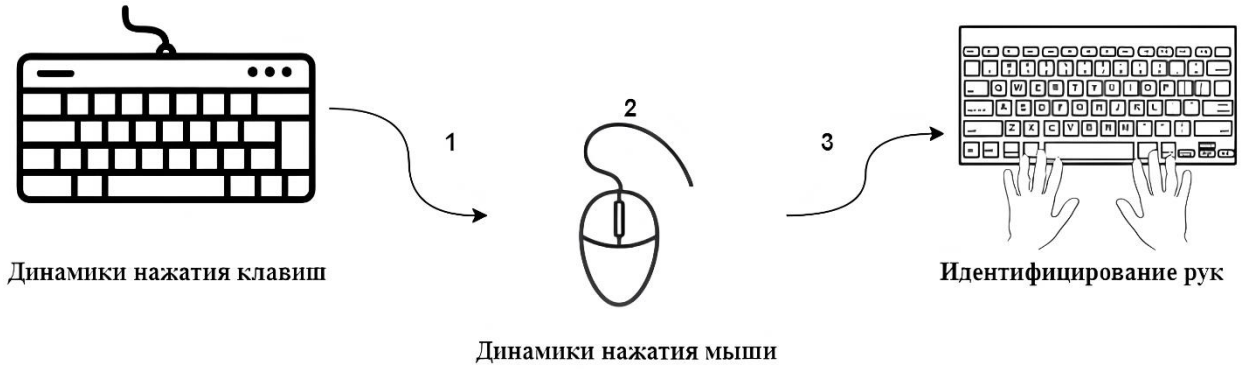


Рис. 2.1. Разработка метода поведенческой биометрии

В своих расчетах и в принятии решений поведенческие биометрические измерения зависят от трех основных концепций, которые должны присутствовать и рассчитываться в каждом используемом биометрическом измерении:

- Извлечение характеристик биометрических измерений;
- Расчет порогового значения: решение об обновлении шаблона принимается на основе расстояния между шаблоном и образцом тестовых данных, а также порогового значения, которое отличается от порогового значения, используемого для аутентификации;
- Создание математического алгоритма, доказывающего, что информация о человеке, использующем клавиатуру и мышь, совпадает с информацией о действительном пользователе, хранящейся в базе данных, на основе чего принимается решение о его аутентификации или отклонении.

2.2 Биометрическая модель динамики нажатия клавиш

2.2.1 Описание основных принципов работы динамического нажатия клавиш

Во время Второй мировой войны (Army Signal Corps) обнаружило и определило, что ритм телеграфной манипуляции каждого человека был разным [52]. Национальный научный фонд и Национальное бюро стандартов США провели исследования в начале 1980-х годов, которые доказали, что модели типизации имеют определенные свойства, которые можно определить [52].

Спиллейн первым предложил использовать клавиатуру для измерения динамики нажатия клавиш для идентификации личности [53]. Джэинс, 1980 был первым, кто продемонстрировал использование динамики нажатия клавиш для аутентификации [54, 55, 52].

В последнее время биометрические системы с динамическим нажатием клавиш стали альтернативой системе имени пользователя и пароля, которая имеет ряд недостатков: пароли могут быть забыты, атакованы или переданы другим, и, таким образом, система подвергается риску. Стиль письма пользователя может быть уникальным из-за схожих нейрофизиологических факторов, которые делают письменные подписи уникальными.

В отличие от других биометрических систем, которые обычно требуют дополнительного оборудования и поэтому дороги в реализации, биометрия, основанная на динамике нажатия клавиш, практически бесплатна, т.е. единственное необходимое оборудование — это клавиатура [56, 53].

Динамика нажатия клавиш — это небольшие закономерности и изменения в движениях рук и пальцев на клавиатуре, которые происходят естественным образом по мере набора пользователем текста и представляют собой не интрузивный биометрический показатель, широко используемый для аутентификации [57].

Нажатия клавиш столь же уникальны и неповторимы, как и подписи. Это означает, что всегда будут огромные различия в стиле письма – скорость, ритм, использование заглавных букв (левая или правая клавиша Shift), плавность или неуверенность при печати, орфографические ошибки и последующие исправления и т. д. [57]

Динамику нажатия клавиш можно разделить на две подобласти [58], статическую и непрерывную динамику нажатия клавиш.

1- Статическая динамика нажатия клавиш основана на предопределенной задаче в управляемой среде. Часто пользователю поручают написать фразу или предложение, повторяя их несколько раз, пока его поведение записывается [59]. Затем это поведение используется для создания шаблона для идентификации

пользователя в будущем [60], что выполняется путем повторного написания пользователем той же фразы.

2- Непрерывная динамика нажатия клавиш, с другой стороны, не имеет предопределенной задачи и может выполняться с руководством пользователя или без него [60]. Часто это выполняется путем взаимодействия пользователя с некоторой системой, записывающей его поведение. В отличие от статической динамики нажатия клавиш, это обычно требует больше данных для идентификации пользователя. Практические приложения непрерывной динамики нажатия клавиш часто сосредоточены на обнаружении аномалий, а не на идентификации.

2.2.2. Анализ структуры и характеристики динамики нажатия клавиш

Характеристики динамики нажатия клавиш обычно извлекаются с использованием информации о времени событий нажатия/удержания/нажатия клавиши. Обычно используются время удержания и задержки между двумя клавишами, т. е. временной интервал между отпусканием клавиши и нажатием следующей клавиши.

Существует несколько методик для выделения признаков, характеризующих клавиатурный почерк пользователя [61,56,62]:

- Анализ одиночных нажатий на клавиши клавиатуры;
- Анализ последовательных нажатий на клавиши клавиатуры (то есть анализ нажатий пользователем комбинаций нескольких клавиш – диграфов, триграфов, N-грамм);
- Комбинированный анализ динамики работы пользователя как с одиночными клавишами клавиатуры, так и с их комбинациями.

Третий подход считается наиболее перспективным, поскольку он позволяет учесть больше индивидуальных особенностей пользователя при работе с клавиатурой, анализируя как его работу с отдельными клавишами, так и с их комбинациями.

В подходе, используемом для анализа динамики нажатия клавиш, как показано в таблице 2, было выделено несколько характеристик нажатия клавиш,

как показано на рисунке 2.2, с помощью сочетаний клавиш на рисунке 2.3, с помощью которых идентифицируются стиль письма и движения человека на клавиатуре:

Последовательность клавиш для слова с фиксированной длиной n. Определение: Key_i - последовательность клавши (k), состоящая из n ≥ 2, нажатий клавиш $K = k_1, k_2, \dots, k_n$ где k_i — кортеж в форме k_i = ($Keydown_i$, $Keyup_i$) и 1 ≤ i ≤ n, i ∈ N.

Таблица 2 — Характеристики динамического нажатия клавиш

Действия клавиатуры	**Описание**	**Формула**
Действие одной клавиши	задержка между нажатием и отпусканием клавиши	$DU = Keydown_i - Keyup_i$
Интервал	время, прошедшее между отпусканием и нажатием следующей клавиши	$UD = Keyup_i - Keydown_{i+1}$
Действие одной клавиши и ее удержание	время, которое проходит между нажатием первой клавиши, ее отпусканием и последующим нажатием следующей клавиши	$DUD = Keydown_{i.1} - Keyup_{i.1} - Keydown_{i+1}$
Действие ключа диграфа	время, которое проходит между двумя клавишами при нажатии и последующем отпускании, затем нажатии следующей клавиши и последующем отпускании.	$DUDU = Keydown_{i.1} - Keyup_{i.1} - Keydown_{i+1} - Keyup_{i+1}$
Действие ключа триграфа	это время, прошедшее между нажатием и отпусканием трех клавши	$DUDUD = Keydown_{i.1} - Keyup_{i.1} - Keydown_{i+1} - Keyup_{i+1} - Keydown_{i+2}$

процесс называется N-грамм	это время, затрачиваемое на нажатие и отпускание четырех или более клавиш, в зависимости от длины введенного пароля	$DUDUD = Keydown_{i.1} - Keyup_{i.1} - Keydown_{i+1} - Keyup_{i+1} - Keydown_{i+2} - Keyup_{i+2} - Keydown_{i+n} - Keyup_{i+n}$
Время выдержки	Это время, прошедшее при двойном нажатии одной и той же клавиши	$DC = (Keydown_i - Keyup_i) + (Keydown_i - Keyup_i)$;
N-грамм отпускания	время, которое проходит между отпусканием четырех или более клавиш	$UUU = Keyup_{i.1} - Keyup_{i+1} - Keyup_{i+2} - Keyup_{i+n}$;
Триграф отпускания	время, прошедшее с момента отпускания двух клавиш и отпускания следующей клавиши	$UUU = Keyup_{i.1} - Keyup_{i+1} - Keyup_{i+2}$;
Диграф отпускания	время, которое проходит между отпусканием одной клавиши и последующим отпусканием следующей клавиши	$UU = Keyup_{i.1} - Keyup_{i+1}$
Диграф нажатия и отпускания	сочетание анализа нажатия клавиш с диграфом, представляет собой время, прошедшее между нажатием и отпусканием двух клавиш и нажатием следующей клавиши	$DUDUD = Keydown_{i.1} - Keyup_{i.1} - Keydown_{i+1} - Keyup_{i+1} - Keydown_{i+2}$
Shift + Alt	Действие кода диграфа= 16 и 18	$F = T_{key_{16}} + T_{key_{18}} - T_{key_{i-1}}$

Ctrl + Z	Действие кода диграфа= 17 и 90	$F = T_{key_{17}} + T_{key_{90}} - T_{key_{i-1}}$
Backspace	Действие одной клавиши=8	$F = T_{key_{8}} - T_{key_{i-1}}$

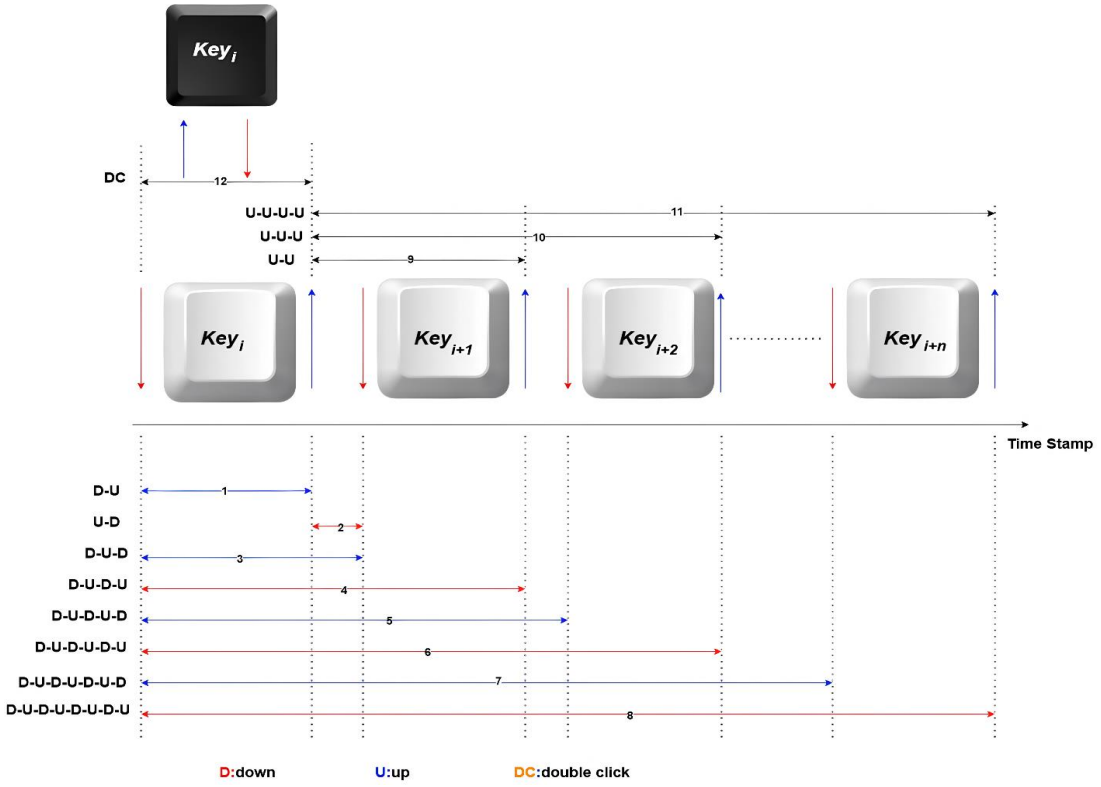


Рис. 2.2. Структуры и характеристики динамики нажатия клавиш

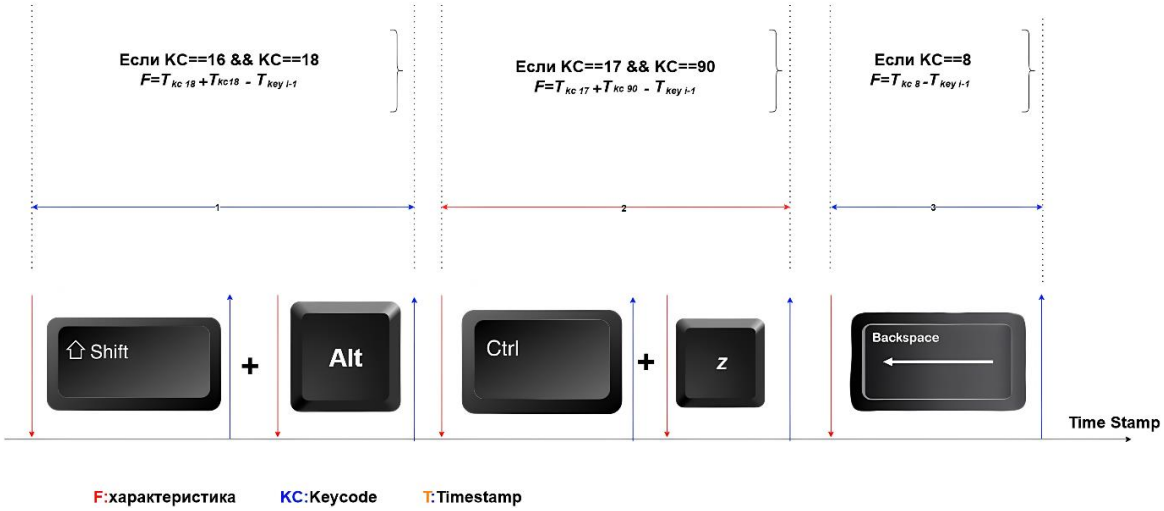


Рис. 2.3. Характеристики нажатия клавиш с помощью сочетаний клавиш

1.2.3 Разработка модели динамики нажатия клавиш для аутентификации.

Механизм аутентификации, основанный на поведенческой биометрии и нажатиях клавиш, был разработан на основе подхода трех-расстояния между данными и Теоремой Пифагора:

1. Евклидово расстояние:

Евклидово расстояние в евклидовом пространстве - длина отрезка прямой линии, соединяющей две точки, результатом расчета является кратчайшее расстояние между двумя точками в размерном пространстве. Евклидово пространство — это двух или трехмерное пространство, в котором применяются аксиомы и постулаты евклидовой геометрии [189, 190].

Эта метрика основана на теореме Пифагора и широко используется в различных областях, таких как машинное обучение, анализ данных, компьютерное зрение.

Для математической области евклидово расстояние рассчитывается для одного измерения, двух измерений и более высоких измерений.

$$d(p,q) = \sqrt{(p-q)^2}, \tag{2.1}$$

$$d(p,q) = \sqrt{(p_1-q_1)^2 + (p_2-q_2)^2}, \tag{2.2}$$

$$d(p,q) = \sqrt{(p_1-q_1)^2 + (p_2-q_2)^2 + \cdots + (p_n-q_n)^2} = \sqrt[2]{\sum_{i}^{n} (p_i - q_i)^2}, \tag{2.3}$$

Для большей точности евклидово расстояние рассчитывается с использованием расстояния стандартного отклонения для каждого объекта, которое определяется следующим образом:

$$d(p,q) = \sqrt[2]{\sum_{i}^{n} (p_i - q_i \,/\, \alpha_i)^2}, \tag{2.4}$$

Всегда положительное и симметричное расстояние означает, что расстояние от точки X до точки Y такое же, как и от точки Y до X.

Но в области динамики нажатия клавиш стиль письма человека принципиально отличается от стиля письма другого человека, так как при

написании пароля человек может следовать в соответствии со своим стилем определенным путем, чтобы дойти до букв на кратчайшем расстоянии, а другой человек может пройти более длинный путь, чтобы добраться до финальной точки. Например, возьмем пароль (you), как показано на рисунке 2.4 Мы проиллюстрируем два стиля письма, которым следуют два разных человека.

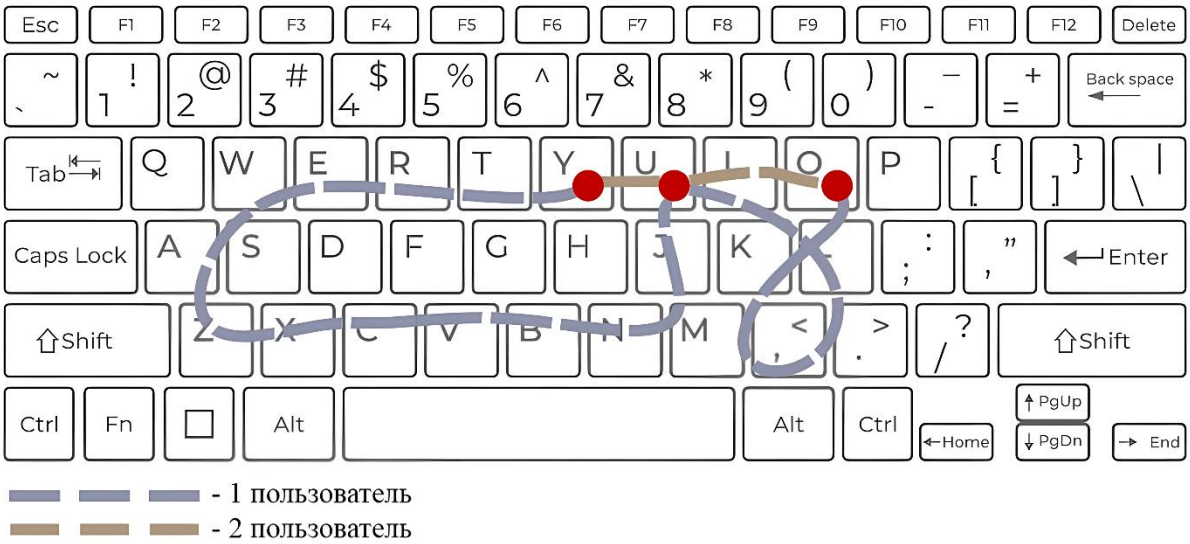


Рис. 2.4. Стиль набора текста на клавиатуре

Следовательно, исходя из этого принципа, один и тот же человек может со временем написать свой пароль по-другому. Он может пройти расстояние, большее, чем-то расстояние, которое было при обучении. Поэтому вычисление порогового значения на основе евклидова расстояния может оказаться неэффективным, когда время пройдет. Поэтому евклидово расстояние было объединено с Манхэттенским расстоянием, чтобы найти лучшее пороговое значение для пользователя, поскольку Манхэттенское расстояние следует схеме извилистых дорог для достижения цели.

2. Манхэттенское расстояние

Манхэттенское расстояние — это метрика, используемая для определения расстояния между двумя точками на сетке. В отличие от евклидова расстояния, которое измеряет кратчайшую возможную линию между двумя точками, манхэттенское расстояние измеряет сумму абсолютных разностей между координатами точек [194]. Этот метод называется «манхэттенским расстоянием»,

потому что, подобно такси, проезжающему по сетке улиц Манхэттена, оно должно двигаться вдоль линий сетки.

$$d_T(p,q) = |p - q|, \tag{2.5}$$

$$d(p,q) = \big||p - q|\big|_T = |p_1 - q_1| + |p_2 - q_2|, \tag{2.6}$$

$$d_T(p,q) = \sum_{i=1}^{n} |p_i - c_i|\,, \tag{2.7}$$

Кроме того, для большей точности Манхэттенское расстояние рассчитывается с использованием расстояния стандартного отклонения для каждого объекта, которое определяется следующим образом:

$$d_T(p,q) = \sum_{i}^{n} |(p_i - q_i)|\,/\alpha_i\,, \tag{2.8}$$

Манхэттенское расстояние и евклидово расстояние — это меры, используемые для расчета расстояния между двумя или более точками. Основное различие между ними заключается в том, как измеряется это расстояние. Манхэттенское расстояние вычисляет сумму абсолютных разностей координат, а евклидово расстояние вычисляет квадратный корень из суммы квадратов разностей координат. В пространстве, напоминающем сетку, манхэттенское расстояние больше подходит для измерения расстояний вдоль линий сетки, а евклидово расстояние больше подходит для измерения расстояния по прямой линии.

Как показано на рисунке 2.5, евклидово расстояние, обозначающее синий цвет, проходит кратчайшее расстояние для достижения цели, так что его значение (1) известно, но манхэттенское расстояние, обозначающее красный цвет, требует более длинного и извилистого пути к цели, так что части каждой стороны (1,2) манхэттенского расстояния не известны, а вычисляется общая сумма. Следовательно, чтобы узнать значение сторон манхэттенского расстояния, используется расстояние Чебышева.

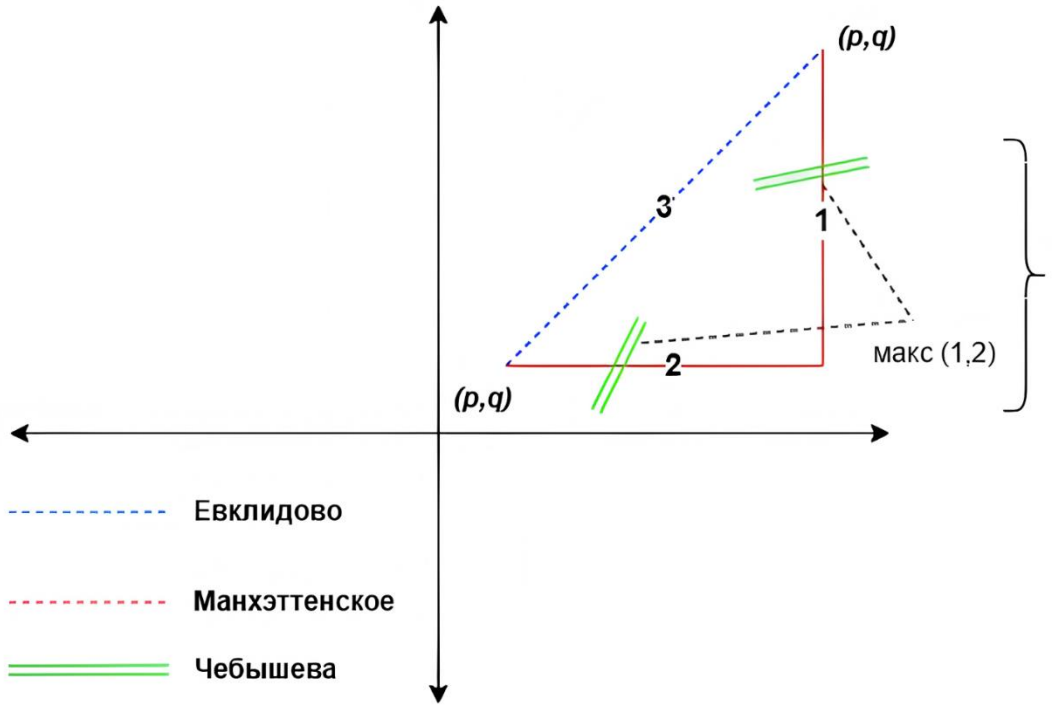


Рис. 2.5. Путь по трем траекториям движения

3. Расстояние Чебышева

Расстояние Чебышева, названное в честь русского математика Пафнутия Чебышева, определяется, как максимальная разность между координатами двух точек вдоль любой одной оси.

Расстояние Чебышева подчеркивает максимальный сдвиг в любом направлении координат, что имеет решающее значение в сценариях, где движение не ограничивается горизонтальными или вертикальными траекториями, а включает любую прямую линию [190-192].

$$d_{Chebyshev}(x,y) = lim_{p\to\infty}(\sum_i |x_i - y_i|^p)^{\frac{1}{p}} = \max|x_i - y_i|\,, \qquad (2.9)$$

Расстояние Чебышева используется для нахождения длины стороны Манхэттенского расстояния, как показано на рисунке 2.5.

замечено, что когда объединяется три пробела и пишем на клавиатуре любые две буквы, мы получаем прямоугольный треугольник, как показано на рисунке 2.5, но, когда мы пишем на клавиатуре три и более буквы, мы получаем группу из прямоугольных треугольников и группу равносторонних квадратов, как показано на рисунке 2.6.

2.2.4. Процесс построения модели пользователя на этапе обучения по результатам динамики нажатия клавиш.

Основная цель этапа обучения — найти уникальное пороговое значение для каждого пользователя в зависимости от его стиля письма на клавиатуре на основе особенностей, которые были извлечены от его нажатий клавиш. Поэтому пороговое значение является основным направлением процесса аутентификации, и это барьер, который пользователь не должен пересекать.

Поэтому строится и рассчитывается пороговое значение, подходящее всем пользователям, чтобы оно было гибким и в то же время строгим. Поэтому для ограничения порогового значения использовались Манхэттенское, Евклидово расстояния и расстояние Чебышева, чтобы найти лучшее значение.

Пороговое значение рассчитывается на основе используемого подхода путем расчета угла противолежащего катета прямоугольного треугольника (Тригонометрические соотношения), для каждой нажатой клавиши, а также расчета площадей всех квадратов параллелограмма [194], как показано на рисунке 2.6.

Стороны прямоугольного треугольника с помощью теоремы Пифагора вычисляются по формуле:

$$c^2 = a^2 + b^2 = bc^2 = ab^2 + ac^2, \tag{2.10}$$

где c – евклидово расстояние; a – расстояние Чебышева; b – манхэттенское расстояние.

В геометрии прямоугольный треугольник — это треугольник с одним из прямых углов, то есть две стороны прямоугольного треугольника образуют угол 90 градусов.

По теореме Пифагора угол вычисляется по формуле:

$$\sin\alpha = \frac{a}{c} = \frac{\sum_{i=1}^{n}|p_i - c_i|}{\sqrt[2]{\sum_{i}^{n}(p_i - q_i)^2}}, \tag{2.11}$$

$$a = c \,.\sin\alpha = \sqrt[2]{\sum_{\boldsymbol{i}}^{\boldsymbol{n}}(p_i - q_i)^2} \,.\, sin^{-1}, \tag{2.12}$$

где α – катет, противолежащий углу; a – противолежащий катет; c – Гипотенуза.

$$\cos\alpha = \frac{b}{c} = \frac{\max|x_i - y_i|}{\sqrt[2]{\sum_i^n (p_i - q_i)^2}}, \tag{2.13}$$

$$b = c\,.\cos\alpha = \sqrt[2]{\sum_{\boldsymbol{i}}^{\boldsymbol{n}} (p_i - q_i)^2}\,.cos^{-1}, \tag{2.14}$$

где α – катет, прилежащий углу; b – прилежащий катет.

$$tg\ \alpha = \frac{a}{b} = \frac{\sum_{i=1}^n |p_i - c_i|}{\max|x_i - y_i|}, \tag{2.15}$$

$$a = b\,.tg\ \alpha = ma\,x|x_i - y_i|\,.tg^{-1}, \tag{2.16}$$

где α – катет, прилежащий углу.

Квадрат — это правильный четырехугольник, стороны которого равны по длине и перпендикулярны, образуя четыре прямых угла. Квадрат можно образовать сложением двух прямоугольных равнобедренных треугольников по гипотенузе [194].

Свойства квадрата в области углов

1. Все четыре стороны квадрата имеют одинаковую длину, то есть они равны:

$$AB = BC = CD = AD, \tag{2.17}$$

2. Все четыре угла квадрата прямые:

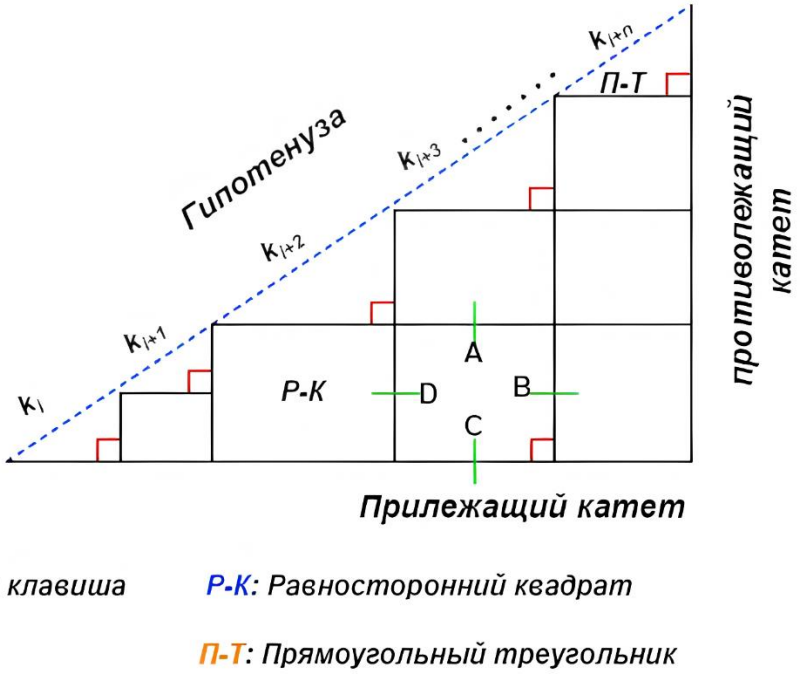
$$\angle ABC = \angle BCD = \angle CDA = \angle DAB = 90^{\circ}, \tag{2.18}$$

3. Сумма углов квадрата равна 360 градусов:

$$\angle ABC + \angle BCD + \angle CDA + DAB = 360^{\circ}, \tag{2.19}$$

$$threshould = \frac{\sqrt{\left(\sum_{i=1}^n \alpha_{right\ triangle} + \sum_{i=1}^n \frac{\beta_{square}}{4}\right)}}{\mu_n}, \tag{2.20}$$

Где $threshould$ – Пороговое значение; $\alpha_{right\ triangle}$ – Противоположный угол прямоугольного треугольника; β_{square} – Квадратный угол; μ –Количество нажатых и отпущенных клавиш на клавиатуре.



К: клавиша *Р-К: Равносторонний квадрат*

П-Т: Прямоугольный треугольник

Рис. 2.6 Геометрическое формирование подхода пользователя к нажатию клавиш

1.2.5 Рабочая среда этапа обучения

После успешной регистрации информации о пользователе в базе данных начинается этап обучения, как показное на рисунке 2.8:

- Пользователь вводит свой пароль и повторяет его три раза в отведенном текстовом поле как показное на рисунке 2.7.
- Создается биометрическая модель для каждого пользователя, извлекаются характеристики динамики нажатия клавиш и рассчитывается временная метка для каждой характеристики. Значения модели хранятся на сервере в файле с расширением CSV.
- Рассчитывается значение динамического порога для каждого пользователя путем нахождения углов прямоугольных треугольников и углов квадрата, образованного по временным меткам нажатия и отпускания клавиш на клавиатуре, с помощью трех типов расстояний между точками: евклидово, манхэттенские и Чебышева.

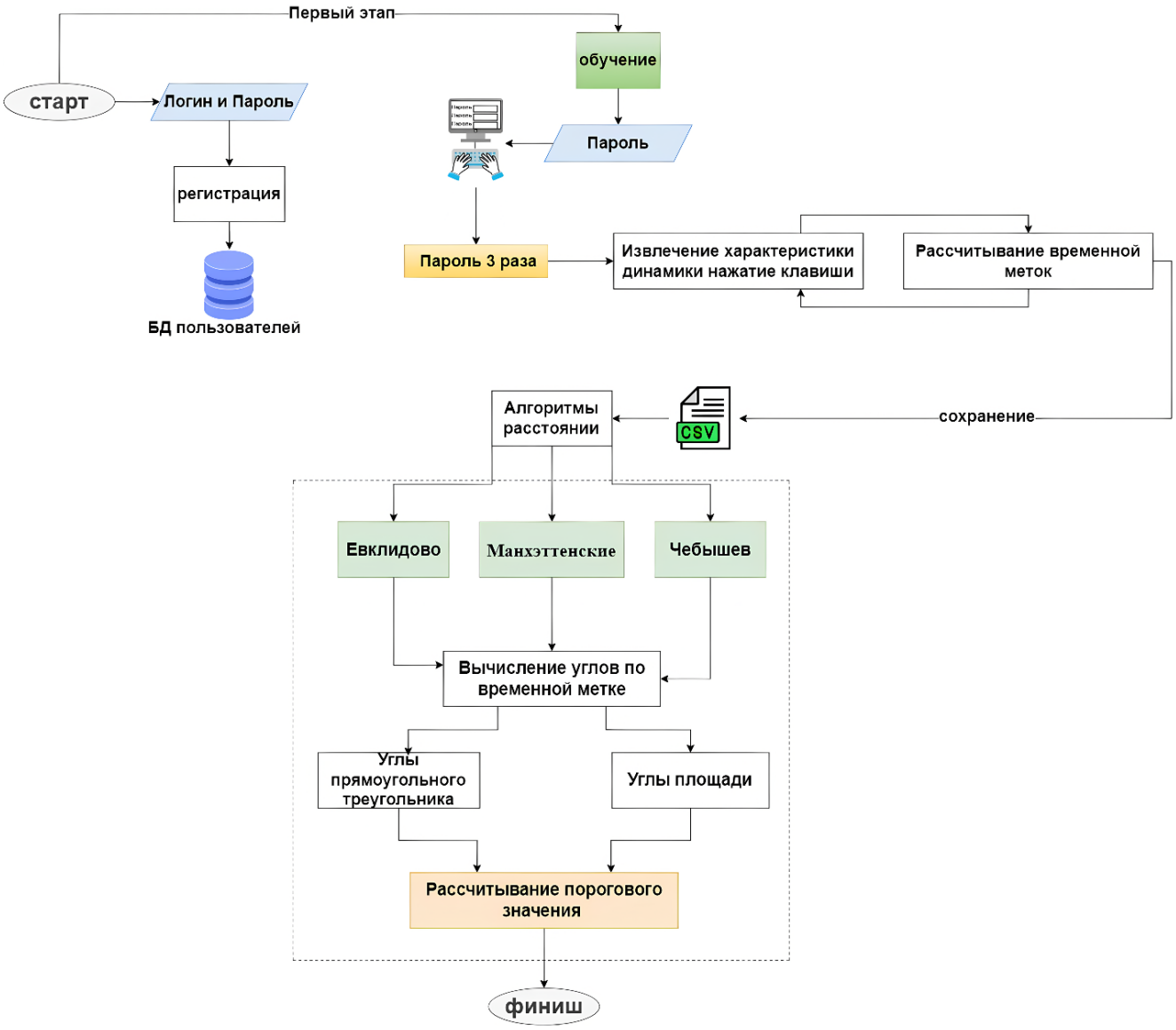


Рис. 2.7. Блок-схема этапа обучения динамики нажатия клавиш.

Рис. 2.8. Внедрение этапа обучения

2.2.6 Процесс подтверждения модели пользователя на этапе тестирования по результатам динамики нажатия клавиш.

Этап проверки и тестирования зависит от определения общей площади каждого из прямоугольных треугольников и квадратов, поскольку они представляют все клавиши, которые пользователь нажимал и отпускал, используя свой другой стиль движения на клавиатуре, как показано на рисунке 2.6**.**

Первая часть общей площади значений временных меток, рассчитанная с помощью функций динамики нажатия клавиш, извлеченных для каждого пользователя, рассчитывается путем расчета площади прямоугольного треугольника. Используется формула:

$$T = \frac{1}{2}ab = \frac{ma\,x|p_i - q_i|\,.\ \sum_{i=1}^{n}|p_i - q_i|}{2}, \tag{2.21}$$

где a – основание, образованное от расстояния Чебышева; b – высота, образованная от манхэттенского расстояния p, q – значение временной метки для каждой клавиши, набираемой на клавиатуре.

Вторая часть общей площади рассчитывается путем расчета площади квадрата для каждой временной метки, образованной Манхэттенским расстоянием. Площадь вычисляется по формуле:

$$A = l^2 = ma\,x|p_i - q_i|^2\,, \tag{2.22}$$

где l – основание, образованное от расстояния Чебышева;

$$area_{test} = \sum_i \left(\frac{\frac{ma\,x|p_i - q_i|\,.\ \sum_{i=1}^{n}|p_i - q_i|}{2} + ma\,x|p_i - q_i|^2}{\mu_n} \right), \tag{2.23}$$

где $area_{test}$ – общее пространство значений временных меток на этапе тестирования; μ_n–все клавиши, введенные на клавиатуре в процессе входа в систему и ввода пароля

2.2.7 Рабочая среда этапа тестирования:

– Пользователь вводит логин и пароль для входа в систему. После этого выполняется этап сравнения введенной им информации с тем, что есть в

базе данных. как показное на рисунке 2.9.

– Создание биометрической модели для каждого пользователя, повторное извлечение характеристики динамики нажатия клавиш и расчёт временной метки для каждой характеристики. Значения модели хранятся на сервере в файле с расширением TXT .
– Вычисление общей площади прямоугольных треугольников и квадратов, образованных на основе значений временных меток извлеченных характеристик динамики нажатия клавиш.
– На последнем этапе производится сравнение, чтобы определить, значение меньше или равно общей площади порогового значения, полученного на этапе обучения. Если да, пользователь считается действительным и переходит к следующему этапу "идентификация рук", в противном случай считается недействительным.

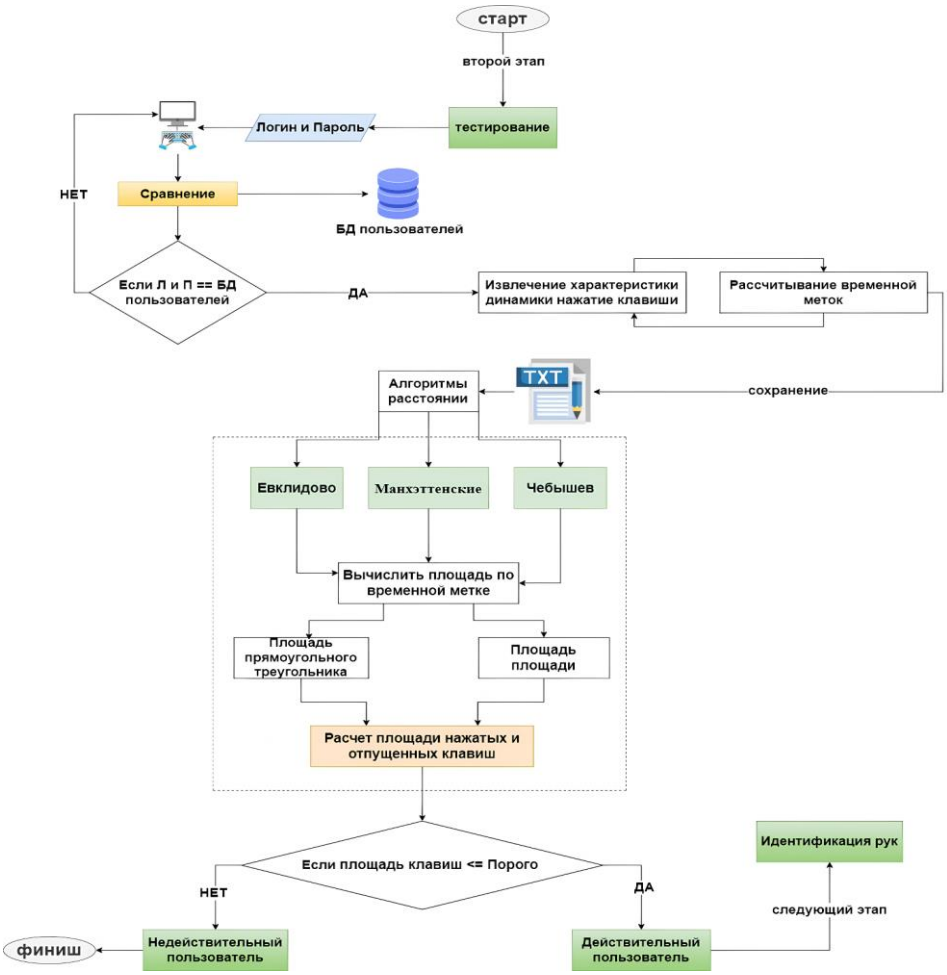


Рис. 2.9. блок-схема этапа тестирования динамики нажатия клавиш.

2.3 Определение руки по использованию клавиатуры

2.3.1 Описание основных принципов работы мягкой биометрии:

Мягкая биометрия: это характеристики, которые люди используют, чтобы различать своих сверстников. Это физические характеристики внешнего вида человека, связанные с человеком, или поведенческие характеристики, которые можно использовать в целях классификации. Эти категории, в отличие от классического биометрического случая, создаются и проверяются с течением времени людьми, где их можно понять и извлечь автоматически, чтобы дать описание человека, которое идентифицирует его или ее среди группы его или ее сверстников. Например, возраст, пол, цвет волос, цвет глаз, форма бровей и ресниц.

Биометрические характеристики, которых недостаточно для аутентификации пользователя, такие как рост, пол, кожа, глаза, цвет волос, и которые основаны на различиях черт людей (уникальное представление личности), и эти характеристики доступны всем [70-73]. Мягкая биометрия позволяет уточнить поиск реального пользователя в базе данных, что приводит к сокращению вычислительного времени. Например, если результат «захвата» биометрических данных определяет, что посетитель сайта – мужчина, согласно единице мягкой биометрии, стандартная система биометрической аутентификации может ограничить поле поиска до пользователя мужчины без учета женщин. Мягкая биометрия для аутентификации позволяет определять эмоциональное состояние (может быть обнаружено на 84 %), пол (на 90 %), одной или двумя руками набран текст на клавиатуре (на 80 %). Наиболее многообещающие результаты аутентификации по мягкой биометрии основаны на классификации уверенности, нерешительности, нервозности, расслабления, печали и усталости с точностью от 77 до 88 %, определении, левша или правша пользователь, и определении возрастной группы [71,73,74].

Мягкая биометрия состоит из вспомогательной информации, извлеченной из первичных биометрических образцов (таких как лицо, отпечатки пальцев, рука и радужная оболочка глаза).

Первое исследование дискриминационных возможностей мягких биометрических признаков было проведено французский учёный Альфонсом Бертильоном (1956) в девятнадцатом веке, который представил идею подхода к идентификации личности, основанного на биометрических, морфологических и антропометрических признаках [68].

Термин «мягкая биометрия» был введен в Джейн и др (2004) для описания набора характеристик, которые предоставляют некоторую информацию о человеке, но не являются уникальными для этого человека, в основном из-за отсутствия отличительности и постоянства. Определение мягких биометрических признаков развивалось с течением времени, связывая мягкие биометрические признаки с «метками», которые люди используют для описания друг друга [69].

Как описано в Данчевой и др. (2011), чтобы повысить эффективность систем классификации и распознавания даже в сложных условиях и сделать их более гибкими, используется мягкая биометрия в сочетании с дополнительными функциями. Они используются в криминалистических системах, при обнаружении преступлений и в разведывательных приложениях для классификации набора характеристик человека [67].

2.3.2 Разработка модели идентификации рук на основе динамики нажатия клавиш.

Для повышения степени безопасности на первом этапе процесса аутентификации разработана дополнительная модель, основанная на динамике нажатия клавиш для определения и распознавания руки, печатающей на клавиатуре на основе законов кинематики.

Абстрактная наука о движении или кинематика — это одна из отраслей науки о движении, которая описывает концепцию физического движения объектов без учета причины движения, такой как массы или силы её также называют геометрией движения. Следовательно, это противоположность науке о движении (динамике), которая занимается силами и взаимодействиями, которые производят движение или влияют на него [193].

Для определения типа рук на клавиатуре была использована кинематика, поскольку без использования искусственного интеллекта и внешних устройств сложно определить и измерить силу, влияющую на движение пользователя при переходе от одной клавиши к другой, в зависимости от длины пароля. Поэтому целью исследовательского проекта является разработка механизма аутентификации без использования внешних устройств.

В этой части исследования, использована клавиатура (QWERTY) для идентификации пишущей руки пользователя - пишет ли он одной или двумя руками.

QWERTY на сегодняшний день является наиболее широко используемой раскладкой английской клавиатуры. Название «QWERTY» происходит от первых шести клавиш на этих клавиатурах: QWERTY. Эта клавиатура была разработана в 1874 году американским изобретателем пишущих машинок Кристофером Шульцем и позже использовалась для компьютерных клавиатур. Хотя это, возможно, не самая эффективная раскладка для письма на английском языке, поскольку существуют новые конструкции QWERTY, такие как раскладка Дворжака [75].

Для исследования клавиатура была разделена на восемь частей, четыре ряда и два столбца, чтобы облегчить процесс нахождения нажимаемой и отпускаемой клавиши как показано на рисунке 2.10.

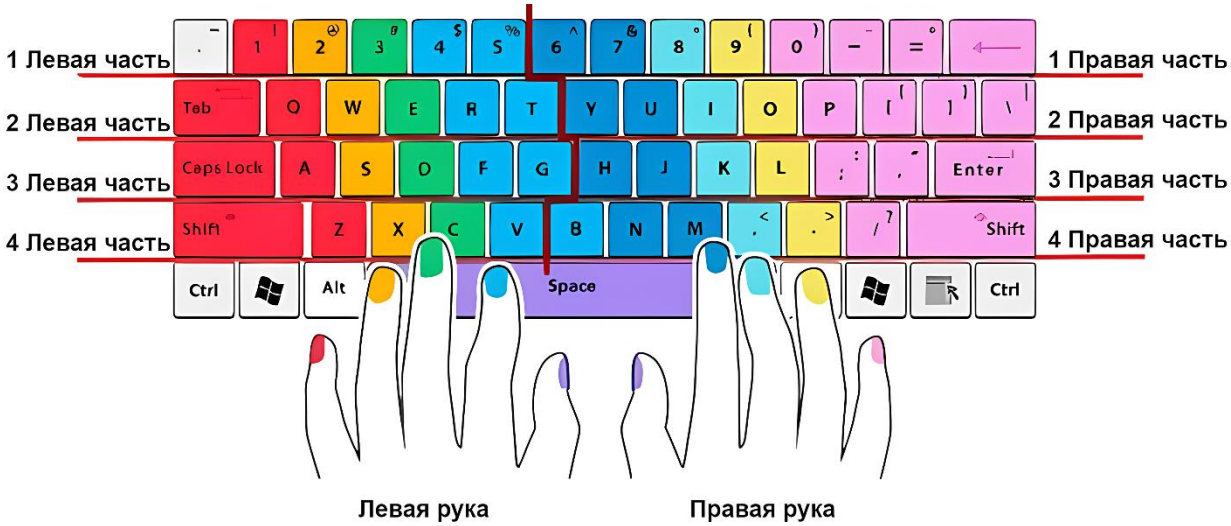


Рис. 2.10. Разделение клавиатуры для определения пишущих рук (1-2)

Каждой цифре, букве или символу на клавиатуре присвоено значение в шестнадцатеричной системе, и соответственно было создано 8 матриц на основе разделения клавиатуры на части как показное в таблице 3.

Американский стандартный код обмена информацией (ASCII, аббр. от англ. American Standard Code for Information Interchange) представляет собой стандарт кодирования символов для электронной связи. Коды ASCII представляют текст в компьютерах, телекоммуникационном оборудовании и других устройствах. В связи с техническими ограничениями компьютерных систем на момент его изобретения ASCII имеет всего 128 кодовых точек, из которых только 95 являются печатными символами, что серьезно ограничивает его возможности. Современные компьютерные системы используют Unicode, который имеет миллионы кодовых точек, но первые 128 из них совпадают с набором ASCII [76, 78, 79].

ASCII частично был разработан на основе телеграфного кода. Его первое коммерческое использование было в Teletype Model 33 и Teletype Model 35 в качестве семибайтного кода телетайпа, продвигаемого службами передачи данных Bell. Работа над стандартом ASCII началась в мае 1961 г. с первого заседания подкомитета X3.2 Американской ассоциации стандартов (ныне Американский национальный институт стандартов – сокр. ANSI). Первое

издание стандарта было опубликовано в 1963 г., В 1967 г. – было серьезно переработано, последнее обновление произошло в 1986 г. По сравнению с более ранними телеграфными кодами, предлагаемые коды Bellu Piece и ASCII были упорядочены для более удобной сортировки (т. е. расстановки в алфавитном порядке) списков и дополнительных функций для устройств, отличных от телетайпов [73, 79,188].

Первоначально основанный на (современном) английском алфавите ASCII кодирует 128 указанных символов в семибитные целые числа.

95 закодированных символов можно раскодировать: к ним относятся цифры от 0 до 9, строчные буквы от “a” до “z”, прописные буквы от “A” до “Z” и символы пунктуации. Кроме того, исходная спецификация ASCII включала 33 непечатаемых управляющих кода, созданных в моделях телетайпов; большинство

из них уже устарели, хотя некоторые из них все еще широко используются, например, carriage return, linefeed и коды табуляции [80].

Код ASCII используется для расчета значения характеристики динамики нажатия клавиш. Например, строчная буква “e” будет представлена в кодировке ASCII как двоичное число 1101001 = шестнадцатеричное 69 (e – девятая буква) = десятичное 105.

Однако при случайном и беспорядочном нажатии кнопок клавиатуры (например, кликнуть клавишу «Q», затем – «K», а после – «Z») сложно найти общее расстояние между нажатым пользователем клавишами. Основываясь на законах кинетической физики, расстояние определяется, как сумма полного движения тела, независимо от направления движения, совершаемого этим телом. Выходит, расстояние является стандартной величиной, а также его можно определить, как длину – путь между начальной и конечной точками. Т. е. итоговое расстояние можно определить, измерив дистанцию, соединяющую каждую клавишу, использованную при движении рук / руки от начальной к конечной точки.

Таблица 3 — коды символов и клавиш по ASCII

Часть	**символов и клавиш**	**коды**
Первая левая часть	~,1,2,3,4, 5, @, $, %	126,49,50,51,52,53,64,36,37
Первая правая часть	6,7,8,9,0,-, =, ^, &, *, (,), _, +	54,55,56,57,48,45,61,94,38,42,40,41,95,43
Вторая левая часть	Tab,q,w,e,r,t,y,u,I,o,p,[,],\ , {,},|	9,113,119,101,114,116,121,117,105,111,112 , 91,93,92,123,125,124
Третья левая часть	Capslook,a,s,d,f,g	20,97,115,100,102,103
Третья правая часть	h, j, k, l,’,”	104,106,107,108, 59,39,58,34
четвертая левая часть	z,x,c,v	122,120,99,118

Четвертая правая часть	b,n,m,,,.,/, <,>,?	98,110,109,44,46,47,60,62,63

Определение места каждого нажатия и отпускания клавиши на клавиатуре чрезвычайно важно для определения типа руки. Это связано с тем, что буквы и символы на клавиатуре расположены не на одной прямой линии, а в нескольких рядах. чтобы найти скорость руки и пройденное расстояние, необходимо определить место движения каждой кнопки, рассчитывается по формуле:

$$mk_{mc} = \begin{pmatrix} m_i c_i & m_i c_{i+1} & m_{i+1} c_{i+n} \\ m_{i+1} c_i & m_{i+1} c_{i+1} & m_{i+1} c_{i+n} \\ \vdots & \cdots & \cdots \\ m_8 c_i & m_8 c_{i+1} & m_8 c_{i+n} \end{pmatrix}, \quad (2.24)$$

где mk – матрица расположении клавиш; m номер части разделенной клавиатуры; c – расположение кнопки в части.

Независимо от того, какая задействована рука (правая или левая), скорость набора одной рукой заметно ниже, чем при использовании обеих рук, так как в последнем случае расстояние, которое человек проходит при переключении с одной буквы на другую, короче, чем при наборе одной рукой.

Исследователи А. Перейра, Д.Л. Ли, Х. Садишкумар, Ч. Ларош, Д. Оделл и Д. Рempел опубликовали результаты исследования, проведенного с целью изучения влияния расстояния между клавишами на скорость набора текста, количество допускаемых ошибок, удобство использования, активность мышц предплечья и положение запястья. Исследование было сосредоточено на конструкции традиционной механической клавиатуры, а не на экранной, создаваемой программным обеспечением [79, 81, 82]).

На основании исследований А. Перейра и других ученых в 1970-х гг. в работе [83] указано, что на обычное расстояние между клавишами на клавиатуре больше влияет отраслевая практика, чем вопросы эргономики (скорость набора текста, биомеханика, частота ошибок и удобство использования). Исследователи отмечают, что Международная организация по стандартизации (ISO), Американский национальный институт стандартов и Общество человеческого

фактора и эргономики (ANSI/HFES) рекомендуют, чтобы горизонтальное и вертикальное межцентровые расстояния (при взгляде на клавиатуру сверху) составляли 19 мм + / –1 мм, хотя не все, но большинство конструкций клавиатуры соответствуют этим стандартам.

Взаимосвязь производительности и расстояния между клавишами была рассмотрена в исследованиях А. Перейра и др. исследователей. Так, в Японии исследователи пришли к выводу, что у людей с маленькими пальцами не наблюдается снижение производительности при различном расстоянии между клавишами (диапазон – от 15 до 19,7 мм); в то же время производительность действительно была снижена у людей с большими пальцами, если расстояние между клавишами составило 16 мм или меньше (исследователи предостерегают от приложения выявленных результатов к населению США или других стран из-за различий в антропометрии рук) [82,83]. В работе [61] показано, что увеличено время ввода и частота ошибок с использованием цифровых клавиатур при увеличении расстояния с 19 до 21 мм. Обзор литературы 1972 г., в котором рассматривались параметры конструкции клавиатуры того времени, дал основание полагать, что оптимальное расстояние между центрами клавиш составляет 18,1 мм [73]. В работах [84, 85] было выявлено, что при совместном рассмотрении скорости набора текста, частоты ошибок и предпочтений пользователя интервал в 19 мм был лучшим (по сравнению с другими: 14,3; 16,6; 21,4).

Ни в одном из вышеупомянутых исследований не рассматривалось ключевое влияние расстояния на биомеханические или физиологические показатели. Понимая, что люди с пальцами меньшей длины скорее всего лучше адаптируются к сокращению расстояния между клавишами, А. Перейра сосредоточил внимание на людях с более длинными пальцами, и в статье [70] автор в первую очередь исследует горизонтальное расстояние между клавишами.

2.3.3 Процесс построения модели пользователя на этапе обучения по результатам идентификации рук.

Определение направления движения набора текста на клавиатуре очень важно для получения расстояния и расчета скорости, которую преодолевает пользователь при наборе пароля одной или двумя руками. Поэтому движение использовалось позиционно между собой координата x и координата y, как показано на рисунке 2.11.

Буквы и символы на клавиатуре расположены частями, разделенными на ряды, и поэтому, когда пользователь переходит от одной строки к другой, это считается перемещением по координате Y, а когда он переходит от одной клавиши к другой клавише по той же строке, это считается позиционным перемещением по координате X, как показное на рисунке 2.11.

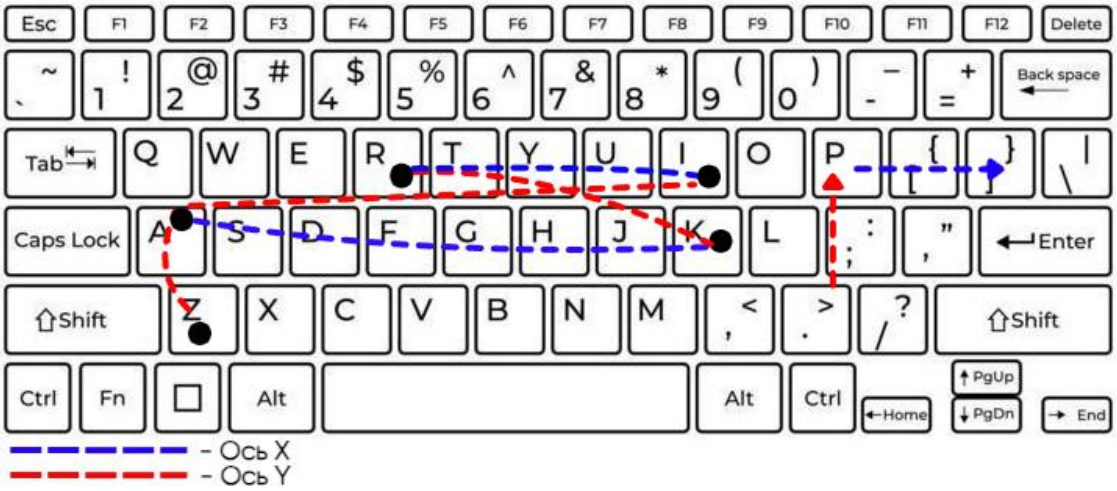


Рис. 2.11. Принцип движения рук по оси x и y

Векторное расстояние для определения руки рассчитывается по формуле:

$$r = (x, y) = x\vec{\imath} + y\vec{\jmath}, \tag{2.25}$$

$$r = key_z\,\vec{\imath} + key_z\,\vec{\jmath} + key_{z+1}\vec{\imath} + key_{z+1}\,\vec{\jmath} + \cdots key_{n+1}\vec{\imath} + key_{n+1}\,\vec{\jmath}, \tag{2.26}$$

$$r_{towhand} = \sum_{z=1}^{n} key_z\,\vec{\imath} + key_z\,\vec{\jmath}\ , \tag{2.27}$$

$$r_{towhand} = \sum_{z=1}^{n} (key_z\,\vec{\imath} + key_z\,\vec{\jmath}\,) \,.\, |mk_{mc} - mk_{m+1\,c+1}|, \tag{2.28}$$

где $r_{towhand}$ – расстояние вектора положения двух рук; $r_{onehand}$ – расстояние вектора положения одной руки; $\vec{i}$ - вектор движения по оси X; $\vec{j}$ – вектор движения по оси Y; key_z – клавиша при нажатии и отпускании; n– длина пароля.

Стандартная векторная скорость для обеих рук рассчитывается по формуле:

$$\vec{v} = \frac{\Delta_r}{\Delta_t} = \frac{d_{\vec{r}}}{d_{\vec{t}}} = \frac{d_x}{d_t}\ i + \frac{d_y}{d_t}\ j, \tag{2.29}$$

$$\vec{v} = \dot{x}\vec{i} + \dot{y}\vec{j}, \tag{2.30}$$

$$|\vec{v}| = \sqrt{\dot{x}\vec{i} + \dot{y}\vec{j}}, \tag{2.31}$$

$$|\vec{v}|_{towhand=}\sqrt{\dot{x}\vec{i} + \dot{y}\vec{j}} = \sum_{i=1}^{n} \sqrt{\frac{r_{towhand}}{t_{i+1} - t_i}}, \tag{2.32}$$

$$r_{towhand} = \frac{|\vec{v}|_{towhand}}{\Delta_t}, \tag{2.33}$$

где $|\vec{v}|_{towhand=}$ – стандартная векторная скорость для обеих рук;

Стандартная векторная скорость для одной рукой рассчитывается по формуле:

$$|\vec{v}|_{onehand=}\sqrt{\dot{x}\vec{i} + \dot{y}\vec{j}} = \sum_{i=1}^{n} \sqrt{\frac{r_{onehand}}{t_{i+1} - t_i}}, \tag{2.34}$$

$$r_{onehand} = \frac{|\vec{v}|_{onehand}}{\Delta_t}, \tag{2.35}$$

где $|\vec{v}|_{onehand}$ – стандартная векторная скорость для одной руки;

Высчитывается расположение всех букв, которые были написаны обеими руками или одной рукой в качестве пароля, по формуле:

$$position_{towhand} = \frac{\sum_{i=1}^{n} r_{towhand}}{\mu}, \tag{2.36}$$

$$position_{onehand} = \frac{\sum_{i=1}^{n} r_{onehand}}{\mu}, \tag{2.37}$$

где $position_{towhand}$ – положение букв относительно пространства при письме двумя руками; $position_{onehand}$ – положение букв относительно пространства при письме одной руки μ–количество нажатых и отпущенных клавиш.

2.3.4 Рабочая среда этапа обучения

Функция идентификации пишущей руки работает на этапе обучения при выполнении первого этапа динамического нажатия клавиш. Они представляют собой извлеченные характеристики динамического нажатия клавиш и временных меток, такие же используется на этапе идентификации руки. Таким образом, этап обучения незаметно сокращает время и не вызывает затруднений у пользователя, как показано на рисунке 2.12.

1. Пользователь вводит пароль, используя свой собственный стиль письма, либо одной рукой, либо двумя руками, в зависимости от привычного ему способа, и трижды повторяет пароль в текстовом поле;
2. Создается биометрическая модель для каждого пользователя и извлекаются те же динамические характеристики нажатий клавиш, которые были извлечены на первом этапе, а также временные метки для каждой характеристики. Значения модели хранятся на сервере в файле с расширением CSV.
3. Расчет значения векторного расстояния по позиционному перемещению как для двуручного письма, так и для одноручного письма.
4. На последнем этапе производится сравнение: превышает ли векторное расстояние обеих рук векторное расстояние при письме одной рукой, если да, то пользователь пишет двумя руками, в противном случае пользователь пишет одной рукой.

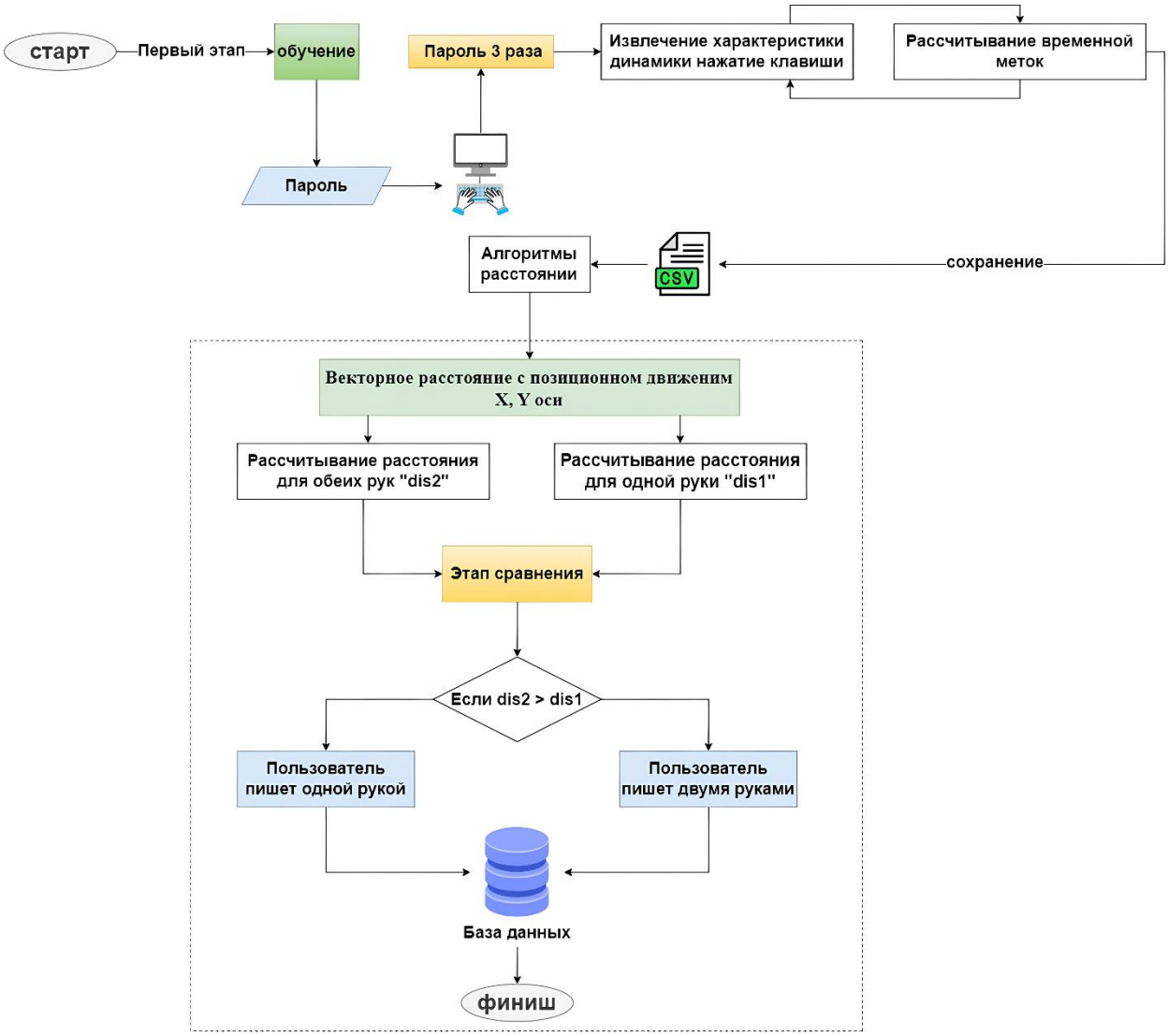


Рис. 2.12 Блок-схема этапа обучения идентификации рук

2.3.5 Процесс подтверждения модели пользователя на этапе тестирования по результатам идентификации рук.

Этап проверки и тестирования зависит от определения скорости и пройденного расстояния при письме двумя руками и письме одной рукой.

То есть скорость прохождения расстояния двумя руками больше скорости письма одной рукой, как показное на рисунке 2.13 и 2.14. рассчитывается по формуле:

$$spd_{towhand} = |r_{towhand} - position_{towhand}|, \quad (2.38)$$

$$spd_{onehand} = \left|\frac{position_{onehand}}{\mu * 10}\right|, \quad (2.39)$$

Где $spd_{towhand}$ – скорость прохождения дистанции двумя руками; $spd_{onehand}$ – Скорость прохождения дистанции одной рукой;

Рис. 2.13. Динамика нажатия клавиш с абстрактной кинематикой одной руки

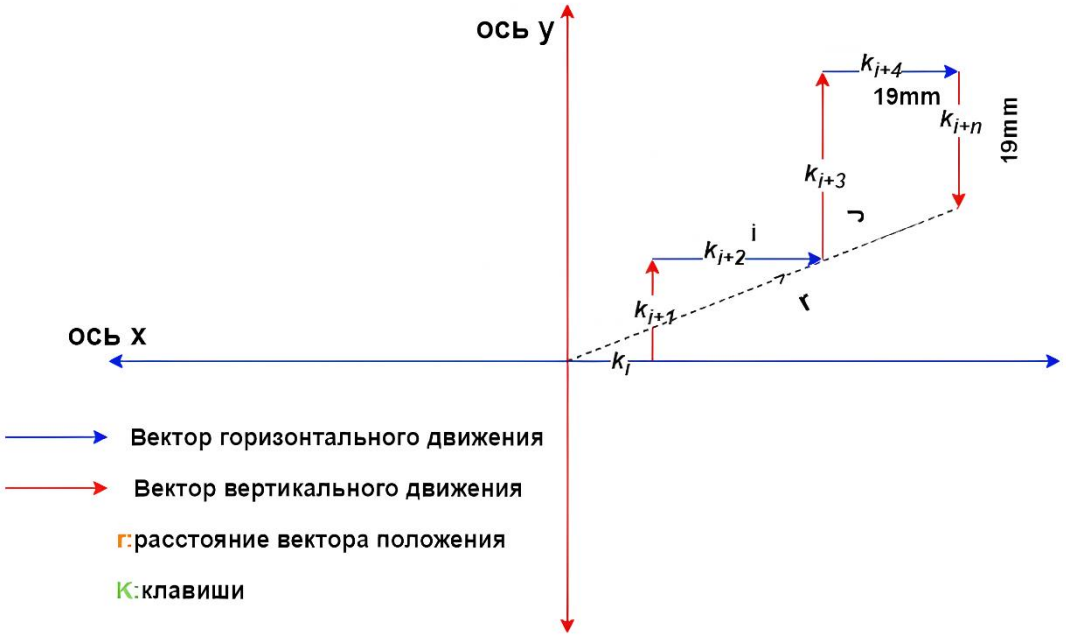


Рис. 2.14. Динамика нажатия клавиш с абстрактной кинематикой двух рук

2.3.6 Рабочая среда этапа тестирования:

Вначале обеспечивается успешное завершение первого этапа проверки динамики нажатия клавиш. Если он не завершен, процесс проверки возвращается к первому этапу, так что пользователь не может перейти к следующему этапу

проверки, если он не проходит предыдущий этап, как показано на рисунке 2.15.

1. Если пользователь проходит первый этап проверки нажатий клавиш, используются временные метки и динамические характеристики извлеченных нажатий клавиш, хранящиеся на сервере в файле с расширением TXT.
2. Алгоритм расчета векторной скорости основан на принципе позиционного движения по координатам x и y, рассчитывается для каждого скорость прохождение дистанции одной и двумя руками.
3. Сравнение производится, если с скорость прохождения дистанции двумя руками больше скорости прохождения дистанции одной рукой, то пользователь пишет двумя руками, в противном случае он пишет одной рукой, и значение сохраняется на сервере в файле с расширением CSV.
4. После этого производится сравнение со значениями, которые были рассчитаны на этапе обучения. Если значение, хранящееся в файле с расширением CSV, равно такому же значению, хранящемуся в базе данных, пользователь считается действительным и переходит в следующей этап, представляющий собой динамику нажатая мышьи, в противном случае пользователь считается недействительным, и сеанс прекращается или пользователь повторяет этап тестирование заново.

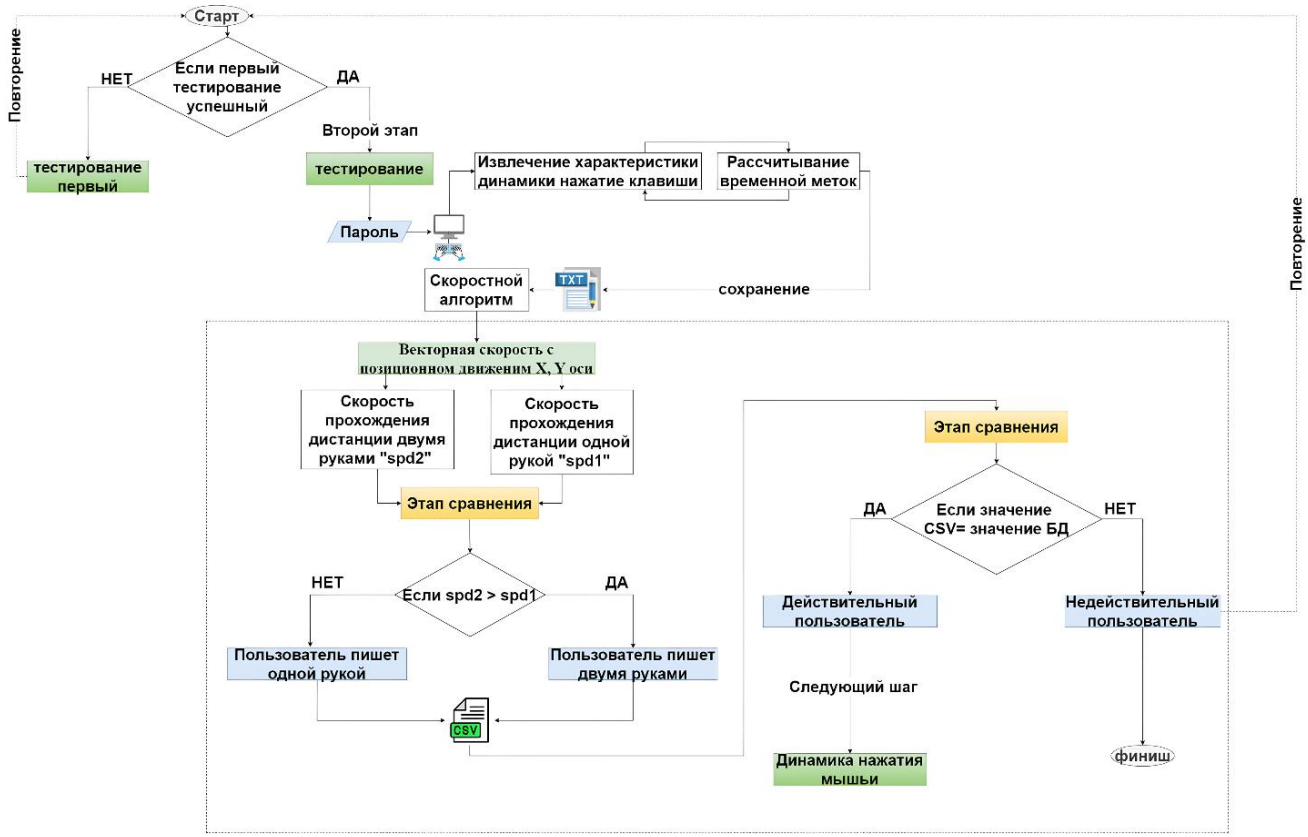


Рис. 2.15. блок-схема этапа тестирования идентификации рук

2.4 Биометрическая модель динамика движений мыши

2.4.1 Описание основных принципов работы динамики мыши

Аутентификация пользователей является неотъемлемой частью любой политики безопасности для любой системы, целью которой является проверка личности пользователей с целью предоставления пользователям доступа к их собственным учетным записям и взаимодействиям в системе, предотвращения операций взлома и кражи их учетных данных, а также предотвращения смешивания своих учетных данных с учетными данными другого пользователя.

Биометрические системы были самыми используемыми, включая динамику движения мыши, которая показывает стиль движения, а также давления на мышь, которые различаются у всех людей. Динамика движения мыши — интересный и недорогой биометрический метод, для которого нет необходимости в дополнительном датчике. Это позволяет биометрической системе аутентифицировать или идентифицировать человека на основе того, как человек

щелкает и перемещает указатель мыши по местоположению окна для ввода пароля. Динамика движений мыши относится к категории поведенческой биометрии, то есть шаблон пользователя отражает тот или иной аспект его поведения.

Впервые динамика мыши для повторной аутентификации ранее изучалась как автономная биометрия в 2004 году Пусара и Бродли [34]. К сожалению, их исследование не дало окончательных результатов, так как в нем участвовало всего одиннадцать пользователей, что побудило авторов сделать вывод о том, что биометрия мыши недостаточна для повторной аутентификации пользователя. Они также предложили проверять, двигалась ли мышь каждые 100 миллисекунд, поскольку временной интервал для каждого движения мыши составляет 100 миллисекунд, что с точки зрения процессора компьютера является очень долгим временем, на основе поведения движения мыши с ложноположительным показателем 0,43% и ложноотрицательным показателем 1,75%.

Позже, в 2005 году, Ахмед и Траоре [86] предложили подход, объединяющий динамику нажатия клавиш с динамикой мыши. При достижении очень высокой точности, количество действий мыши, необходимых для проверки личности пользователя, слишком велико, чтобы быть практичным. В частности, их эксперимент требует до 2000 совокупных действий мыши, прежде чем пользователь может быть распознан, что непрактично для развертывания в реальном времени.

Биометрия динамики мыши — это поведенческая биометрическая технология, которая извлекает и анализирует характеристики движения устройства ввода мыши, когда пользователь компьютера взаимодействует с графическим пользовательским интерфейсом в целях идентификации. Это серия движений, то есть жестов, причем каждый жест представляет собой конкретный и непрерывный физический процесс, инициируемый и завершаемый пользователем [88,89]. Основная цель отслеживания мыши — глубже понять поведение пользователя и сделать выводы о его намерениях. Есть отслеживание глаз, но отслеживание мыши дешевле и лучше. Каждое событие мыши, выполняемое в системе, будет аутентифицировано.

Область взаимодействия человека и компьютера особенно заинтересована в изучении человеческого поведения, поскольку оно может пролить свет на деятельность человека. Отслеживание мыши и клавиатуры, согласно предыдущим исследованиям HCI, может дать более полную картину поведения пользователя при наличии высоких когнитивных нагрузок, таких как принятие решений и разработка действий [89].

Динамика мыши рассматривается как пример поведенческой биометрии, представленной в Университете Виктории в исследовательской лаборатории информационной безопасности и объектных технологий (ISOT) в 2003 году. Основные преимущества динамики мыши по сравнению с другими биометрическими технологиями заключаются в том, что она позволяет динамически и пассивно отслеживать пользователя. Соответственно, ее можно использовать для непрерывного отслеживания подлинных и самозваных пользователей во время сеансов вычислений. Когда пользователь использует устройство мыши, анализируются ее характеристики действий. В динамике мыши некоторые характеристики извлекаются из действий мышью пользователем. Затем эти характеристики сохраняются. Когда пользователь хочет получить доступ к системе позже, система сравнивает его/ее действия с сохраненными и решает, является ли он/она подлинным или самозванцем [87].

Когда пользователь пытается войти в компьютерную систему, динамика мыши требует от него только ввода имени для входа и выполнения определенной последовательности операций с мышью [90-92]. Извлеченные поведенческие характеристики, основанные на движениях мыши и щелчках мыши, сравниваются с профилем законного пользователя. Соответствие аутентифицирует пользователя; в противном случае ему будет отказано в доступе. Более того, характеристики поведения мыши пользователя можно постоянно анализировать во время его последующего использования компьютерной системы для мониторинга идентификации или обнаружения вторжений [93].

2.4.2 Анализ структуры и характеристики динамики мыши

Поведение мыши основано на двух основных типах операций, щелчке и движении мыши. Затем каждая операция мыши анализировалась индивидуально и переводилась в несколько характеристик, как показное в таблице 4. Этот динамический процесс разделил эти характеристики на две категории:

1. характеристики, характеризующие общие свойства поведения мыши во время взаимодействия, такие как статистика одиночного и двойного щелчка как показное на рисунке 16.
2. характеристики, отображающие подробные динамические процессы поведения мыши, такие как скорость движения.

Таблица 4 — Характеристики динамики мыши

Действия мыши	**Описание**	**Формула**
Один щелчок	Событие на мыши при нажатии и отпускании правой или левой кнопки.	$m_{du} = time_{mdu_i} - time_{mud_i}$ $m_{ud} = time_{mud_i} - time_{mdu_i}$
Двойной щелчок	Событие на мышке при двойном одновременном щелчке по одному и тому же элементу	$m_{dudu} = (time_{mdu_i} - time_{mud_i}) + (time_{mdu_{i+1}} - time_{mud_{i+1}})$
Движение мыши	Это естественное движение мыши без щелчков (Данные горизонтальной координаты оси x, Данные вертикальной координаты оси y)	$m_{mov} = (x_i - y_i) + (x_{i+1} - y_{i+1}) + \cdots + (x_i - y_i)$
Движение и щелчок	Это движение мыши с последующим щелчком по элементу или двойным щелчком мыши.	$m_{mov_{oc}} = (x_i - y_i) + mdu$ $m_{mov_{dc}} = (x_i - y_i) + mdudu$
Перетаскивание	Это событие, которое начинается при постоянном нажатии на клавишу мыши ее движении, а затем отпусканим клавиши.	$m_{dp} = time_{md_i} - m_{mov} - time_{mu_i}$
Тишина	Ситуация без каких-либо действий мыши	x=20 миллисекунд

		$ms = \begin{cases} 1, & m_{du} - m_{du} == 0 \\ 1, & m_{mov} - m_{mov} == 0 \end{cases}$
Скорость движения мыши	Это скорость движения мыши при перемещении между элементами по горизонтальной линии, вертикальной линии или кривой.	$msp_{mov} = \frac{m_{mov}}{time_i - time_{i-1}}$

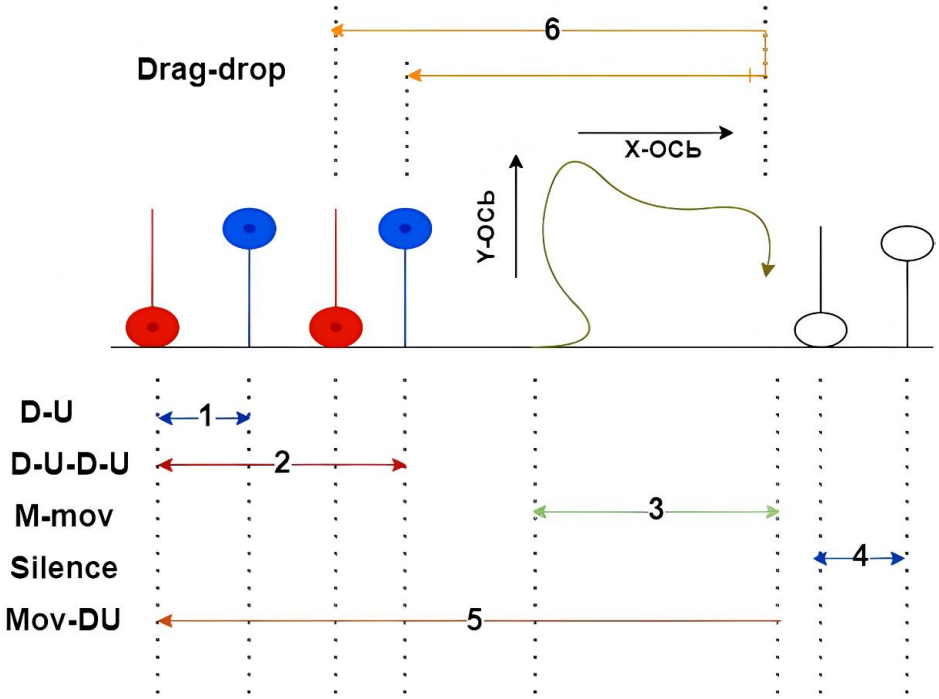


Рис. 2.16. Характеристики поведения мыши

Процесс обучения динамики мыши ограничивается перемещением при нажатии и отпускании и другими особенностями динамики мыши в текстовом поле для написания пароля. Аналогично на этапе проверки ограничивается диалоговым окном для указания логина и пароля. Следовательно, чтобы снизить процент ложных отклонений, необходимо построить соответствующий порог, который варьируется от порога, используемого в динамике нажатий клавиш. В дополнение к Манхэттенскому расстоянию использовалось евклидово расстояние и расстояние Чебышева. В динамике мыши используется, для увеличения порогового значения, наряду с Манхэттенским расстоянием, Минковское расстояние, поскольку оно возвращает кратчайшее расстояние в форме кривой между двумя точками. Таким

образом, общая площадь равно увеличивается, и действительному пользователю разрешено войти в систему с соответствующим ему порогом, как показное на рисунке 2.17.

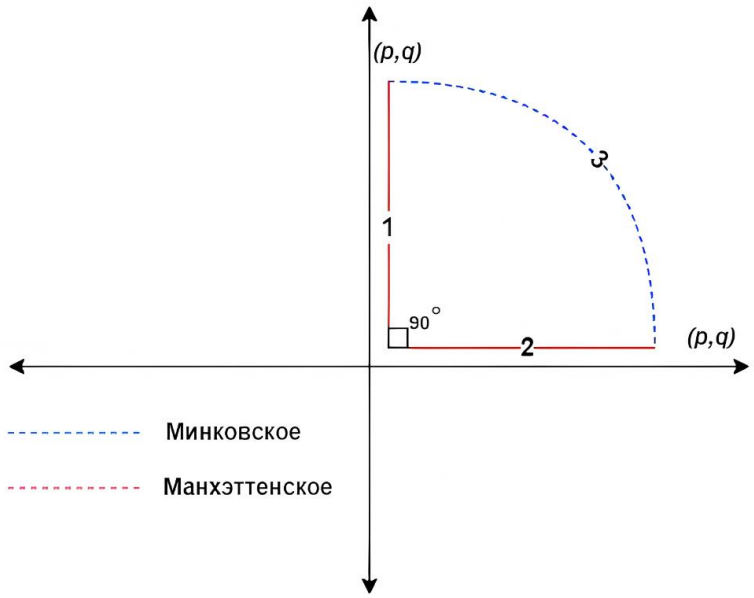


Рис. 2.17. Расстояние, образованное Минковскими и Манхэттенскими метриками

Расстояние Минковского — это универсальная метрика, используемая в нормированных векторных пространствах, названная в честь немецкого математика Германа Минковского [191,192]. Это обобщение нескольких известных мер расстояния, что делает его фундаментальным понятием в различных областях, таких как математика, информатика и анализ данных.

По своей сути расстояние Минковского предоставляет способ измерения расстояния между двумя точками в многомерном пространстве, что делает его особенно полезным, так это его способность охватывать другие метрики расстояния как особые случаи, в первую очередь через параметр p. Этот параметр позволяет расстоянию Минковского адаптироваться к различным проблемным пространствам и характеристикам данных.

Расстояние Минковского рассчитывается по формуле:

$$D(X,Y) = s\left(\sum_{i=1}^{n}|x_i - y_i|^p\right)^{\frac{1}{p}} = \sqrt[p]{\sum_{i=1}^{n}|x_i - y_i|^p}\,, \qquad (2.40)$$

где x, y – две точки в n-мерном пространстве; p – параметр, определяющий тип расстояния.

Когда p установлено равным 1, расстояние Минковского становится манхэттенским расстоянием. Используется формула:

$$D_{Manhattan}(x,y) = \left(\sum_{i=1}^{n}|x_i - y_i|^1\right)^{\frac{1}{1}} = \left(\sum_{i=1}^{n}|x_i - y_i|\right), \quad (2.41)$$

Когда p установлено равным 2, расстояние Минковского становится расстоянием Евклида. Рассчитывается по формуле:

$$D_{Euclidean}(x,y) = \left(\sum_{i=1}^{n}|x_i - y_i|^2\right)^{\frac{1}{2}} = \sqrt{\sum_{i=1}^{n}|x_i - y_i|^2}, \quad (2.42)$$

2.4.3 Процесс построения модели пользователя на этапе обучения по результатам динамики мыши.

Пороговое значение определяется для каждого пользователя на этапе обучения на основе того, как он использует мышь, и в соответствии с характеристиками, которые были извлечены для каждого пользователя и описаны в таблице 3. Пороговое значение создается с использованием расстояния Минковского и Манхэттенского расстояния помимо площади четверти круга.

Чтобы получить лучшее пороговое значение для динамики мыши, кривая в четверть круга рассчитывается с использованием расстояния Минковского. Экспериментально было определено значение p, чтобы найти лучшую кривую для порогового значения, которое составило 0,9.

$$curve_{c_{quadrant}} = D_{Minkowski}(X,Y) = \sqrt[0.9]{\sum_{i=1}^{n}|x_i - y_i|^{0.9}}, \quad (2.43)$$

Квадрант формируется так, как показано на рисунке 32. Площадь четверти круга рассчитывается с помощью Манхэттенского расстояния, чтобы вычислить значение радиуса четверти.

$$A_{quadrant} = \frac{1}{4}\pi R^2 = \frac{\pi}{4} . \frac{\sum_{i=1}^{n} |x_i - y_i|^2}{2}, \tag{2.44}$$

$$\frac{\pi \sum_{i=1}^{n} |x_i - y_i|}{4}, \tag{2.45}$$

где $A_{quadrant}$ – площадь четверти круга; R – луч это одна из сторон четверти круга, и здесь он представляет собой Манхэттенское расстояние; x и y – операции с извлеченными характеристиками динамики мыши;

Пороговое значение динамики мыши в зависимости от площади четверти круга и ее окружности

$$threshould = \sqrt{\left(A_{quadrant} + R\right)} + \frac{\pi D}{4}, \tag{2.46}$$

$$= \sqrt{\left(A_{quadrant} + R\right)} + \frac{\pi 2R}{4}, \tag{2.47}$$

где $threshould$ – Пороговое значение; D – Окружность четверти круга.

2.4.4 Рабочая среда Этап обучения:

1. Пока пользователь на этапе обучения пишет пароль в назначенном текстовом поле, он использует мышь и перемещается между полями на оси кадра, содержащей три текстовых поля, как показано на рисунке 2.19.

2. Создается биометрическая модель для каждого пользователя и извлекаются те же динамические характеристики нажатий клавиш, которые были извлечены на первом этапе, а также временную метку для каждой характеристики. Значения модели хранятся на сервер в файле с расширением CSV как показано на рисунке 2.18.

3. Вычисляется четверть круга с использованием расстояния Минковского и Манхэттенского расстояния, затем вычисляется площадь полученной фигуры и длина дуги, чтобы вычислить пороговое значение для каждого пользователя.

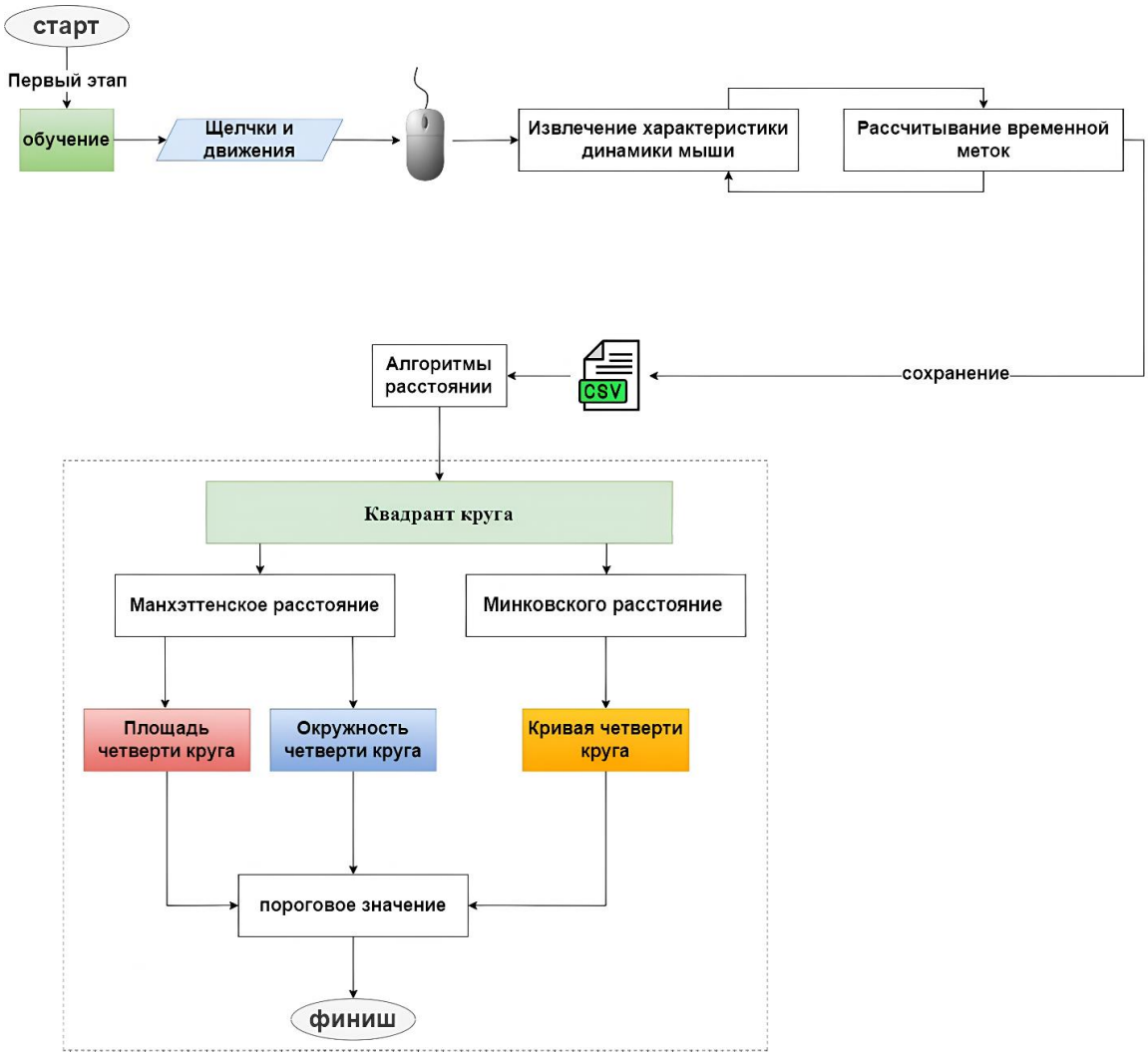


Рис. 2.18. Блок-схема этапа обучения динамики мыши

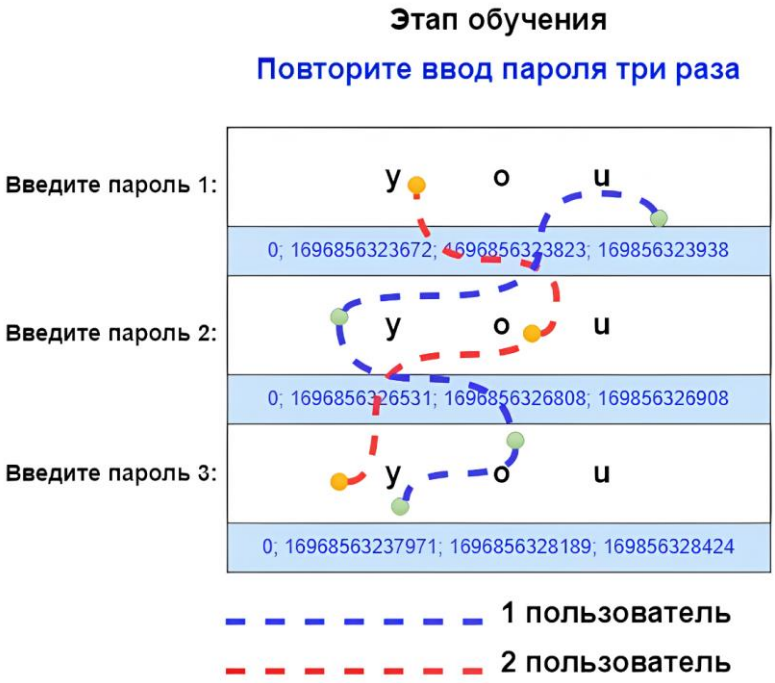


Рис. 2.19. Внедрение этапа обучения динамики мыши

2.4.5 Процесс подтверждения модели пользователя на этапе тестирования по результатам мыши.

Этап проверки и тестирования зависит от расчета расстояния Минковского, чтобы найти кривую прямоугольного сечения и сравнить ее с пороговым значением.

2.4.6 Рабочая среда этапа тестирования

Вначале обеспечивается успешное завершение первого этапа проверки динамики нажатия клавиш. Если он не завершен, процесс проверки возвращается к первому этапу, так что пользователь не может перейти к следующему этапу проверки, если он не проходит предыдущий этап проверки, как показано на рисунке 2.20.

1. Когда пользователь проходит первый и второй этапы теста, извлекаются характеристики и особенности динамики мыши, показанные в таблице 33, а временная метка для каждого из них рассчитывается и сохраняется на сервере в файле с расширением TXT.
2. Через расстояние Минковского рассчитывается кривая в четверть круга, значение которой индивидуально для каждого пользователя при использовании мыши.
3. На последнем этапе сравнивается, меньше ли значение кривой или равно пороговому значению, рассчитанному на этапе обучения. Если ответ положительный, пользователь считается действительным и ему разрешено войти в систему. в противном случае пользователь считается недействительным, и сеанс прекращается или пользователь повторяет этап тестирование заново.

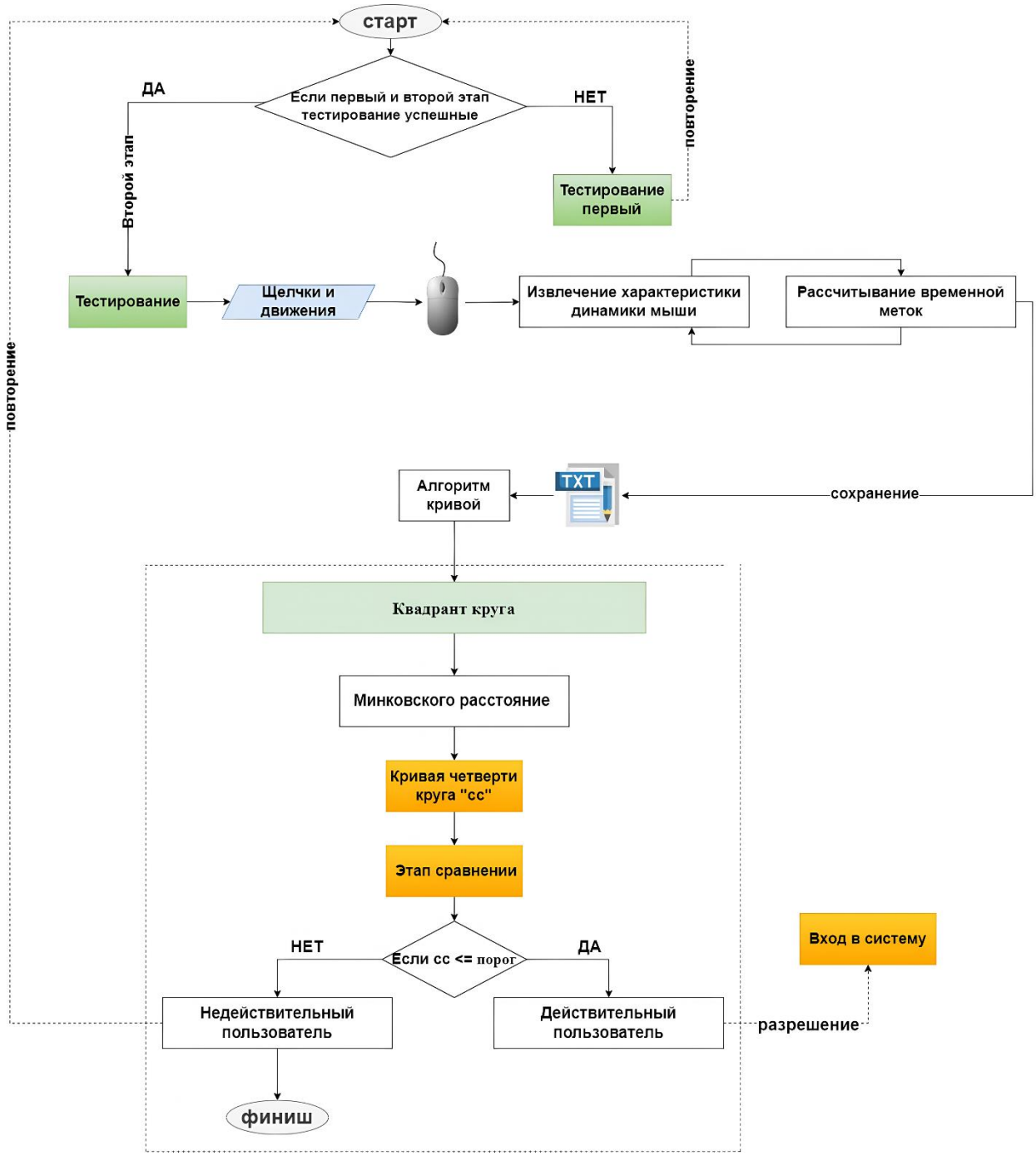


Рис. 2.20. Блок-схема этапа тестирования динамики мыши

Выводы по 2-й главе

Этот раздел содержит три научных результата по аутентификации пользователей на основе поведенческих биометрических измерений. В данном разделе проводилось исследование и разработка методов обработки данных, характеризующих динамику работы пользователей с клавиатурой и мышью.

В результате проведенного исследования были достигнуты следующие результаты:

- Веб-приложение, включающее этап обучения и проверки, для аутентификации через нажатия клавиш и динамики мыши, и идентификации почерка.
- Предлагается подход к повышению точности динамического нажатия клавиш путем использования и комбинирования Манхэттенского расстояния, Евклидова расстояния, расстояния Чебышева и закона Пифагора для прямоугольных треугольников. Данный подход позволил решить проблему для определения точного порогового значения и тем самым повысить качество аутентификации пользователей в среднем на 4% по сравнению с используемой в существующих работах стандартизацией признаков.
- Для повышения точности определения руки, используемой для печати на клавиатуре, был предложен подход, основанный на разделении клавиатуры на восемь частей и использовании законов кинематики для определения скорости и местоположения каждой нажатой клавиши на клавиатуре, что ранее не использовался для решения этой проблемы. Повысилось качество аутентификации пользователей в среднем на 10% по сравнению с используемой в существующих работах стандартизацией признаков.
- На первом этапе аутентификации, основанном на движении мыши пользователя и его перемещении между текстовыми полями для ввода логина и пароля, что считается ограниченным движением и, таким образом, уровень ложного принятия высок. Для решения этой проблемы был предложен подход, зависящий от объединения Манхэттенского расстояния и расстояния Минковского и, таким образом, повышения точности измерения порогового значения на основе четверти кривой круга, который отличает каждого пользователя по его движению и повышает качество аутентификации пользователей в среднем на 2% по

сравнению с используемой в существующих работах стандартизацией признаков.

- Была проведена серия экспериментов, в результате которой было подтверждено высокое качество работы предложенных алгоритмов. В результате использования предложенной комбинации алгоритмов достичь качества распознавания порядка 0.97 ROC AUC, что значительно превосходит качество аутентификации пользователей при использовании методов, рассмотренных в существующих научных работах.
- Поведенческие биометрические данные изменчивы с течением времени из-за своей баллистической природы и поэтому имеют высокий уровень ложных срабатываний. Чтобы решить эту проблему, биометрические данные пользователя обновляются новым пороговым значением после каждой успешной аутентификации.

ГЛАВА 3. 2. Методика трех факторной аутентификации пользователей для веб-приложения.

3.1 Описание основных принципов работы одноразового пароля OTP

В военной и правительственной сфере в 1960-х и 1970-х годах двадцатого века методы шифрования использовались для безопасной связи, что является тем же контекстом, что и система OTP, поскольку ее идея подразумевала использование специальных кодов для аутентификации, но они были бумажными кодов, поскольку пользователи вручную вводили уникальный код, напечатанный на карточке или бумажной ленте, и отсюда берет свое начало технология OTP [95, 96].

S/Key, разработанный Лесли Лампортом в 1984 году, был ранним примером системы аутентификации на основе OTP. Он использовал криптографический хэш, одностороннюю функцию шифрования для генерации последовательности одноразовых паролей на основе секретной парольной фразы пользователя [97].

В 1990-х годах были введены алгоритмы одноразовых паролей на основе времени (TOTP) для генерации OTP на основе текущего времени. Эти алгоритмы полагались на общий секретный ключ между сервером и устройством пользователя, обеспечивая синхронизированную по времени генерацию OTP [95].

Широкое распространение мобильных телефонов в 2000-х годах привело к разработке мобильных OTP решений. Мобильные приложения и доставка OTP на основе SMS стали популярными методами генерации и доставки одноразовых паролей пользователям [98, 99].

Одноразовый пароль (One-Time Password, OTP) — это способ аутентификации, при котором пользователю выдается уникальный пароль, который можно использовать только один раз для входа в систему или подтверждения личности [100, 101].

Этот метод обеспечивает дополнительный уровень безопасности, так как пароль действителен только в течение одного сеанса входа. Когда он использован,

он больше не действителен, что делает его невозможным для повторного использования злоумышленниками [101].

это автоматически сгенерированная числовая или буквенно-цифровая строка символов, которая аутентифицирует пользователя для одной транзакции или сеанса входа [102].

OTP более безопасен, чем статический пароль, особенно пароль, созданный пользователем, который может быть слабым и повторно использоваться в нескольких учетных записях.

OTP могут заменить традиционную информацию для входа в систему аутентификации или могут использоваться в дополнение к ней для добавления еще одного уровня безопасности [103].

Для генерации или получения одноразовых паролей часто применяются смартфоны. После того, как пользователь доказал, что владеет конкретным телефоном, он может использовать приложение аутентификатора, генерирующее пароли OTP — в этом случае телефон служит генератором кодов. Или же OTP могут отправляться в устройство через SMS [99].

Одноразовые пароли усиливают существующие системы идентификации и паролей, добавляя в них динамически генерируемые идентификационные данные [102].

Синхронные токены для создания одноразового пароля используют приватный ключ пользователя и текущее время.

Асинхронные токены используют Challenge Response Authentication Mechanism (CRAM) — серию протоколов, в которых сервер отправляет запрос, а токен должен сгенерировать правильный ответ [100].

Преимущества одноразовых паролей [102, 104]

Внедрение одноразовых паролей (OTP) может предложить более безопасную альтернативу или даже дополнение к запоминаемому статическому паролю в рамках процесса многофакторной аутентификации.

Это связано с тем, что скомпрометированный пароль будет бесполезен для того, кто пытается взломать учетную запись или приложение.

При использовании статических паролей хакер или мошенник, получивший пароль пользователя, будет иметь доступ к потенциально конфиденциальной информации до тех пор, пока этот пароль не будет изменен. В еще худшем случае тот, кто взломал эту учетную запись, может изменить пароль до того, как его законный владелец сможет изменить его и защитить свою информацию.

Благодаря своей одноразовой природе OTP могут защитить приложение или учетную запись, так что даже в случае, если злоумышленник перехватит пароль, он не сможет повторно использовать пароль при второй попытке. Пользователь, который станет жертвой фишинговой атаки или вредоносного ПО, перехватывающего нажатия клавиш, все равно будет защищен. Информация останется в безопасности от обычных методов кражи паролей.

В ходе исследования проекта OTP использоваться с динамикой нажатия клавиш с целью повышения безопасности путем создания многофакторной аутентификации и снижения риска фишинга.

Кибератаки определяются как любая цифровая попытка украсть, нарушить или получить несанкционированный доступ к вычислительной среде/инфраструктуре с целью кражи контролируемой информации [105].

Фишинг, который считается одним из основных типов кибератак, с которыми сталкиваются пользователи онлайн-сервисов, является опасным и все более распространенным явлением. Lastdrager определил фишинг как «масштабируемый акт обмана, при котором используется выдача себя за другое лицо для получения информации от цели» [106].

Этот термин был придуман как серьезная киберугроза в 1996 году, когда фишеры украли информацию об учетных данных пользователей America Online (AOL) [107].

В сентябре 2024 года, the Anti-Phishing Working Group (APWG) сообщила о 342.092 фишинговых веб-сайтах [108].

По данным APWG, это самый высокий ежемесячный показатель за всю историю отчетности. На рисунке 31 показано количество случаев фишинга с первого квартала 2024 года по третий квартал 2024 года.

Рис. 3.1. Обнаружение уникальных фишинговых сайтов

Согласно недавнему исследованию, проведенному IBM и Ponemon Institute, годовые расходы на фишинговые атаки в США значительно возросли в 2023 году, достигнув уровня, при котором крупные американские компании теперь выплачивают 4,76 миллиона долларов в год [109].

До сих пор стратегии борьбы с фишингом не были достаточно эффективными, чтобы снизить вероятность попыток фишинга. Данные пользователей должны быть защищены от фишинговых атак, научив их тому, как сообщать о фишинговых письмах и что делать в случае их получения. Это связано с тем, что фишеры стремятся найти недостатки и уязвимости в конкретном решении, чтобы их можно было использовать для выполнения успешной атаки.

Например, Национальный центр кибербезопасности Иорданского Хашимитского Королевства запустил платформу для повышения осведомленности компаний и учреждений с целью защиты их бизнеса от риска фишинга посредством повышения осведомленности сотрудников, технических обновлений и резервного копирования [110]. Этот подход сопоставим с тем, который реализовал Стэнфордский университет, который теперь предлагает услугу по повышению осведомленности о фишинге [111].

В результате многие организации в настоящее время работают над предотвращением несанкционированного доступа к своим системам путем ужесточения требований к аутентификации.

Биометрическая аутентификация является перспективной тенденцией для борьбы с фишинговыми атаками. Существует ряд различных систем, которые применяют биометрическую информацию в качестве средства идентификации людей, как в случае гражданской, правительственной и медицинской идентификации. Схемы биометрической аутентификации набирают популярность по сравнению с другими типами аутентификации в последние годы, поскольку они обеспечивают высокую безопасность для защиты личности людей и легко сочетаются с традиционными методами аутентификации. Биометрическую информацию можно приблизительно разделить на физиологические и поведенческие характеристики. Биометрическая информация, используемая в методах физиологической аутентификации, выводится из физических черт человека, таких как отпечатки пальцев и распознавание лица. Однако измерение этих характеристик очень дорого для развертывания, как и сопутствующее оборудование. Напротив, поведенческие характеристики основаны на том, что пользователи узнали или приобрели, что отличает их от других. К ним относятся нажатие клавиш и динамика мыши. Из множества возможных биометрических характеристик динамика нажатия клавиш является наиболее популярной и широко изучалась для целей распознавания [112]. Недавние исследования изучали эффективность динамики нажатия клавиш с целью повышения уровня безопасности в системах аутентификации. Эти исследования различались по подходу, принимая различные классы динамики нажатия клавиш (например, свободный и фиксированный текст), методы классификации шаблонов (такие как статистическое и машинное обучение) и экспериментальные среды (контролирующие или неконтролирующие). Все они дали многообещающие результаты, но результаты, полученные с использованием свободного текста, несомненно, более безопасны, чем те, которые получены с использованием фиксированного текста. Более того, многочисленные исследования пытались

изучить внутренние преимущества динамики нажатия клавиш со свободным текстом в обеспечении непрерывной и не интрузивной аутентификации.

В ответ на эту проблему было предложено множество способов уменьшить влияние попыток фишинга. Эти решения можно разделить на нетехнические (образовательные), технические (программное обеспечение) или их сочетание. Технические решения можно разделить на две категории: предотвращение и обнаружение фишинга. Обнаружение фишинга подразумевает проверку медиаконтента на веб-сайте. Он может подтвердить подлинность веб-сайта, используя изображение или URL [113].

Чтобы уменьшить влияние попыток фишинга, был разработан ряд методов обнаружения фишинга, включая поисковые системы [114], панели инструментов безопасности, машинное обучение [115], черные/белые списки [116] и методы, основанные на визуальном сходстве [117]. Улучшая безопасность веб-сайтов, решения по предотвращению фишинга направлены на предотвращение попыток фишинга. Это достигается путем внедрения специальных методов защиты платформ взаимодействия с пользователем и процедур аутентификации на практике [118].

было подчеркнуто, как поведенческая биометрическая аутентификация может использоваться для предотвращения фишинга. Вышеупомянутый автор обсудил, как динамика нажатия клавиш может использоваться для предотвращения фишинговых атак, а также для обеспечения того, чтобы фишеры не выдавали себя за пользователей [119].

Использование свободного текста было специально изучено в исследовании [120, 121], в котором использовались евклидовы концепции для определения расстояний между парами клавиш на основе их расположения на клавиатуре. Время полета важнее времени задержки, что является интригующим выводом этой работы, который напрямую связан с текущим исследованием. Настоящее исследование оценивает этот вывод. С показателем ложного принятия (FAR) 21% и показателем ложного отклонения (FRR) 17% наилучшие результаты показали функции downdown (DD), up-down (UD) и up-up (UU).

В этой части текущего исследования проекта изучается эффективность интеграции динамики нажатия клавиш свободного текста с системой одноразовых паролей (OTP) путем генерации случайного слова из пароля и аутентификации пользователя не только на основе случайного совпадения слова, но и также аутентификация на основе поведенческих биометрических данных динамики нажатия клавиш, таким образом, достижение безопасная и сильная аутентификация и предотвращение фишинговых атак [104].

3.2 Разработка модели одноразового пароля (OTP) на основе динамики нажатия клавиш.

3.2.1 Процесс построения модели пользователя на этапе обучения по результатам одноразового пароля (OTP) на основе динамики нажатия клавиш.

Индекс Жаккара (Jaccard index) использовался для измерения процента сходства между набором образцов паролей и набором случайных образцов слов, чтобы принять решение о том, какое расстояние использовать для создания порогового значения на основе процента сходства [122].

Индекс Жаккара — это статистика, используемая для измерения сходства и разнообразия наборов выборок. Обычно он определяется соотношением двух объемов (площадей или объемов), объема пересечения, деленного на объем объединения, также называемого пересечением над объединением (IoU) [122].

был разработан Grove Karl Gilbert в 1884 году как коэффициент проверки (v) [123] и теперь часто называется критическим индексом успеха в метеорологии. Позже был независимо разработан Paul Jaccard, первоначально дав французское название модулю сообщества (modulus de la Community), и снова независимо придуман T. Tanimoto. Следовательно, в некоторых областях его также называют индексом Танимото или коэффициентом Танимото[122,123].

Сходство Жаккара — это мера сходства между двумя асимметричными бинарными векторами или, можно сказать, способ найти сходство между двумя

наборами. Это распространенная мера близости, используемая для вычисления сходства двух элементов, таких как два текстовых документа. Индекс находится в диапазоне от 0 до 1. Диапазон, близкий к 1, означает большее сходство в двух наборах данных.

Он обозначается J и также называется индексом Жаккара, коэффициентом Жаккара, несходством Жаккара и расстоянием Жаккара. Он часто используется в науке о данных и машинном обучении, таких как интеллектуальный анализ текста, электронная коммерция, система рекомендаций и т. д. рассчитывается по формуле:

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}, \tag{3.1}$$

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B|}, \tag{3.2}$$

$$J_\mu(A,B) = \frac{\mu(A \cap B)}{\mu(A \cup B)}, \tag{3.3}$$

где J – является сходство Жаккара; A, B – образцы коллекций; J_μ – является сходство Жаккара включая вероятностные меры; μ – размер измеримое пространство, $\mu(A \cup B) \neq 0$.

Расстояние Жаккара, которое измеряет различие между выборками, является дополнительным к индексу Жаккара и получается путем вычитания индекса Жаккара из 1[124, 125]. рассчитывается по формуле:

$$d_J(A,B) = 1 - J(A,B) = \frac{|A \cup B| - |A \cap B|}{|A \cup B|}, \tag{3.4}$$

$$d_\mu(A,B) = 1 - J_\mu(A,B) = \frac{\mu(A \,\Delta\, B)}{\mu(A \cup B)} = \frac{\mu(A \cup B\,) - \mu(A \cap B)}{\mu(A \cup B)}, \tag{3.5}$$

где d_J– является Расстояние Жаккара; d_μ– расстояние Жаккара для мер, включая вероятностные меры.

Метрики — это подмножество мер, определяемое четырьмя математическими свойствами, которыми они обладают: не отрицательность; тождественность неразличимых; симметрия; и неравенство треугольника. Мы можем описать каждую из них, соответственно, с помощью индекса Жаккара:

$$J(S,T) \geq 0, \tag{3.6}$$

$$J(S,T) = 1 \Leftrightarrow S = T, \quad (3.7)$$

$$J(S,T) = J(T,S), \quad (3.8)$$

$$J(S,U) \geq J(S,T) + J(T,U), \quad (3.9)$$

Легко увидеть, как индекс Жаккара удовлетворяет первым трем свойствам, поскольку его максимальные границы равны $0 \leq J(S,T) \leq 1$, а уравнение 1 не изменится, если поменять местами S и T.

3.2.2 Анализ механизма генерации случайного одноразового пароля

Случайный пароль был сгенерирован на основе символов введенного пароля с использованием трех функций**:**

1. Strlen: встроенная функция, которая возвращает длину заданной строки. Она вычисляет длину строки, включая все пробелы и специальные символы [125].
2. str_repeat: встроенная функция, которая используется для создания новой строки путем повторения заданной строки фиксированное количество раз. Она принимает строку и целое число в качестве аргументов и возвращает новую строку, которая генерируется путем повторения строки, переданной в качестве аргумента, число раз, определенное целым числом, переданным в качестве аргумента этой функции [126].
3. str_shuffle: встроенная функция, которая используется для случайного перемешивания всех символов строки, переданной функции в качестве параметра. Когда передается число, она обрабатывает число как строку и перемешивает ее. Эта функция не вносит никаких изменений в исходную строку или число, переданное ей в качестве параметра. Вместо этого она возвращает новую строку, которая является одной из возможных перестановок строки, переданной ей в качестве параметра [127].
4. Substr: возвращает часть строки, эта функция допускает 3 параметра:
   - string_name: в этом параметре мы передаем исходную строку или строку, которую необходимо вырезать или изменить.
   - start_position: это относится к позиции исходной строки, из которой необходимо извлечь часть. Здесь мы передаем целое число. Если целое

число положительное, оно относится к началу позиции в строке с начала. Если целое число отрицательное, то оно относится к началу позиции с конца строки.

– string_length_to_cut: это относится к длине части строки, которую необходимо вырезать из исходной строки. Если целое число положительное, оно относится к началу с start_position и извлечению длины с начала. Если целое число отрицательное, то оно относится к началу с start_position и извлечению длины с конца строки. Если этот параметр не передан, то функция substr() вернет строку, начиная с start_position до конца строки.

Создание матрицу пароля, который вводит пользователь, и матрицу случайных слов из самого пароля.

$$M_{password} = \begin{pmatrix} key_i \\ key_{i+1} \\ \vdots \\ key_{i+n} \end{pmatrix}_{1\times n}, \tag{3.10}$$

$$M_{random} = \begin{pmatrix} key_i \\ key_{i+1} \\ \vdots \\ key_{i+m} \end{pmatrix}_{1\times m}, \tag{3.11}$$

где $M_{password}$ – оригинальная матрица паролей; M_{random} – Случайный массив паролей; key_i – содержимое пароля представляет собой код буквы, цифру или символ; m, n – представляет собой длину размер пароля, поскольку принадлежит множеству натуральных чисел, $m, n \in N$;

Расположение матрицы случайных паролей так, чтобы уменьшить процент разброса в вычислениях между ней и основной матрицей паролей, как показано на рисунке поочередно.

Операции линейной алгебры используются с матрицами и объединяют их для облегчения процесса сравнения и упорядочивания случайной матрицы. Все элементы первой матрицы делятся на элементы второй матрицы по порядку, рассчитывается по формуле:

$$A = \sum_{i=1}^{n} \frac{M_{password_i}}{M_{random_i}}, \tag{3.12}$$

$$A = \begin{pmatrix} M_{pass_i}/M_{ran_i} & M_{pass_i}/M_{ran_{i+1}} & M_{pass_i}/M_{ran_{i+n}} \\ M_{pass_{i+1}}/M_{ran_i} & M_{pass_{i+1}}/M_{ran_{i+1}} & M_{pass_{i+1}}/M_{ran_{i+n}} \\ \vdots & \cdots & \ldots \\ M_{pass_n}/M_{ran_i} & M_{pass_n}/M_{ran_{i+1}} & M_{pass_n}/M_{ran_{i+n}} \end{pmatrix}, \tag{3.13}$$

где A – Представляет новую матрицу путем деления первой матрицы на вторую; M_{pass} – элементы основного массива паролей; M_{ran} – элементы случайного массива паролей;

Используя операции линейной алгебры, матрица А преобразуется в двоичную матрицу для облегчения процесса передачи и упорядочивания случайного пароля. Рассчитывается по формуле:

$$A` = \left(A_i + \left(-1.(A_{i)}\right)\right), \text{где } \forall A_i \ni]-\infty; 1[\cup]1; +\infty[, \tag{3.14}$$

$$A` = \begin{pmatrix} 0 & 1 & \cdots & A_n \\ 1 & 0 & \cdots & \vdots \\ 1 & 1 & \cdots & \vdots \end{pmatrix}, \tag{3.15}$$

где $A`$ – Матрица из нулевых чисел, кроме одного;

В линейной алгебре транспонирование матрицы — это оператор, который переворачивает матрицу по ее диагонали; то есть он меняет местами индексы строк и столбцов матрицы А, создавая другую матрицу, часто обозначаемую как A^T

Транспонирование матрицы было введено в 1858 году британским математиком Артуром Кэли. В случае логической матрицы, представляющей бинарное отношение R, транспонирование соответствует обратному отношению R^T. Рассчитывается по формуле:

$$A^T{}_{ij} = A_{ji}, \tag{3.16}$$

$$A^T = \begin{pmatrix} 0 & 1 & \cdots & A_n \\ 1 & 0 & \cdots & \vdots \\ A_n & A_n & \cdots & \vdots \end{pmatrix}, \tag{3.17}$$

где A^T – Матрица переноса; i –Ряды; j – Колонны;

В матрице сортировки важно то, что двоичное значение один перемещается к ближайшему значению слева в каждой строке. Рассчитывается по формуле:

$$A_{mul_{ij}} = A`.A^T = A`_{i1}.A^T{}_{1j} + A`_{i2}.A^T{}_{2j} + \cdots + A`_{in}.A^T{}_{nj} = \sum_{k=1}^{n} A`_{ik}.A^T{}_{kj}, \quad (3.18)$$

$$A_{mul} = \begin{pmatrix} A_{mul_{ij}}/_n & A_{mul_{ij+1}}/_n & A_{mul_{ij+n}}/_n \\ A_{mul_{i+1,j}}/_n & A_{mul_{i+1j+1}}/_n & A_{mul_{i+1j+n}}/_n \\ \vdots & \cdots & \ldots \\ A_{mul_{i+n,j}}/_n & A_{mul_{i+n,j+1}}/_n & A_{mul_{i+n,j+n}}/_n \end{pmatrix}, \quad (3.19)$$

где A_{mul} – Матричное умножение между матрицей переноса и нулевой матрицей; k – Колонны;

Процесс перестановки строк и столбцов, если одно из значений всей строки равно нулю, и таким образом операции перестановки выполняются между следующей строкой так, чтобы результатом в итоге стала диагональная матрица с результатом единица.

$$A_{Switching} = A_{mul_{ij}} \rightarrow A_{mul_i-1,j-1}, \text{где } A_{mul_{ij}} = 0, \quad (3.20)$$

$$A_{Switching} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ 0 & 0 & 1 & \ldots \end{pmatrix}, \quad (3.21)$$

где $A_{Switching}$– Диагональная матрица порядка для одного.

Случайная матрица упорядочивается согласно упорядочение перестановки строк и столбцов матрице из каждой строки берется место как показное на рисунке 3.2, содержащее значение 1.

$$sort_{ran} = \left(A_{Switching_i}\ A_{Switching_{i+1}}\ \ldots\ A_{Switching_{i+n}}\right), A_{Switching} \supset \{1\}, \quad (3.22)$$

где $sort_{ran}$– матрица ранжирования.

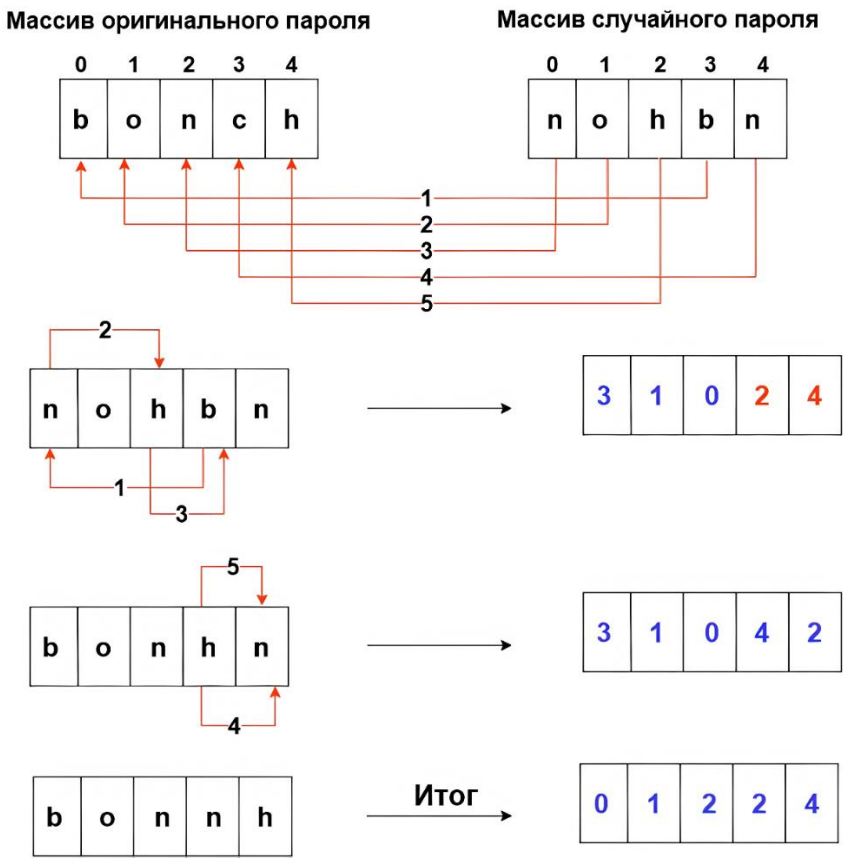


Рис. 3.2. Рандомизация случайного массива

В этой части пороговое значение рассчитывается для нажатий клавиш на первом этапе процесса тестирования и аутентификации, когда пользователь входит в систему, чтобы получить лучшее новое пороговое значение, потому что на этом этапе из пароля генерируется случайное значение. И поэтому возникает необходимость создать пороговое значение, соответствующее событию и времени. Пороговое значение будет рассчитываться для каждых двух последовательных букв.

Пороговое значение рассчитывается на основе расположений случайного массива по отношению к расположениям основной матрицы паролей, в первом случае расположения случайного массива упорядочены в соответствии с положениями массива паролей; в втором случае если местоположения случайного массива не упорядочены по отношению к местоположениям матрицы паролей; в третий случае, если местоположение повторяется в случайной матрице. Вводя его на клавиатуре, это происходит потому, что при генерации случайного слова оно не упорядочивается в соответствии с тем, что было введено. Это буквы, похожие на

введенные, но в их порядке одна из букв может быть удалена из слова, а одна из букв может повторена, становясь таким образом словом, похожим на то, что было введено пользователем.

Пароль представлен в формате массива по формуле:

$$time_{password} = \begin{pmatrix} time_{key_i} - time_{key_{i+1}} \\ time_{key_{i+1}} - time_{key_{i+2}} \\ \vdots \\ time_{key_{i+n}} - time_{key_{i+n+1}} \end{pmatrix}, \quad (3.23)$$

где $time_{password}$ – массив пароля; $time_{key_i}$ – Значение временной метки для каждой нажатой и отпущенной клавиши; n – размер пароля;

Случаи рандомного расположения клавши:

– в первом случае расположения случайного массива упорядочены в соответствии с положениями массива паролей, если клавиш и местоположение одинаковы в двух массивах по формуле массива, как показное на рисунке 3.3 часть первая.

$$time_{sort_{ran_1}} = \begin{pmatrix} time_{password_i} - time_{password_{i+1}} \\ time_{password_{i+1}} - time_{password_{i+2}} \\ \vdots \\ time_{password_{i+n}} - time_{password_{i+n+1}} \end{pmatrix}, \quad (3.24)$$

где $time_{sort_{ran\_1}}$ – первый случай вычисление ожидаемого значения расстояния для клавши;

– в втором случае клавиш повторяется последовательно в случайном массиве, пропорциональном расположению первого клавиша в массиве исходный пароль, как показное на рисунке 3.3 часть вторая.

$$time_{sort_{ran_2}} = \begin{pmatrix} time_{sort_{ran_{i-1}}} - time_{sort_{ran_i}} \\ \vdots \\ time_{sort_{ran_{i-n}}} - time_{sort_{ran_n}} \end{pmatrix}, \quad (3.25)$$

где $time_{sort_{ran\_2}}$ – второй случай вычисление ожидаемого значения расстояния для клавши;

– в третий случае расположение клавиш случайного массива не в порядке и несовместимо относительно расположено исходной матрицы паролей,

как показное на рисунке 3.3 часть тертая.

$$time_{sort_{ran_3}} = \begin{pmatrix} (time_{password_i} - time_{password_{i+1}}) - time_{password_{i+2}} \\ \vdots \\ (time_{password_{i+n}} - time_{password_{i+n+1}}) time_{password_{i+n+2}} \end{pmatrix}, \quad (3.26)$$

где $time_{sort_{ran\_3}}$ – третий случай вычисление ожидаемого значения расстояния для клавши;

Пороговое значение для случайного пароля зависит от трех случаев расчета ожидаемого расстояния от исходного пароля

$$threshould_{ran} = \sqrt{\frac{\sum_{i=1}^{n} time_{sort_{ran-1}} + time_{sort_{ran-2}} + time_{sort_{ran-3}}}{\mu}}, \quad (3.27)$$

где $threshould_{ran}$ – пороговое значение Случайный пароль; μ – Случайная длина пароля.

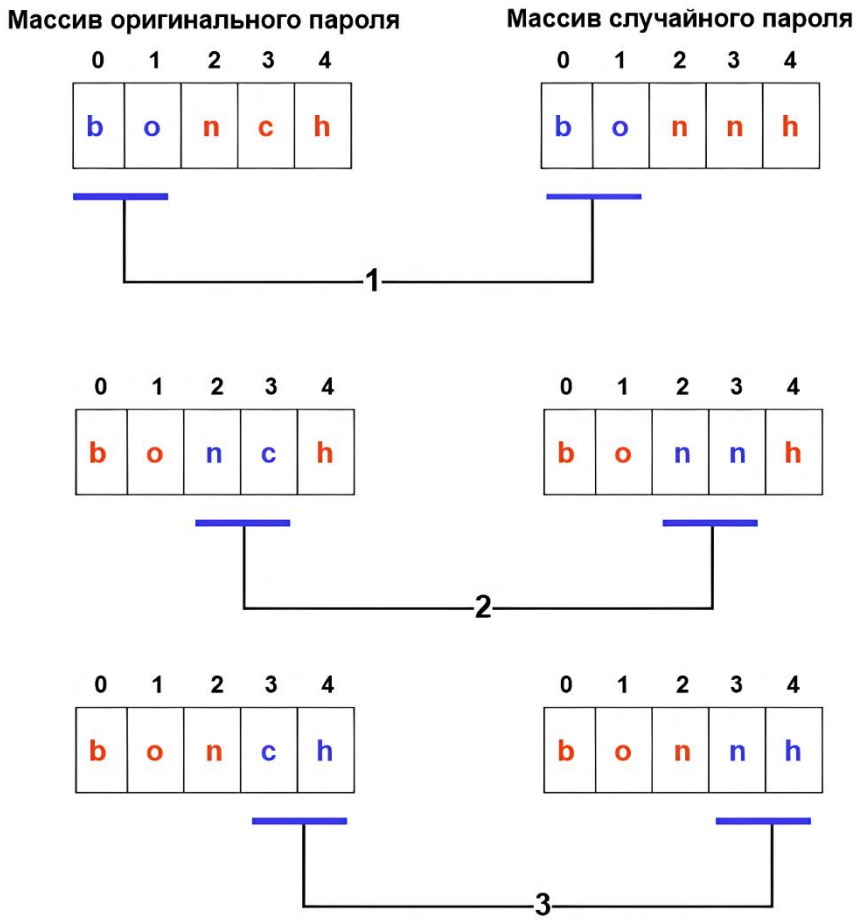


Рис.33. Случаи рандомного расположения клавши

3.2.3 Рабочая среда Этап обучения

Сначала проверяется успешность этап тестирование динамики мыши, в случае неудачи исследование повторяется снова, а в случае успеха начинается фаза обучения система OTP как показное на рисунке 3.4.

1. Из файла с расширением TXT извлекаются характеристики динамики нажатия клавиш, которые пользователь ввел его пароль на первом этапе тестирования. После этого значение временной метки для каждого клавши вычисляется и сохраняется в файле с расширением CSV.
2. Сгенерирование случайного однорукого пароля из исходного пароля и отправление его пользователю по электронной почте и одновременно сохранение его в базе данных каждого пользователя.
3. Реализация алгоритма округления случайного пароля путем создания двух матриц для исходного пароля и случайного пароля путем выполнения операций над матрицей для Рассчитывание невырожденной матрица случайного пароля.
4. Этап расчета порогового значения для каждого местоположения в матрице осуществляется путем расчета экземпляров случайных положений матрицы.

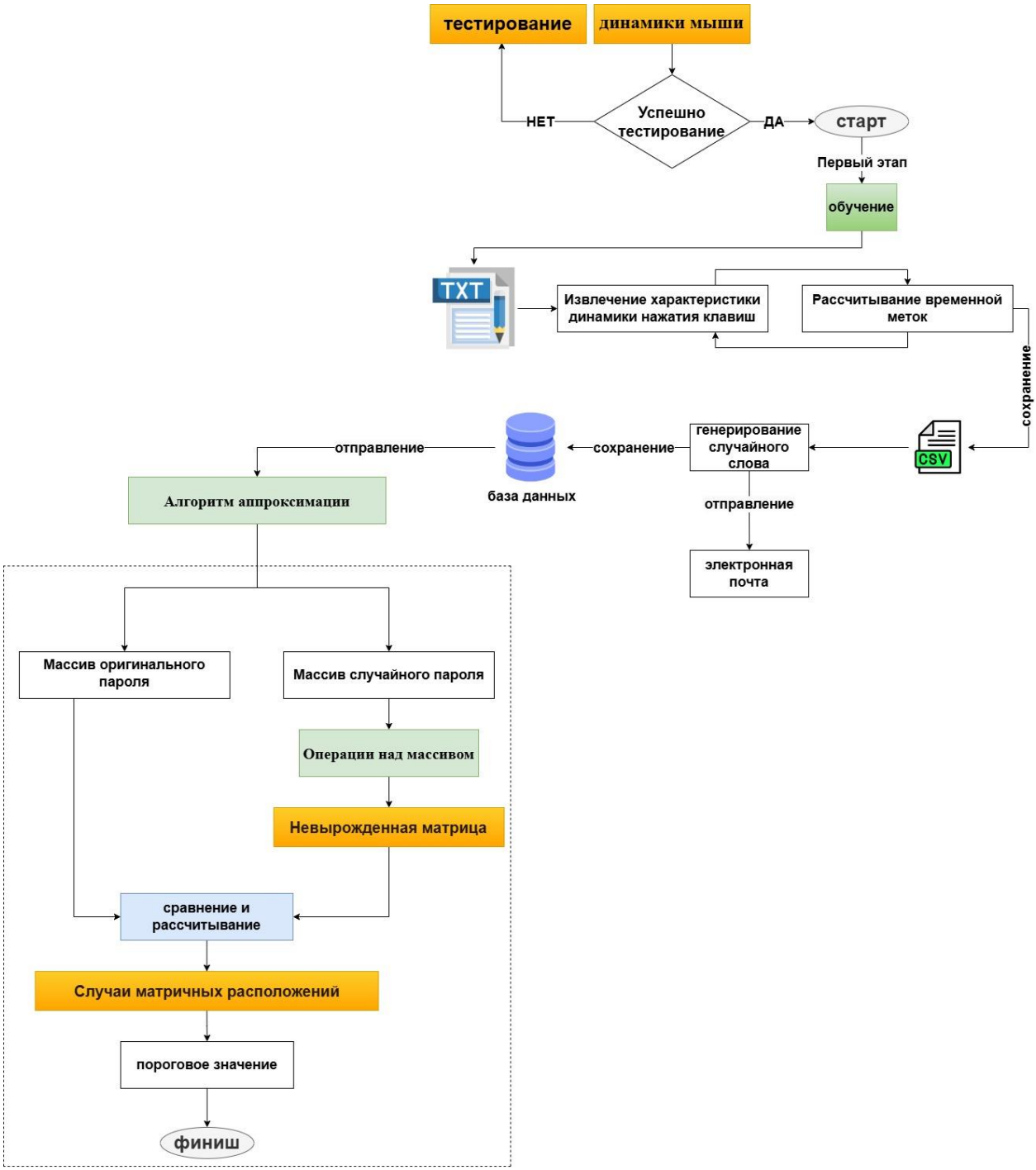


Рис. 3.4. Блок-схема этапа обучения OTP

3.3 Процесс подтверждения модели пользователя на этапе тестирования по результатам одноразового пароля (OTP) на основе динамики нажатия клавиш.

Этап тестирования и аутентификация зависят от индекса сходства и расстояния Жаккара между двумя матрицами.

В исследовательском проекте было показано, что процент сходства очень важен и чувствителен для расчета расстояния между буквами, которые пользователь набирает на клавиатуре, и, следовательно, чем выше процент сходства, тем лучше для пользователя, потому что он может написать случайный однообразного пароль, так же, как он написал свой собственный пароль, и чем ниже процент сходства, тем больше будет повторов букв или удаления одной из букв, и таким образом оно считается новым словом и пользователем должен переписать пропорционально размеру введенного им пространства пароля и пороговому значению.

Хорошая мера должна быть способна различать немного отличающиеся наборы шаблонов и очень отличающиеся наборы шаблонов. Индекс Жаккара обладает несколькими свойствами, которые делают это возможным, а также свойствами, которые усиливают универсальность меры [122, 124].

Коэффициент сходства Жаккара измеряется с помощью характеристик динамики нажатия клавиш. Рассчитывается по формуле:

$$J\left(M_{password}, M_{random}\right) = \frac{\left|M_{password} \cap M_{random}\right|}{\left|M_{password} \cup M_{random}\right|}, \tag{3.28}$$

$$\frac{key_{pass_i}, key_{pass_{i+1}}, \dots, key_{pass_{i+n}} \cap \ key_{ran_i}, key_{ran_{i+1}}, \dots, key_{ran_{i+n}}}{key_{pass_i}, key_{pass_{i+1}}, \dots, key_{pass_{i+n}} \cap \ key_{ran_i}, key_{ran_{i+1}}, \dots, key_{ran_{i+n}}}, \tag{3.29}$$

где $J\left(M_{password}, M_{random}\right)$ – Мера сходства Жаккара;

Использование расстояния Жаккара с динамикой нажатий клавиш показывает нам, что чем меньше расстояние, тем более похожи две матрицы, и наоборот. Рассчитывается по формуле:

$$d_J\left(M_{password}, M_{random}\right) = 1 - J\left(M_{password}, M_{random}\right), \tag{3.30}$$

$$\frac{\left|M_{password} \cup M_{random}\right| - \left|M_{password} \cap M_{random}\right|}{\left|M_{password} \cup M_{random}\right|}, \tag{3.31}$$

$$\frac{\sum_{i=1}^{n} |\, key_{pass_i} \cup key_{ran_i}| - |key_{pass_i} \cap key_{ran_i}|}{\sum_{i=1}^{n} |\, key_{pass_i} \cup key_{ran_i}|}, \tag{3.32}$$

где $d_J\left(M_{password}, M_{random}\right)$– расстояние Жаккара;

С помощью индекса Жаккара рассчитывается процент сходства между двумя матрицами, чтобы найти расстояние Жаккара для сходства, на основе которого выбирается тип расстояния, которое будет использоваться на этапе проверки и аутентификации.

Евклидово расстояние и Манхэттенское расстояние используются для проверки значения введенного случайного слова и его соответствия расстоянию исходного пароля, поэтому, если расстояние Жаккара больше 0.5, используется Манхэттенское расстояние. поскольку расстояние Манхэттена измеряет путь вдоль линий сети, а это означает, что расстояние Манхэттена измеряет самое длинное расстояние, к которому можно получить доступ между двумя точками, чтобы избежать ложный коэффициент отклонения. Однако если расстояние Жаккара меньше или равно 0.5, будет использоваться евклидово расстояние, поскольку евклидово расстояние измеряет прямой путь вдоль линий сетки, то есть евклидово расстояние измеряет кратчайшее расстояние, которого можно достичь между двумя точками, чтобы избежать ложный коэффициент принятия.

$$dis_{ran} = \begin{cases} \text{Евклидовое}, & 0.5 \le d_J\left(M_{password}, M_{random}\right) \\ \text{Манхэттенское}, & 0.5 > d_J\left(M_{password}, M_{random}\right) \end{cases}, \quad (3.33)$$

Евклидово расстояние случайной матрицы на основе расстояния Жаккара, если оно меньше или равно 0.5, рассчитывается по формуле.

$$d_E = \sqrt{\left(key_{r_i} - key_{r_{i+1}}\right)^2 + \left(key_{r_{i+2}} - key_{r_{i+3}}\right)^2 + \cdots + \left(key_{r_{i+n}} - key_{r_{i+n}}\right)^2}, \quad (3.34)$$

$$d_E = \sqrt[2]{\sum_{\boldsymbol{i}}^{\boldsymbol{n}} (key_{r_i} - key_{r_i})^2}, \quad (3.35)$$

где d_E – евклидово расстояние; key_r– клавиша, которая была нажата и отпущена здесь, представляет значение ее временной метки.

$$d_M = \left|key_{r_i} - key_{r_{i+1}}\right| + \left|key_{r_{i+2}} - key_{r_{i+3}}\right| + \cdots + \left|key_{r_n} - key_{r_{i+n}}\right|, \quad (3.36)$$

$$d_M = \sum_{i=1}^{n} \left| key_{r_i} - key_{r_i} \right|, \quad (3.37)$$

где d_E – манхэттенское расстояние;

Случайное ключевое слово отправляется пользователю по электронной почте, как показное на рисунке 3.5, через библиотеку Миллера (PHPMailer).

PHPMailer — это библиотека кода для безопасной и простой отправки (транспортировки) писем через PHP-код с веб-сервера (MUA на сервер MSA) [129].

Отправка писем напрямую через PHP-код требует высокого уровня знаний стандартов протокола SMTP и связанных с ними проблем (таких как возврат каретки), и уязвимостей в отношении инъекций писем для рассылки спама. С 2001 года PHPMailer является одним из популярных решений для этих проблем на PHP [130].

Преимущество PHPMailer

1. Вероятно, самый популярный в мире код для отправки электронной почты из PHP;
2. Используется многими проектами с открытым исходным кодом: WordPress, Drupal, 1CRM, SugarCRM, Yii, Joomla! и многими другими;
3. Интегрированная поддержка SMTP — отправка без локального почтового сервера;
4. Отправка писем с несколькими адресами «Кому», «Копия», «Скрытая копия» и «Ответить»;
5. Составные/альтернативные письма для почтовых клиентов, которые не читают электронную почту в формате HTML;
6. Добавление вложений, в том числе встроенных;
7. Поддержка содержимого UTF-8 и кодировок 8bit, base64, binary и quote-printable;
8. Аутентификация SMTP с механизмами LOGIN, PLAIN, CRAM-MD5 и XOAUTH2 через транспорты SMTPS и SMTP+STARTTLS;
9. Автоматическая проверка адресов электронной почты;
10. Защита от атак с внедрением заголовков;

11. Сообщения об ошибках на более чем 50 языках;
12. Поддержка подписей DKIM и S/MIME;
13. Совместимость с PHP 5.5 и более поздними версиями, включая PHP 8.2;
14. Пространство имен для предотвращения конфликтов имен.

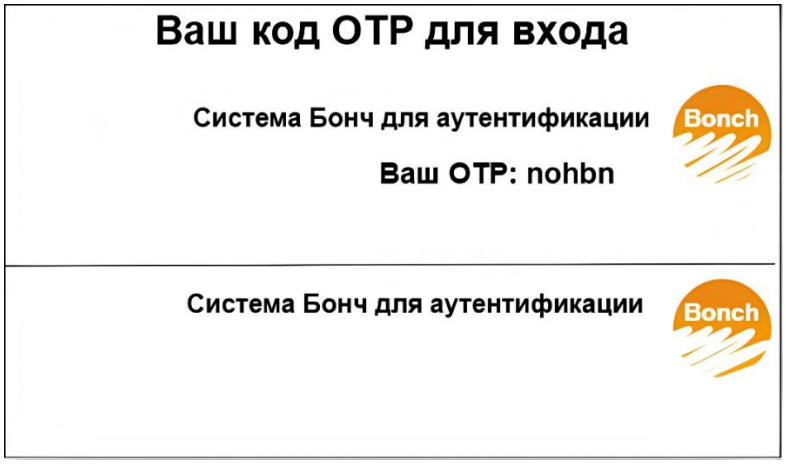


Рис. 3.5. Отправка случайного слова по электронной почте

3.3.1 Рабочая среда этапа тестирования

1. На этапе тестировании пользователь получает случайный пароль по электронной почте, после чего вводит случайное слово в назначенное текстовое поле как показное на рисунке 3.7.
2. Извлечение характеристик динамики нажатия клавиш случайного слова, вычисление значения временной метки для каждой клавиши и сохранение его в файле с расширением TXT как показное на рисунке 3.6.
3. Индекс сходства Жаккара и расстояние Жаккара рассчитываются для определения типа расстояния измерения, которое будет использоваться для проведения процесса исследования и сравнения, либо с использованием Манхэттенского расстояния, либо Евклидова расстояния.
4. Сначала сравнивается совпадение случайного пароля, введенного пользователем, со случайным паролем, хранящимся в базе данных. После этого выполняется сравнение, чтобы определить, меньше или равно значение расстояния пороговому значению то пользователь действителен и

ему разрешен вход в систему, в противном случае пользователь считается недействительным и сеанс завершается.

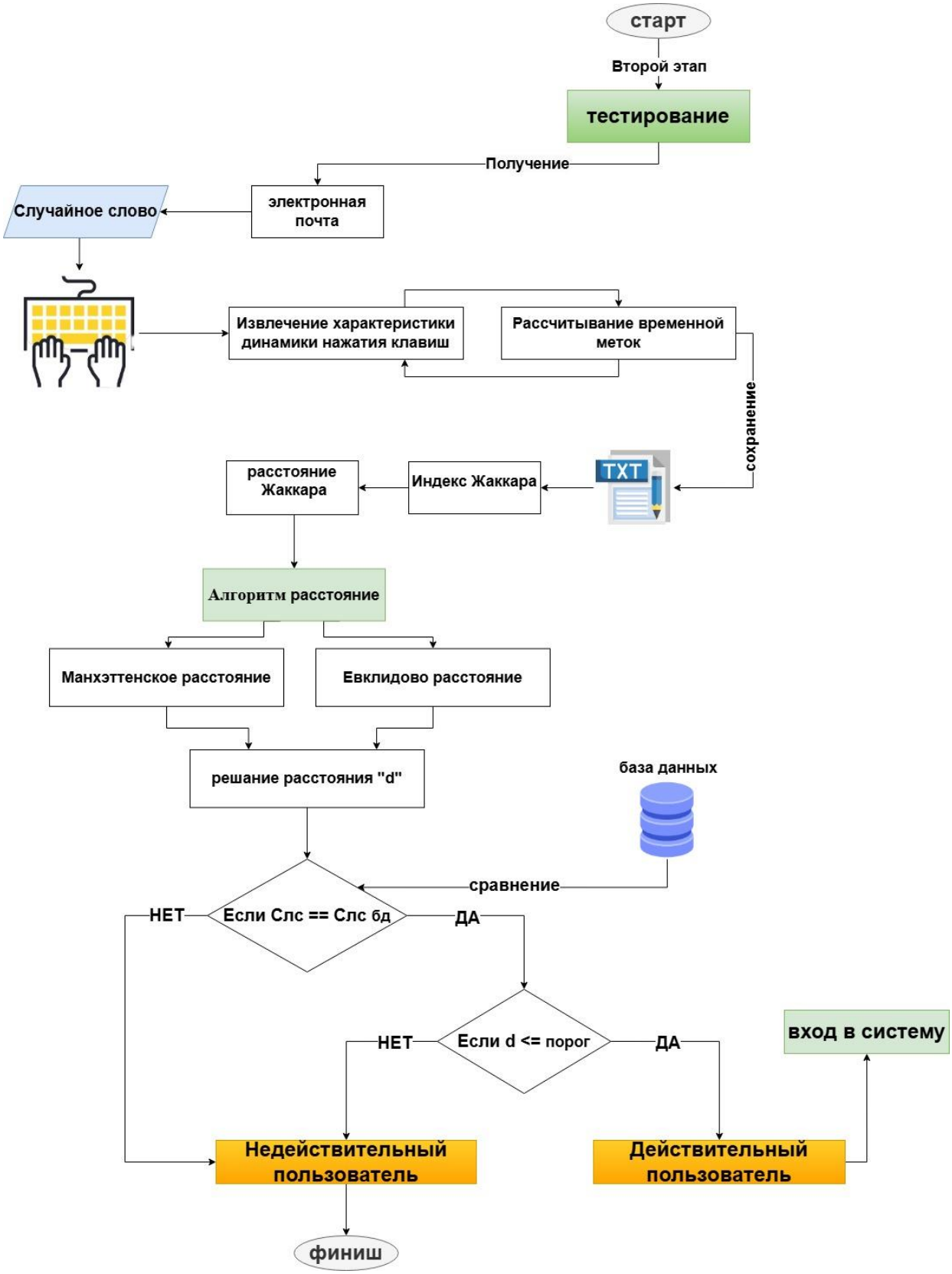


Рис. 3.6. Блок-схема этапа тестирования OTP

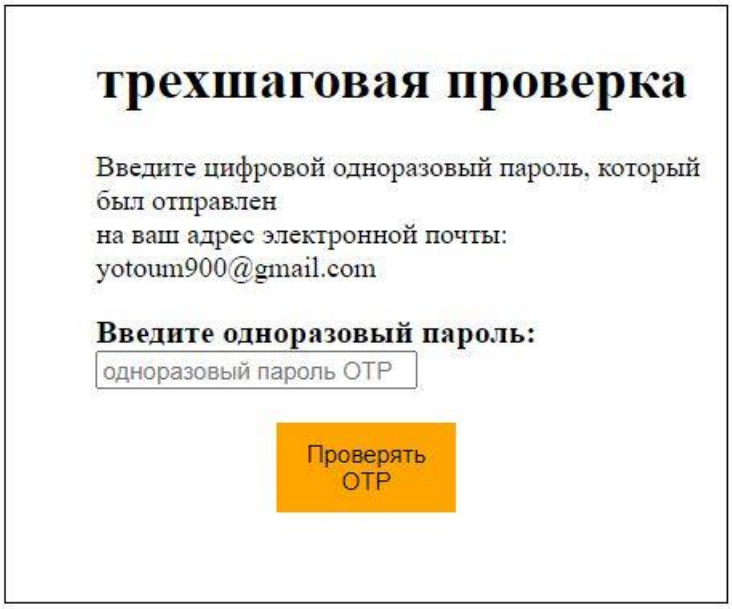


Рис. 3.7. Внедрение этапа тестирования OTP

Выводы по 3**-й главе**

В данном разделе проводились исследование и разработка модели многофакторной аутентификации на основе поведенческой биометрии нажатий клавиш с использованием метода одноразовых паролей (OTP).

Этот подход не использовался в предыдущих исследованиях и показал многообещающие результаты в повышении уровня защиты и снижении фишинговых атак.

В результате проведенного исследования были достигнуты следующие результаты:

- Предлагается подход к генерации случайного пароля из мастер-пароля, введенного пользователем на этапе проверки в первой части. В связи с этим, повышается точность выбора шаблона биометрических данных для проверки.
- Для снижения уровня ложного принятия предлагается матричный подход, основанный на случайных позициях ключей относительно ключевого слова, что позволяет улучшить выбор точного порогового значения.
- Предлагается подход к проверке нажатия клавиш, основанный на расстоянии сходства Жаккара, позволяющий выбор расстояния

Манхэттена или Евклидово расстояния для выполнения процесса проверки. Повысилось качество аутентификации пользователей в среднем на 8% по сравнению с используемой в существующих работах стандартизацией признаков.

- Была проведена серия экспериментов, в результате которой было подтверждено высокое качество работы предложенных алгоритмов. В результате использования предложенной комбинации алгоритмов удалось достичь качества распознавания порядка 0.9316 ROC AUC, что значительно превосходит качество аутентификации пользователей при использовании методов, рассмотренных в существующих научных работах.

ГЛАВА 4. Динамическое Непрерывная аутентификации пользователей и субъектов доступа для веб-приложения в процессе работы

4.1 Описание основных принципов работы непрерывной аутентификации

Одним из заметных подходов, набирающих популярность в последние годы, является непрерывная аутентификация, парадигма которой выходит за рамки традиционных методов аутентификации. Непрерывная аутентификация стремится установить динамичный и непрерывный процесс проверки, постоянно отслеживая поведение пользователя, чтобы гарантировать, что предполагаемый пользователь получает доступ к системе. Таким образом, непрерывная аутентификация использует модель поведения пользователя при взаимодействии с системой для подтверждения его личности; следовательно, в результате получается более целостный и безопасный подход к безопасности системы по сравнению с традиционными методами. Это позволяет более надежной системе гарантировать, что любой пользователь, получающий доступ к защищенной системе, является предполагаемым пользователем этой системы, гарантируя, что конфиденциальная информация будет сохранена в безопасности [5].

Непрерывная аутентификация, также известная как активная аутентификация, была введена в 2012 году для обработки новых методов проверки личности пользователя [5].

Биометрия поведения на основе программного обеспечения была основным акцентом для сбора данных сеанса и выяснения того, использовал ли реальный пользователь систему в любой конкретный момент.

В непрерывной аутентификации динамика мыши является многообещающим подходом для того, чтобы идентифицировать аномалии. В отличие от других биометрических методов, таких как голос или отпечатки пальцев, динамика мыши не требует специальных устройств для сбора данных. Более того, после первоначальной аутентификации с использованием паролей или других учетных данных процедура обнаружения может быть плавно включена в обычную компьютерную деятельность пользователей, предлагая не интрузивный метод

контроля личности или постоянной аутентификации. Тщательный анализ динамики мыши дан Ямпольским и др. [131].

Динамика мыши — термин для описания использования мыши пользователем. Метод предложен в качестве анализа поведенческой биометрии для аутентификации пользователя с свободным интерфейсом [132–135]. Используя незначительные, но заметные изменения в мелкой моторике человека, динамика мыши позволяет однозначно идентифицировать пользователей. Кроме того, она требует меньше личной информации от пользователя, чем другие поведенческие биометрические показатели [133]. В ранее заявленных веб-средах динамика мыши является наиболее простым и распространенным способом сбора данных. Таким образом, она является сильным претендентом на будущие методы онлайн-аутентификации пользователей на основе поведения [134, 135]. Помимо недавней известности динамики мыши, алгоритмы машинного и глубокого обучения стали широко использоваться в условиях с большим количеством данных и часто превосходят традиционные статистические методы в ряде этих проблемных областей [136–142].

Физическая и поведенческая биометрия — это две категории, на которые делится биометрическая аутентификация [143]. Физическая биометрия использует голосовые сканеры [144], радужную оболочку глаза [145] и отпечатки пальцев [146] для аутентификации людей, используя их отличительные физические характеристики. Поведенческая биометрия, которая включает в себя динамику прикосновений [147] и динамику нажатия клавиш [148, 149], также работает на основе предпосылки, что поведение человека в целом для определенных задач достаточно отличается, чтобы его можно было использовать для аутентификации пользователя. По сравнению с физической биометрией, поведенческая биометрия для идентификации привлекла особое внимание, поскольку она менее навязчива, имеет более широкий спектр применения и не требует внешних датчиков. Кроме того, было показано, что динамика мыши является непрерывным, ненавязчивым и легким подходом к динамической аутентификации пользователя [34, 150–152].

Динамика мыши широко использовалась для аутентификации личности в течение последних десяти лет благодаря своим преимуществам, включающим в себя непрерывный мониторинг пользователя без беспокойства пользователей и отсутствие дополнительного оборудования [53].

Каждая система или метод аутентификации имеет характеристики, которые отличают ее от других. В книге (Advances in User Authentication) объясняются характеристики хорошей системы непрерывной аутентификации. Решения непрерывной аутентификации, в целом, предлагают дополнительную степень безопасности по сравнению со статическими методами аутентификации. Реализация этого метода имеет решающее значение в высокозащищенных системах, которые включают постоянную проверку личности пользователя. Как показное в таблице 5 характеристики являются существенными для хорошей системы непрерывной аутентификации [5], и показное в таблице 6 Преимущества и разница непрерывной и статической аутентификации.

Таблица 5 — Характеристики системы непрерывной аутентификации

Характеристика	**Описании**
Независимость от компьютерной системы	Метод непрерывного подтверждения личности пользователя должен работать с различными программами и не зависеть от системного оборудования. Постоянное использование анализа движения мыши и нажатия клавиш для аутентификации пользователя должно быть совместимо с различными приложениями и аппаратными модальностями.
Ненавязчивость	личности пользователей должны легко и ненавязчиво проверяться непрерывной системой. Пароли и PIN-коды являются примерами интрузивных подходов, которые требуют больше времени и усилий пользователя, что значительно повышает стоимость аутентификации. В результате пользователь может заниматься своими повседневными делами без участия системы. Система может немедленно заблокировать доступ к системе и запросить проверку подлинности, если обнаружит самозванца.
Более быстрый ответ	необходимо более быстрое выполнение модальностей аутентификации, поскольку

	непрерывная аутентификация принимает решения на ходу. Технология, требующая небольшого обучения пользователя, должна обеспечивать решения для непрерывной проверки личности, которые являются быстрыми, надежными и простыми в использовании.
Поведенческие атрибуты	поведенческая биометрия необходима для того, чтобы система непрерывной аутентификации распознавала людей в изменяющейся во времени обстановке. Эти модальности используются в непрерывной аутентификации, поскольку они недороги, легко интегрируются с текущей системой и развертываются без особых усилий. Поведенческие атрибуты работают пассивно в системе и собирают данные об использовании пользователем (через распознавание клавиатуры и движения мыши).

Таблица 6— Преимущества и разница непрерывной и статической аутентификации

Параметр	**Непрерывная аутентификация**	**Статическая аутентификация**	**Доказательство**
Повышенная безопасность	Успешная система непрерывной аутентификации учитывает человеческие и экологические элементы в дополнение к надежности. Благодаря постоянному мониторингу поведения пользователя и подтверждению того, что в систему в течение сеанса входит одно и то же лицо.	Однократные аутентифицированные входы похожи на регистрацию кого-либо на входе в том, что сложно следить за их поведением после того, как они получат доступ к системе. Современные контексты, такие как общие учетные записи, удаленные рабочие наборы и многотерминальный доступ, слишком динамичны для управления статическими методами	По данным отчета Verizon Data Breach Investigations Report 2024, 88% нарушений, связанных со взломом, вызваны кражей учетных данных, что на 180% больше, чем в предыдущем году [154].

	Непрерывная аутентификация добавляет дополнительную степень защиты. Помимо предотвращения нежелательного доступа, это может помочь защититься от различных угроз, включая phishing и stuffing.	аутентификации. Используя учетные данные для входа или имитацию активности пользователя, хакеры обнаружили методы обхода мониторинга бездействия. Следовательно, даже если мониторинг бездействия имеет свои преимущества, его следует использовать в сочетании с другими мерами безопасности, чтобы обеспечить комплексное решение по обеспечению безопасности.	
Снижение риска мошенничества	Можно выявить любые нарушения или изменения в поведении пользователя, которые могут указывать на возможную попытку мошенничества, регулярно наблюдая за его действиями. Таким образом, можно снизить вероятность мошенничества с идентификацией и финансовым	Поскольку статическая аутентификация зависит от постоянного пароля или других неизменных учетных данных, которые могут быть украдены путем фишинга, кражи или других вредоносных попыток, она уязвима для мошенничества.	• По данным Центра ресурсов по краже личных данных, в 2024 году было зафиксировано 2242 утечки данных, в результате которых было раскрыто более 940 миллионов записей. Это составляет 70 процентов от общего числа на конец 2024 года, что делает рекордно высокое количество компрометаций маловероятным в 2025 году [155].

	мошенничество м.		• ОСТИН, Техас 2024 — 3,1 млрд долларов убытков из-за мошенничества. Таковы общие потери, подсчитанные в последнем отчете Ассоциации сертифицированны х экспертов по мошенничеству (ACFE) «Профессиональное мошенничество 2024: отчет для наций». На основании своих выводов ACFE подсчитала, что организации ежегодно теряют 5% дохода из-за мошенничества.[15 6]
Улучшенный пользовательск ий опыт	Непрерывная аутентификаци я устраняет необходимость во внешних устройствах, таких как многофакторна я аутентификаци я, которые могут быть трудоемкими и отвлекающими для пользователя.	Пользовательский опыт неудовлетворительн ый. Пользователям часто предлагается пройти MFA только один раз, в начале сеанса. Но иногда, в том числе при получении доступа к личным данным или выполнении важных	• Согласно отчету LastPass, среднестатистическ ий сотрудник компании должен помнить 191 пароль, что приводит к усталости от паролей и разочарованию.[157] • Опрос, проведенный PYMNTS и Entersekt, показал,

	Она обеспечивает более плавный и интуитивный опыт, анализируя поведение пользователя без вмешательства.		что 85% потребителей предпочитают биометрическую аутентификацию паролям, называя удобство основной причиной.[158]

Было предложено много подходов к непрерывной аутентификации. Две широкие категории: аутентификация на основе движения и биометрическая аутентификация. Биометрическая аутентификация использует определенные биометрические идентификаторы для непрерывной аутентификации личности пользователя. В этом подходе предыдущие ученые исследовали распознавание речи для проверки личности пользователя (Thomas & Preetha Mathew, 2023) [159]. Несмотря на то, что этот подход показал многообещающие результаты, поскольку он основан на слуховой информации, которую пользователь может не часто предоставлять при использовании системы, он нелегко реализуем. Другой подход использовал распознавание лиц в качестве биометрического идентификатора для аутентификации личности пользователя (Zhang et al., 2016) [160]. Хотя этот подход широко используется в безопасности мобильных систем, его нелегко переносить на настольные приложения. Кроме того, распознавание лиц основано на визуальных данных, которые могут меняться в зависимости от среды пользователя, что часто приводит к менее благоприятным прогнозам (Smith-Creasey et al.,2018) [161]. Некоторые другие подходы включают отслеживание глаз и мониторинг сердечного ритма; Оба метода сталкиваются с теми же проблемами, что и другие методы биометрической аутентификации, поскольку они либо используют нестандартное оборудование, либо слишком зависят от среды пользователя, чтобы их можно было использовать в динамических условиях (Jacob & Karn, 2003; Cheung & Vhaduri, 2020)[162, 163].

Для аутентификации на основе движений основными используемыми методами являются динамика нажатия клавиш, динамика касания и динамика мыши (Sayed et al., 2013) [164]. Общей темой среди этих предпочтительных модальностей непрерывной аутентификации является то, что они изучают модели поведения, пока пользователь взаимодействует с системой, и используют их для определения легитимности доступа пользователя к системе. Однако в сфере непрерывной аутентификации динамика мыши выступает как многообещающее направление для повышения безопасности (Quraishi & Bedi, 2022) [165]. Динамика мыши включает анализ уникальных моделей и характеристик пользователей в их движениях мыши. К ним относятся такие параметры, как скорость, траектория, тип действия мыши (перетаскивание, отпускание или щелчок) и т. д. Используя модель взаимодействия пользователя при использовании системы с мышью, можно разработать систему, которая не только аутентифицирует пользователей во время первоначального входа в систему (или вскоре после этого), но и непрерывно проверяет их личность на протяжении всего сеанса. Значимость включения динамики мыши в структуру аутентификации заключается в ее потенциале предлагать неинтрузивные, но высокоэффективные средства идентификации пользователей. В отличие от традиционных методов, которые полагаются на явные действия, такие как ввод пароля, непрерывная аутентификация с использованием динамики мыши может плавно адаптироваться к динамическому взаимодействию пользователя с системой (Chen et al., 2019) [166]. Это не только улучшает пользовательский опыт, но и обеспечивает дополнительный уровень безопасности за счет постоянной проверки личности пользователя на основе его уникальных поведенческих моделей (Mondal & Bours, 2015) [167]. Кроме того, сбор данных мыши пользователя полностью неинтрузивный, поскольку данные мыши не могут содержать конфиденциальную информацию, которая может поставить под угрозу конфиденциальность пользователя.

Непрерывная биометрическая аутентификация уже была предложена для множества сценариев, таких как безопасность кабины самолета, транспортного средства, настольных компьютеров и мобильных устройств (Carrillo, 2003; Crouse,

Han, Chandra, Barbello, & Jain, 2015; Derman & Salah, 2018; Janakiraman, Kumar, Zhang, & Sim, 2005) [168 - 171], доверенных автономных систем (Wang, Abbass, & Hu, 2016)[172], телемедицины (Agrafioti, Bui, & Hatzinakos, 2012) [173], а также онлайн-экзаменов и электронного обучения (Flior & Kowalski, 2010)[174]. Существуют и другие исследования, в которых анализируется биометрическая непрерывная аутентификация. Аль Абдулвахид, Кларк, Стенгель, Фернелл и Райх (2016) [175] фокусируются на том, как проектировать общую систему, Стилиос, Тану, Андроулидакис и Зайцева (2016) [176] фокусируются на поведенческой биометрии в целом, а Махфуз, Махмуд и Элдин (2017) [177] — на поведенческой биометрии смартфонов.

Непрерывная аутентификация на основе поведенческой биометрии привлекательна, поскольку биометрические образцы для распознавания получаются с помощью традиционных устройств ввода (таких как клавиатура, мышь и сенсорный экран) (Stanic, 2013) [178]. Они не требуют дополнительного датчика и не ограничивают действия пользователя при аутентификации (Shen, Cai, & Guan, 2012) [179].

Как правило, биометрическое распознавание на основе динамики клавиатуры, мыши и касания использует определенные наборы действий, известных как события. Например, события мыши включают перемещение курсора, наведение и щелчок, и перетаскивание; события касания включают сжатие, скольжение и касание экрана; события клавиатуры включают нажатие определенных последовательностей клавиш или временной интервал между нажатием и отпусканием клавиши. Это поведенческие биометрические характеристики, которые используют действия человека для подтверждения его личности. Поскольку поведение человека зависит от множества экологических, эмоциональных и биологических факторов, они, следовательно, более непредсказуемы, чем физические биометрические данные, такие как внешность или отпечатки пальцев. Например, выбор периферийного устройства может привести к расхождениям в биометрических профилях пользователей при использовании динамики мыши и клавиатуры (Фридман и др., 2015) [180].

Процесс получения поведенческих характеристик по своей сути является временным, как и биосигналы. С другой стороны, аутентификация невозможна, если выбранная характеристика не используется. Если пользователь занимается деятельностью, которая не влечет за собой ни одного из событий (например, чтение или просмотр видео), у системы нет абсолютно никакой информации для оценки ее общего уровня безопасности. Это один из факторов, которые способствуют использованию этих характеристик в мультимодальных системах (Fridman et al., 2015; Mondal & Bours, 2015) [180, 181].

Можно оценить лингвистический стиль написанного текста, если система способна определить, какие клавиши нажимаются во время непрерывной аутентификации пользователя, наблюдая за динамикой нажатия клавиш пользователями. Это называется стилометрией, и она также извлекает выгоду из динамики нажатия клавиш. Тем не менее, пользователи могут решить не использовать ее из опасений, что они непреднамеренно разгласят личную информацию. Использование шаблонов использования мобильных устройств для проверки людей — еще одна новая разработка. Чтобы выяснить, является ли текущий пользователь тем же самым, который был аутентифицирован ранее, Сентено, Гуан и ван Мурсел (2018) и Валеро и др. (2018) [182, 183] предлагают использовать данные об использовании мобильных устройств, такие как акселерометр, гироскоп, магнитометр и статистика, собранная при взаимодействии с приложениями.

В таблице 7 представлена сводная информация о работе непрерывной аутентификации с использованием динамики мыши и касания, включая данные об опыте каждого метода, его эффективности и количестве действий или движений мыши, необходимых для выполнения непрерывной аутентификации.

Было показано, что большинство методов, используемых для проведения непрерывной аутентификации на основе поведенческих биометрических измерений движения мыши или нажатия клавиш, применяются с использованием систем машинного обучения и нейронных сетей. Они дают многообещающие результаты, при этом эксперимент проводится впервые. Однако в долгосрочной

перспективе результаты могут отличаться и стать противоположными, т.к. нейронная сеть может ошибиться в прогнозировании движения пользователя. В исследовательском проекте мы предлагаем систему непрерывной аутентификации, которая аутентифицирует пользователя и гарантирует, что достоверность сеанса основана на движении мыши. При каждой остановке движения проводится проверка. Проверка проводится на основе измерения расстояний и ускорения, а также с использованием расстояния Левенштейна для определения порогового значения. каждое событие, произошедшее на веб-странице.

Таблица 7— известные исследования о непрерывной аутентификации

Работа	**метод сопоставления**	**События аутинтификации**	**Пользователи**	**Производительность**
[179]	SVM	1000 событий	28	FAR= 2.75% FRR= 3.39%
[184]	Random Forest	30 событий	25	ERR=8.53%
[185]	Bayesian network	7 событий	40	FAR= 4.66% FRR= 0.13%
[186]	SVM	11 событий	41	ERR=2%
[187]	SVM	1 минута	200	ERR=5%
[182]	CNN + SVM	0.5 секунд	100	Accuracy=96.3%

4.2 Разработка модели непрерывная аутентификация на основе динамики мыши.

Используется характеристики, отражающие подробные динамические процессы поведения мыши, относится к динамическому движению как показное в таблице 7.

Таблица 7 — Характеристики динамики мыши при непрерывной аутентификации

Действия мыши	**Описание**	**Формула**
Затишье	Ситуация без каких-либо действий мыши	*x=10 миллисекунд* $ms = \begin{cases} 1, & m_{du} - m_{du_x} == 0 \\ 1, & m_{mov} - m_{mov_x} == 0 \end{cases}$
Скорость движения мыши	Это скорость движения мыши при перемещении между элементами по горизонтальной	$msp_{mov} = \frac{m_{mov}}{time_i - time_{i-1}}$

	линии, вертикальной линии или кривой.	
Дистанция движения мыши	Движение между точками без нажатия	dis_{mov}=19.$\left\|m_{mov_i} - m_{mov_{i+1}}\right\|$
Дистанция движения и щелчка	Движение между точками с нажатием	$dis_{oc} = 19.\left\|m_{mov_{oc_i}} - m_{mov_{oc_{i+1}}}\right\|$ $dis_{dc} = 19.\left\|m_{mov_{dc_i}} - m_{mov_{dc_{i+1}}}\right\|$
x-скорость	Скорость перемещения по оси x	$speed_x = \frac{\sum_{i=1}^{n}(x_i - x_{i+n})}{\sum_{i=1}^{n}(time_i - time_{i+n})}$
y-скорость	Скорость перемещения по оси y.	$speed_y = \frac{\sum_{i=1}^{n}(y_i - y_{i+n})}{\sum_{i=1}^{n}(time_i - time_{i+n})}$
X-ускорение	Ускорение движения по оси x	$acc_x = \frac{speed_x}{\sum_{i=1}^{n}(time_i - time_{i+n})}$
Y-ускорение	Ускорение движения по оси y	$acc_y = \frac{speed_y}{\sum_{i=1}^{n}(time_i - time_{i+n})}$
Ускорение расстояния	Ускорение в зависимости от расстояния межу точками затишья	$acc_d = \frac{acc_x + acc_y}{\sum_{i=1}^{n}}$

Система непрерывной аутентификации основана на Веб-странице, разделенной на 4 строки в соответствии с делением клавиатуры, как показано на рисунке 4.1, где каждый сектор обозначает клавишу и код по оформления (ASCII, аббр. от англ. American Standard Code for Information Interchange)[188], как показано на Таблица 3, и также содержит расстояние между каждой клавишей, равное 19 мм, чтобы облегчить процесс расчета пройденного расстояния при использовании пользователем мыши. Рассчитывается по формуле:

$$ms_{rp} = \begin{pmatrix} r_i p_i & r_i p_{i+1} & r_{i+1} p_{i+n} \\ r_{i+1} p_i & r_{i+1} p_{i+1} & r_{i+1} p_{i+n} \\ \vdots & \cdots & \ldots \\ r_4 p_i & r_4 p_{i+1} & r_4 p_{i+n} \end{pmatrix}, \quad (4.1)$$

где ms – матрица расположения сектора; r - номер части разделенной клавиатуры; p – расположение клавши в строке.

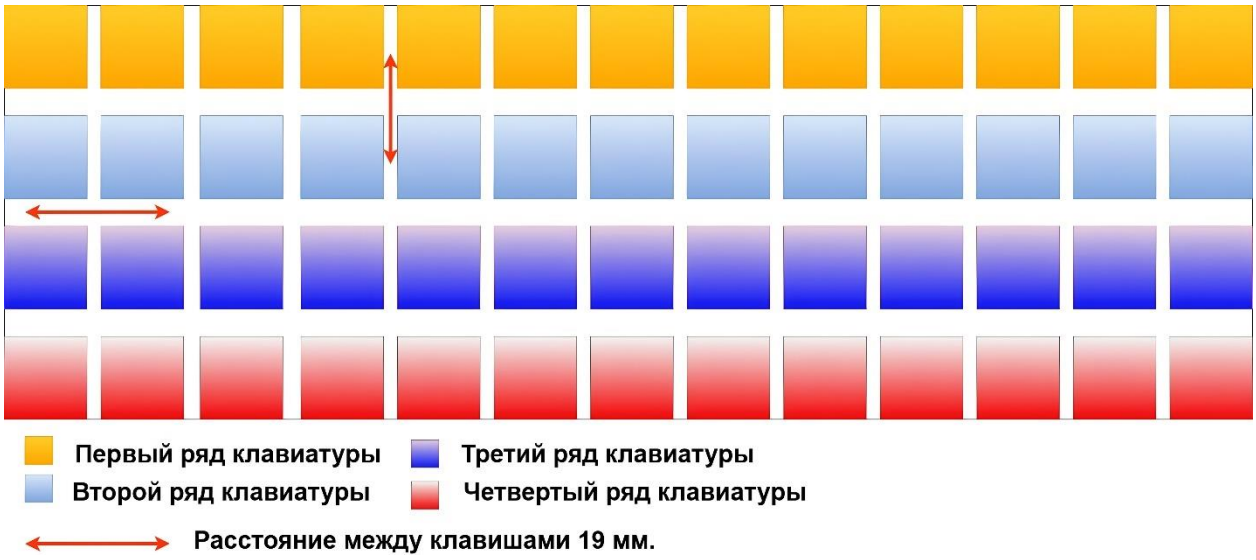


Рис. 4.1. Разделение строк веб-страницы

4.2.1 Процесс построения модели пользователя на этапе обучения по результатам модели непрерывной аутентификации на основе динамики мыши.

Расстояние между квадратами определяется при перемещении по веб-странице. Рассчитывается по формуле:

$$Square_{location} = \left| ms_{rp_i} - ms_{rp_{i+1}} \right|, \quad (4.2)$$

где $Square_{location}$ – расположение квадратов.

$$\vec{v} = \frac{d_{\vec{r}}}{d_{\vec{t}}} = \frac{19mm \,.\, Square_{location}}{t_{i+1} - t_i}, \quad (4.3)$$

$$\vec{a}(t) = \lim_{\Delta t \to 0} \frac{\vec{a}(t + \Delta t) - \vec{v}(t)}{\Delta t} = \frac{\Delta \vec{v}}{\Delta t}, \quad (4.4)$$

$$\vec{a}(t) = \frac{dv_x}{dt}\,\hat{\imath} + \frac{dv_y}{dt}\,\hat{\jmath} + \frac{dv_z}{dt}\,\hat{k}, \quad (4.5)$$

$$\vec{a}(t) = \frac{d^2{}_x}{dt^2}\,\hat{\imath} + \frac{d^2{}_y}{dt^2}\,\hat{\jmath} + \frac{d^2{}_z}{dt^2}\,\hat{k}, \quad (4.6)$$

$$\vec{a}(t) = \frac{\Delta\vec{v}}{\Delta t} = \frac{\vec{v}_{i+1} - \vec{v}_i}{t_{i+1} - t_i}, \tag{4.7}$$

где $\vec{a}(t)$ – стандартная векторная ускорения; Δt – количество временных меток между щелчками, отпусканием и движениями мыши; $\vec{v}(t)$ – стандартная векторная скорость

$$\overline{mva} = \frac{1}{n}\left(\sum_{i=1}^{n} \vec{a}_i\right), \tag{4.8}$$

$$\overline{mva} = \frac{\sum \vec{a}(t)_i - \vec{a}(t)_{i+1} - \cdots - \vec{a}(t)_{i+n}}{\mu - 1}, \tag{4.9}$$

где $\overline{mva}$ – среднее арифметическое значений ускорения; μ – длина строки, образуемой движением мыши.

В ходе исследовательского эксперимента было показано, что перемещение пользователя по веб-странице при использовании мыши зависит от четырех основных движений, как показано на рисунке 4.2:

1. Перемещение из одной точки в другую по прямой линии
2. Перемещение из одной точки в другую длинным прямоугольным движением
3. Движение из одной точки в другую по кривой.
4. Движение из одной точки в другую зигзагообразными способом

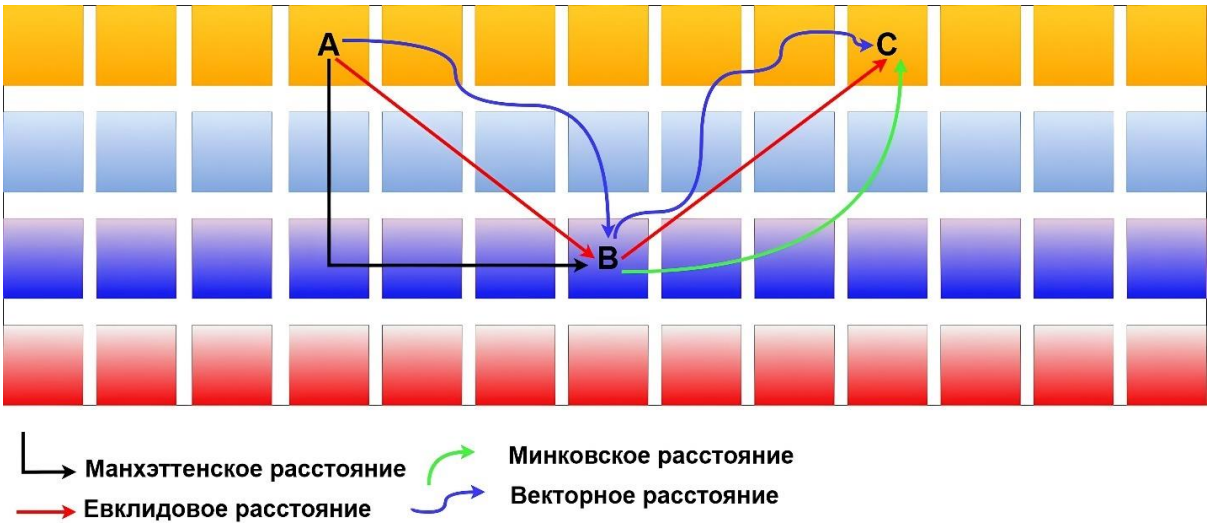


Рис. 4.2. тип движении по веб-странице

В первом случае, если движение происходит между одной точкой и другой по прямой линии по координате x, y или по наклонной, то расстояние между точками вычисляется на основе Евклидова расстояния, поскольку Евклидово расстояние [189, 190] вычисляет кратчайшее расстояние между двумя или более точками на прямой линии. Рассчитывается по формуле:

$$e(ac) = \sqrt[2]{\sum_{i}^{n}((ac_i - ac_i)/\alpha_i)^2}, \quad (3.10)$$

$$e = \sqrt{\left(\frac{\vec{a}(t)_i - \vec{a}(t)_{i+1}}{\overline{mva}}\right)^2 + \left(\frac{\vec{a}(t)_{i+2} - \vec{a}(t)_{i+3}}{\overline{mva}}\right)^2 + \cdots + \left(\frac{\vec{a}(t)_{i+n} - \vec{a}(t)_n}{\overline{mva}}\right)^2}, \quad (4.11)$$

$$e(ac) = \sqrt[2]{\sum_{i}^{n}\left(\frac{\vec{a}(t)_i - \vec{a}(t)_i}{\overline{mva}}\right)^2}, \quad (4.12)$$

где e – Евклидово расстояние; α_i – среднее арифметическое ускорение

Во втором случае, если движение мыши между двумя или более точками имеет вид длинной прямой линии, разделенной на две части, то расстояние между точками рассчитывается на основе Манхэттенского расстояния [189], поскольку оно измеряет наибольшее расстояние между двумя или более точками на осях x и y. Рассчитывается по формуле:

$$m(ac) = \sum_{i}^{n}|(ac_i - ac_i)|/\alpha_i, \quad (4.13)$$

$$m(ac) = \frac{|\vec{a}(t)_i - \vec{a}(t)_{i+1}| + |\vec{a}(t)_{i+2} - \vec{a}(t)_{i+3}| + \cdots + |\vec{a}(t)_{i+n} - \vec{a}(t)_n|}{\mu - 1}, \quad (4.14)$$

$$m(ac) = \sum_{i=1}^{n}\left|\frac{\vec{a}(t)_i - \vec{a}(t)_i}{\mu - 1}\right|, \quad (4.15)$$

где m – Манхэттен расстояние;

Во втором случае, если движение мыши принимает форму кривой между двумя или более точками, расстояние между точками рассчитывается на основе расстояния Минковского [191, 192]. Это происходит потому, что чем выше значение p, тем более изогнута линия соединяющий две точки становится. Таким образом, значение p будет равно трем. Рассчитывается по формуле:

$$mi(ac) = s\left(\sum_{i=1}^{n} |\vec{a}(t)_i - \vec{a}(t)_i|^3\right)^{\frac{1}{3}}, \quad (4.16)$$

$$mi(ac) = \sqrt[3]{\left|\frac{\vec{a}(t)_i - \vec{a}(t)_{i+1}}{\overline{mva}}\right|^3 + \left|\frac{\vec{a}(t)_{i+2} - \vec{a}(t)_{i+3}}{\overline{mva}}\right|^3 + \cdots + \left|\frac{\vec{a}(t)_{i+n} - \vec{a}(t)_n}{\overline{mva}}\right|^3}, \quad (4.17)$$

$$mi(ac) = \sqrt[3]{\sum_{i=1}^{n} \left|\frac{\vec{a}(t)_i - \vec{a}(t)_{i+1}}{\overline{mva}}\right|^3}, \quad (4.18)$$

где mi – расстояние Минковского;

В четвертом случае, если движение мыши между точками образовано зигзагообразной линией, а не прямой, расстояние будет рассчитано на основе расстояния, направленного законами кинематики [193], поскольку он вычисляет расстояние от начальной до конечной точки, независимо от направления движения, будь то прямолинейное, извилистое или зигзагообразное. Поэтому это расстояние подходит для расчета таких движений. Рассчитывается по формуле:

$$\overrightarrow{rv} = \sqrt[3]{\sum_{i=1}^{n} \frac{19mm\ .\ Square_{location}}{\overline{mva}}}, \quad (4.19)$$

где $\overrightarrow{rv}$ – векторное расстояние между точками кривизны.

$$distance_{Total} = \sqrt{\frac{e(ac) + m(ac) + mi(ac) + \overrightarrow{rv}}{4}}, \quad (4.20)$$

где $distance_{Total}$ – общее расстояние, рассчитанное из четырех расстояний.

4.2.2 Рабочая среда этапа обучения

После того как пользователь успешно завершает этап проверки одноразового пароля, процесс непрерывной аутентификации начинает отслеживать перемещения пользователя при выполнении действий и операций на веб-сайте как показное на рисунке 4.3.

1. На этапе обучения характеристики динамики мыши извлекаются из

нажатий и движений, рассчитываются значения временных меток и сохраняются в файле с расширением CSV.

2. Расчет значений кинематики для движения мыши внутри веб-страницы, векторной скорости, векторного расстояния между разделенными квадратами и ускорения, поскольку значение ускорения является основным значением, из которого рассчитываются значения расстояний между каждым движением мыши, поскольку скорость пользователя при использовании мыши может меняться между каждым периодом, так как ускорение может быть похоже на то, когда движение мыши перемещается между текстом поля, кнопками и т. Д как показано на рисунке 4.3.
3. После этого общее расстояние рассчитывается для каждого движения мыши, прямое движение, прямое движение с двумя линиями, случайное движение или изогнутое движение, путем расчета Манхэттенского расстояния, Евклидова расстояния, расстояния Минковского и векторного расстояния.

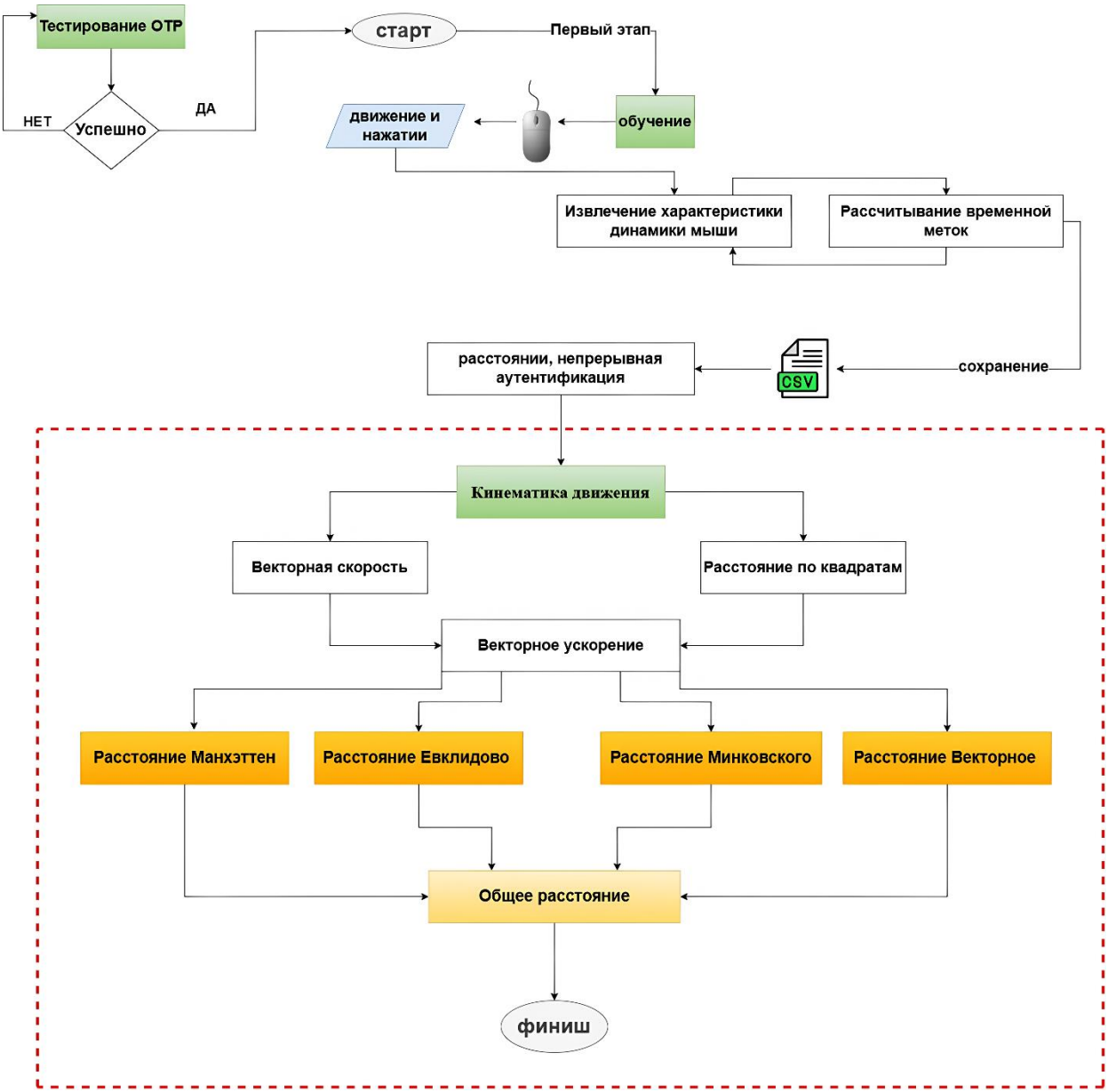


Рис. 4.3. Блок-схема этапа обучения непрерывной аутентификации

4.2.3 Процесс подтверждения модели пользователя на этапе тестирования по результатам модели непрерывной аутентификации на основе динамики мыши.

На этапе проверки сеанса расстояние Левенштейна [194, 195] используется из-за его высокой способности и эффективности в определении степени различия между буквами, сформированными на этапе обучения, и буквами, сформированными на этапе проверки.

Существует много расстояний, которые измеряют сходство или различие между сериями букв или символов, например, расстояние косинусного сходства,

расстояние Жаккара, расстояние Хэмминга и индекс Серенсена-Дайса, но расстояние Левенштейна имело большое значение, поскольку оно вычисляет разницу независимо от длины серии букв или символов.

Расстояние Левенштейна — это строковая метрика для измерения разницы между двумя последовательностями. Расстояние Левенштейна между двумя словами — это минимальное количество односимвольных правок (вставок, удалений или замен), необходимых для замены одного слова на другое. Оно названо в честь советского математика Владимира Левенштейна, который определил метрику в 1965 году [196].

Левенштейн окончил Московский государственный университет в 1958 году, где учился на механико-математическом факультете. После окончания университета он работал в Институте прикладной математики им. М. В. Келдыша [196].

Алгоритм не улучшался более 50 лет и на то были веские причины. Согласно MIT, вполне возможно, что алгоритм Левенштейна — лучшее, что мы когда-либо получим с точки зрения эффективности.

Расстояние Левенштейна также может называться расстоянием редактирования, хотя этот термин может также обозначать более крупное семейство метрик расстояния, известных под общим названием расстояние редактирования. Оно тесно связано с попарным выравниванием строк. Расстояние Левенштейна рассчитывается по формуле:

$$\mathrm{lev}(a,b) = \begin{cases} |a| & \text{□□ } |b| = 0 \\ |b| & \text{□□ } |a| = 0 \\ lev(tail(a), tail(b)) & \text{□□ } h\text{□□□ }(a) = head(b), \\ 1 + min\begin{cases} lev(tail(a), b) \\ lev\big(a, tail(b)\big) \\ lev\big(tail(a), tail(b)\big) \end{cases} & \text{□□}h\text{□□□□□□}. \end{cases}, \quad (4.21)$$

где lev– расстояние Левенштейна; a – номер строки в массиве.; b – номер столбца в матрице; $h\text{□□□}\,(a)$ – первая буква ряда; $h\text{□□□}\,(b)$ – первая буква столбца; $tail$ – серия букв, за исключением первой буквы.

Для применения расстояния Левенштейна используется алгоритм Вагнера-Фишера[197, 198], который вычисляет расстояние редактирования на основе наблюдения, что если мы зарезервируем матрицу для хранения расстояний редактирования между всеми префиксами первой строки и всеми префиксами второй, то мы сможем вычислить значения в матрице путем ее заполнения и, таким образом, найти расстояние между двумя полными строками, как последнее вычисленное значение. Рассчитывается по формуле:

$$D(i,j) = \begin{cases} 0, & i = 0, j = 0 \\ i, & j = 0, i > 0 \\ j, & i = 0, j > 0 \\ \min \begin{cases} D(i,j-1) + 1 \\ D(i-1,j) + 1, \\ D(i-1,j-1) + m(S_1[i], (S_2[j] \end{cases} & j > 0, i > 0 \end{cases}, \quad (4.22)$$

где D – расстояние между двумя строками методом Вагнера-Фишера; i – номер строки в массиве.; j – номер столбца в матрице; S – буквы, образованные движением мыши; m – значение сходства между буквами; min – минимальное значение между расстояниями буквами.

В предыдущих методах пороговое значение вычисляется в процессе обучения, но при непрерывной аутентификации пороговое значение вычисляется в процессе тестирования, на основе букв и символов, образованных движением мыши, как показано на рисунке 4.4. Это связано с тем, что пользователь, впервые зайдя в систему, выполняет ваши операции, такие как оплата или другие операции, согласно веб-сайту, и соответственно его данные динамически извлекаются и ожидают другого процесса ввода.

Пороговое значение очень чувствительно, поскольку является критерием, по которому пользователю разрешается продолжить сеанс или наоборот. Поэтому для достижения высокой степени точности пороговое значение рассчитывается на основе трех случаев в зависимости от расстояния Левенштейна, чтобы уменьшить процент ложного принятия и ложного отвержения.

Случаи порогового значения относительно количества щелчков мыши, как показано по формуле:

$$threshould_i = \begin{cases} threshould_1, & \mu_{lev} > \mu_{Training} \\ threshould_2, & \mu_{lev} < \mu_{Training}, \\ threshould_3, & \mu_{lev} = \mu_{Training} \end{cases} \quad (4.23)$$

где $threshould_i$ – пороговое значение; μ_{lev} – длина букв и символов, сформированных на этапе тестирования; $\mu_{Training}$ – длина букв и символов, сформированных на этапе обучения;

– Если Длина букв и символов, сформированных на этапе тестирования больше длины букв и символов, сформированных на этапе обучения, пороговое значение рассчитывается по формуле:

$$threshould_1 = \frac{distance_{Total}}{\mu_{lev} - (\mu_{lev} - \mu_{Training})}, \quad (4.24)$$

– Если длина букв и символов, сформированных на этапе тестирования меньше длины букв и символов, сформированных на этапе обучения, пороговое значение рассчитывается по формуле:

$$threshould_2 = \frac{distance_{Total}}{\mu_{lev} + (\mu_{lev} - \mu_{Training})}, \quad (4.25)$$

– Если Длина букв и символов, сформированных на этапе тестирования равно длине букв и символов, сформированных на этапе обучения, пороговое значение рассчитывается по формуле:

$$threshould_3 = \frac{distance_{Total}}{\mu_{lev} + 1}, \quad (4.26)$$

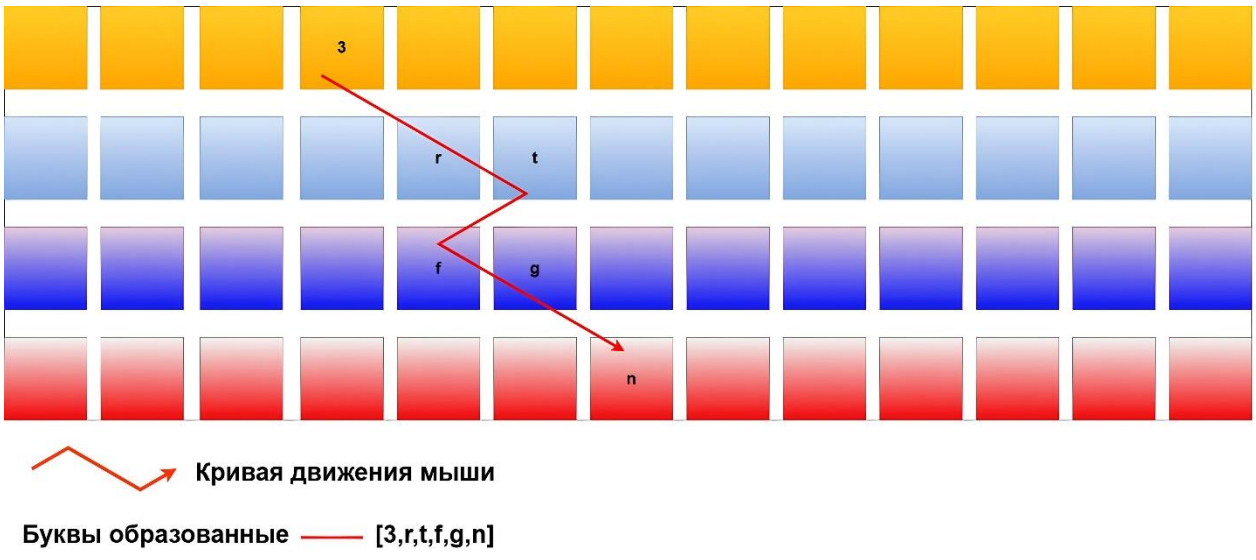


Рис. 4.4. Формирование строкового типа по движению мыши

На этапе обучения пользователь совершает движение мыши внутри веб-страницы. Это не означает, что на этапе проверки пользователь должен совершать то же движение мыши, которое он делал на этапе обучения. Чтобы решить эту проблему и избежать высокого уровня ложных отклонений, в исследовании используются расстояния, упомянутые выше (Евклидово, Манхэттенское, Московского, векторное расстояние), но для достижения высокой степени точности принцип построения расстояния Левенштейна было очень важно для непрерывной аутентификации, поскольку движение пользователя время от времени может меняться. На рисунке 4.5 показан принцип действия расстояния Левенштейна, например, на этапе обучения пользователь совершал движение мышью строкового типа «bonch» по размеру 5, на этапе проверки он совершил движение мышью во время затишья на доли секунды, извлекается строка букв и символов, образованная движением мыши. Вычисляется расстояние Левенштейна, как показное в первой части, оно равно 3 и меньше размера строкового типа, сформированного в процессе обучения. Выбирается первое пороговое значение, пользователь завершает свое движение внутри веб-страницы. Когда снова возникает затишье, строковой тип извлекается и вычисляется расстояние Левенштейна, оно равно 5, тому же значению, что и строковой тип, сформированный в процессе обучения, выбирается третье пороговое значение. Пользователь продолжает совершать свое движение, извлекается строковой тип и вычисляется расстояние Левенштейна, оно равно 8, больше размера строкового типа, сформированного в процессе обучения, выбирается второе пороговое значение, расчета расстояния Левенштейна и выбора порогов продолжается до тех пор, пока система не обнаружит подозрительное поведение и пользователь не завершит сеанс.

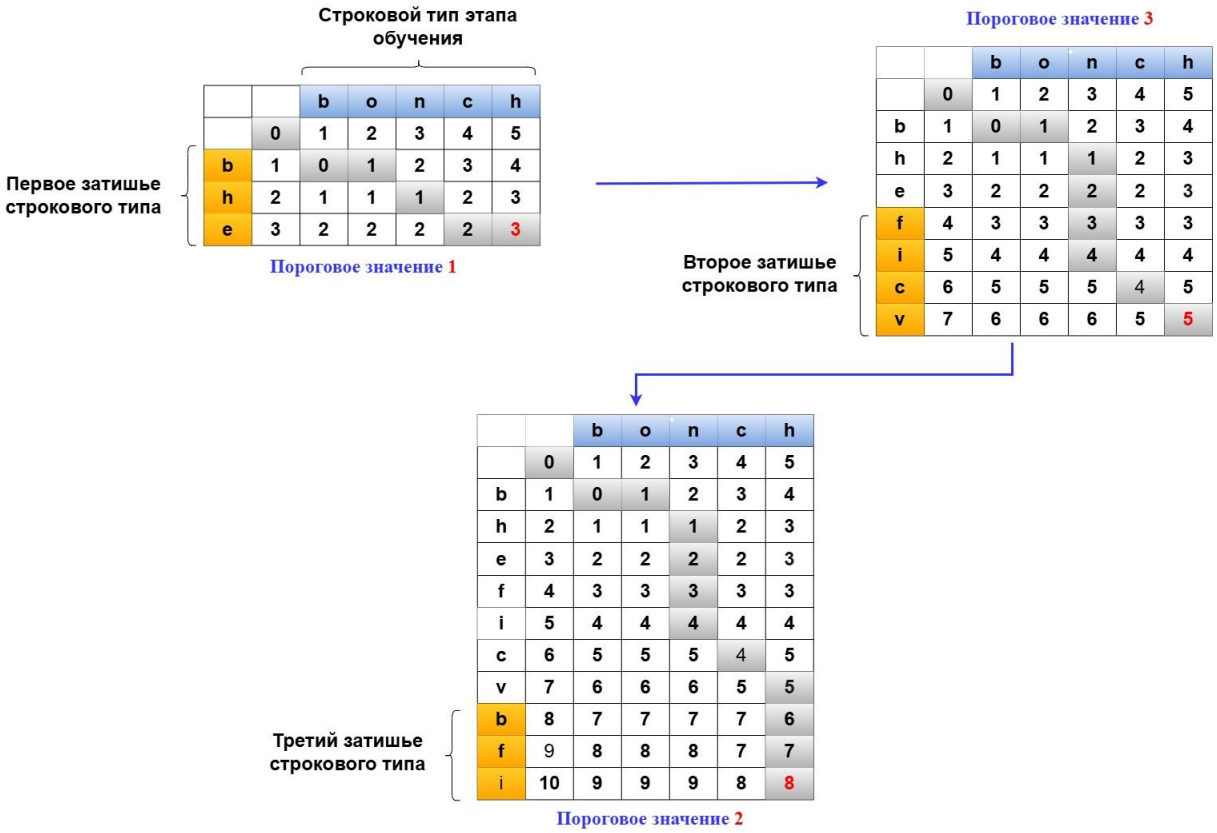


Рис. 4.5. принцип расстояния Левенштейна для динамики мыши

Процесс проверки сеанса продолжается периодически, в зависимости от разницы между двумя строка символов, сформированными на этапе обучения и этапе проверки. На этапе обучения, когда пользователь перемещается по веб-сайту, при наведении курсора на квадраты внутри веб-страницы формируется ряд букв и символов, где каждый квадрат представляет букву или символ на клавиатуре.

Поэтому на этапе проверки действительность пользователя проверяется на каждой паузе с помощью мыши, где в течение этого периода алгоритм извлекает строку букв и символов и рассчитывает расстояние Левенштейна, как показное на рисунке 4.6. И в тот же момент выбирается пороговое значение, на основе которого пользователю разрешается продолжить сеанс или он блокируется.

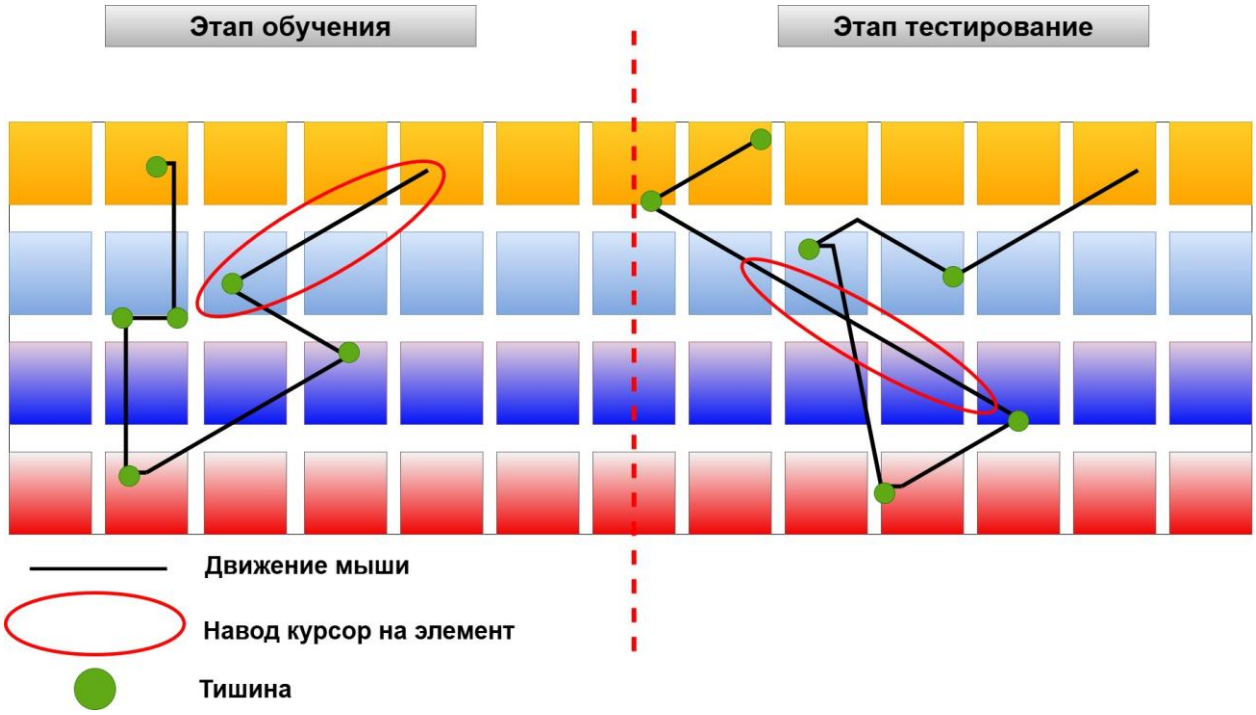


Рис. 4.6. процесс затишья, вычисляющийся расстоянием Левенштейна

4.2.4 Рабочая среда этапа тестирования

После того, как пользователь впервые использует веб-сайт, начинается этап проверки на основе данных, полученных на этапе обучения как показное на рисунке 4.8.

На этапе тестирования характеристики динамики мыши извлекаются из нажатия и движения, рассчитываются значения временных меток и сохраняются в файле с расширением TXT как показное на рисунке 4.7.

1. Расчет и выбор из 3-х порогов на каждом этапе затишья движения мыши, путем расчета расстояния Левенштейна.
2. После этого общее расстояние рассчитывается путем расчета четырех расстояний: Манхэттенского, Евклидова расстояния, расстояния Минковского и векторного расстояния, в соответствии с движением внутри веб-страницы.
3. Затем происходит этап сравнения, где если общее расстояние на этапе проверки меньше или равно пороговому значению, сеанс продолжается, в противном случае сеанс завершается. Процесс расчета общего расстояния для каждого перемещения внутри веб-страницы периодически повторяется

при каждой остановке движения мыши, чтобы гарантировать достоверность сеанса пользователя и, таким образом, этот метод очень эффективен и способен обнаружить подозрительное поведение с минимальными затратами времени и движением мыши на веб-странице.

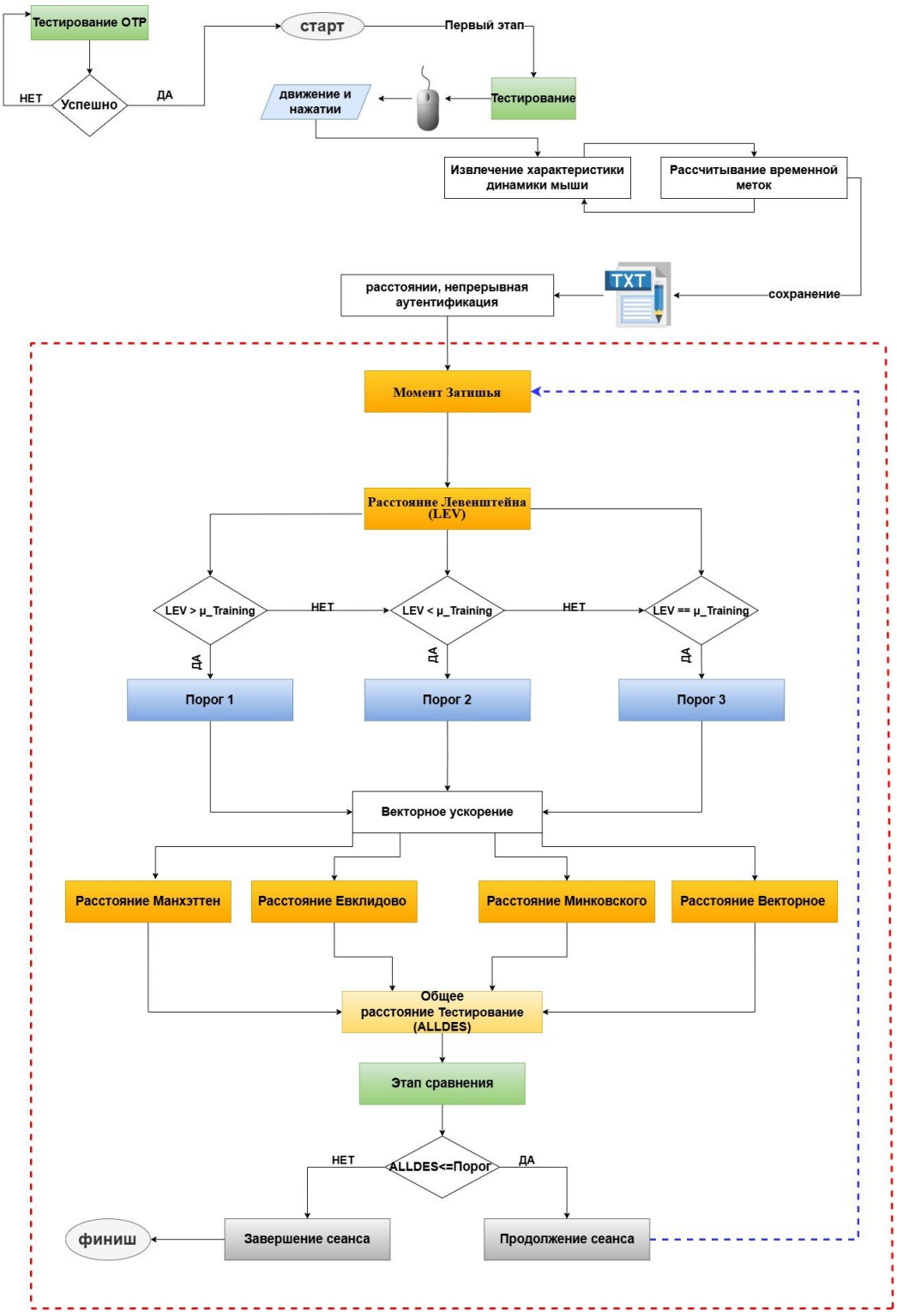


Рис. 4.7. Блок-схема этапа тестирования непрерывной аутентификации

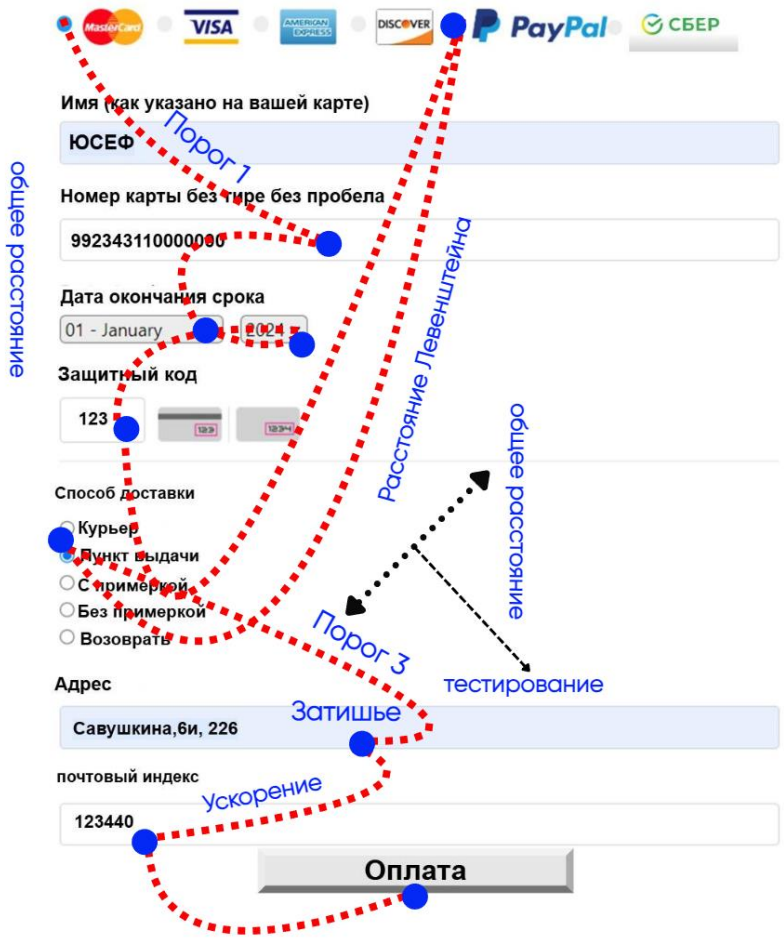


Рис. 4.8. Внедрение непрерывной аутентификации

4.3 Эксперимент по всему научному результату

Для целей данного исследования мы создали новую биометрическую базу данных. Мы собрали данные и провели эксперимент в университетской резиденции, расположенной в России, город Санкт-Петербург, улица Караваевская, дом 34. Был сделан выбор участников из множества национальностей, чтобы исключить возможность предвзятой выборки и тем самым повысить достоверность нашего эксперимента. Участники принадлежат к 6 разным странам, таким образом, всего в этом эксперименте добровольно приняли участие 60 человек, в их числе: 15 человек русской национальности, 16 человек тунисской национальности, 7 человека сирийской национальности, 5 человека иракской национальности, 10 человек алжирской национальности и 7 человек африканской

национальности. Тип рук используемая для набора текста на клавиатуре.15 правша, 5 левша, 40 обе руки.

Для постановки биометрического эксперимента на добровольцах все, что вам нужно компьютер с клавиатурой QWERTY, мышью и приложение для сбора данных о динамике нажатия клавиш и динамике мыши.

Местоположение и положение оборудования зафиксированы и не могут быть перемещены на протяжении всего сеанса для обеспечения достоверности результатов.

Всего было проведено пять эксперимента, чтобы проверить результативность программы.

Во время первого эксперимента, все участники написали свои логин (электронная почта) и пароль, а затем перешли к этапу обучения, повторив введенный ими пароль три раза тем же методом, который они используют при написании на клавиатуре, а также используя движения мыши во время фазы обучения.

После этапа обучения все динамические данные о нажатиях клавиш и движениях мыши были получены и сохранены в модели в программном обеспечении.

После этапа обучения мы попросили участников войти в систему, используя идентификационный логин и пароль, которые были введены во время обучения, а также используя движение мыши во время процесса входа.

Желаемая цель и удовлетворительный результат были достигнуты благодаря успешной аутентификации всех добровольцев и входу в систему без каких-либо проблем. Эффективность динамики нажатия клавиш повышается на 4%, динамики мыши на 2%, определения рук на 10%.

После этого один раз происходит этап генерации случайного слова из того же пароля. Добровольцы ввели случайное слово, полученное по электронной почте, и результаты показали, что система способна аутентифицировать пользователя, даже если слово случайное, потому что система строится динамически. Получная точность аутентификации от 93%.

После успешного завершения процесса аутентификации добровольцев и входа в систему начинается этап непрерывной аутентификации с отслеживания движений, стиля и поведения пользователя во время использования системы. Во время сеанса пользователям предлагается заполнить форму банковского платежа и перемещаться между полями с помощью мыши. Биометрическая система сохраняет свои данные в модели. После этапа обучения пользователь регистрируется и снова входит в систему, заполняя форму банковского платежа еще раз таким же образом. Система не обнаружила никакого подозрительного поведения, поэтому успешность непрерывной аутентификации составила 98%.

На тех же добровольцах был проведен второй эксперимент, чтобы выяснить надежность системы аутентификации. Добровольцу было предложено использовать логин и пароль другого добровольца. Доброволец не мог войти в систему, хотя у него были те же пользовательские данные. Был получен высокий результат точности и безопасности, достигающий примерно 98%.

Во время третьего эксперимента, добровольцев попросили ввести логин и пароль, которыми они владели, но с разницей в манере письма, которой они пользовались во время первого обучения системы. Один из добровольцев на этапе обучения вводил пароль правой рукой, но на этапе аутентификации он ввел его обеими руками, и, таким образом, система обнаружила другое биометрическое значение и не допустила пользователя в систему. Следовательно, степень безопасности и точности мягких биометрических измерений, которые были созданы на основе динамического нажатия клавиши, достигли 95%.

Четвертый эксперимент был проведен на добровольцах. Первый доброволец успешно проводит первый этап аутентификации, после чего начинается другой этап, который генерирует случайное слово через OTP на основе введенного пароля. Второй доброволец завершает этот процесс, получая пароль было сгенерировано случайное слово. В результате вмешательства в сеанс, система отклонила слово, введенное вторым добровольцем, хотя случайное слово было введено правильно. Это связано с тем, что биометрические измерения этого человека отличаются от измерений первого добровольца и последний доброволец считался

недействительным пользователем. Таким образом, уровень безопасности был получен и повышен до 94%.

В ходе пятого эксперимента добровольцам предлагалось, после успешного прохождения аутентификации, ввести данные в форму банковского платежа, но на этот раз с помощью подозрительного движения мыши, которое отличалось от манеры ее движения на этапе обучения. Сеанс добровольцев был немедленно прерван. Таким образом получен высокий результат по скорости обнаружения несанкционированного пользователя, достигший 96%.

В таблице 8 показаны результаты эксперимента по всему научному результату. В него входят следующие параметры: коэффициент ложного принятия, коэффициент ложного отклонения, коэффициент ошибок и площадь под кривой.

Таблица 8 — эксперимент методов по ROC кривая

Метод	FAR	FRR	ERR	AUC	Точность
динамики нажатия клавиш	0.012677	0.014568	0.013623	0.99810	98.7262%
Определение руки	0.044664	0.040262	0.042145	0.96413	95.0216%
динамики мыши	0.015029	0.03556	0.025295	0.9527	96.3639%
OTP	0.049269	0.054649	0.051959	0.9453	93.1631%
НЕП	0.013209	0.029605	0.021407	0.9944	97.1862%

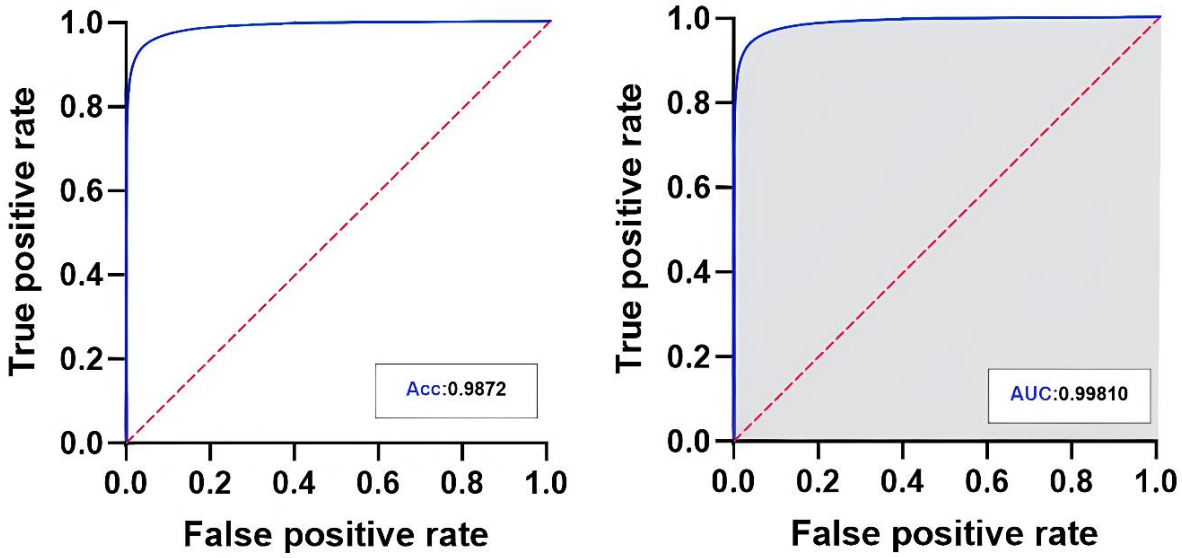


Рис.4.9. ROC-кривая и AUC динамики нажатия клавиш

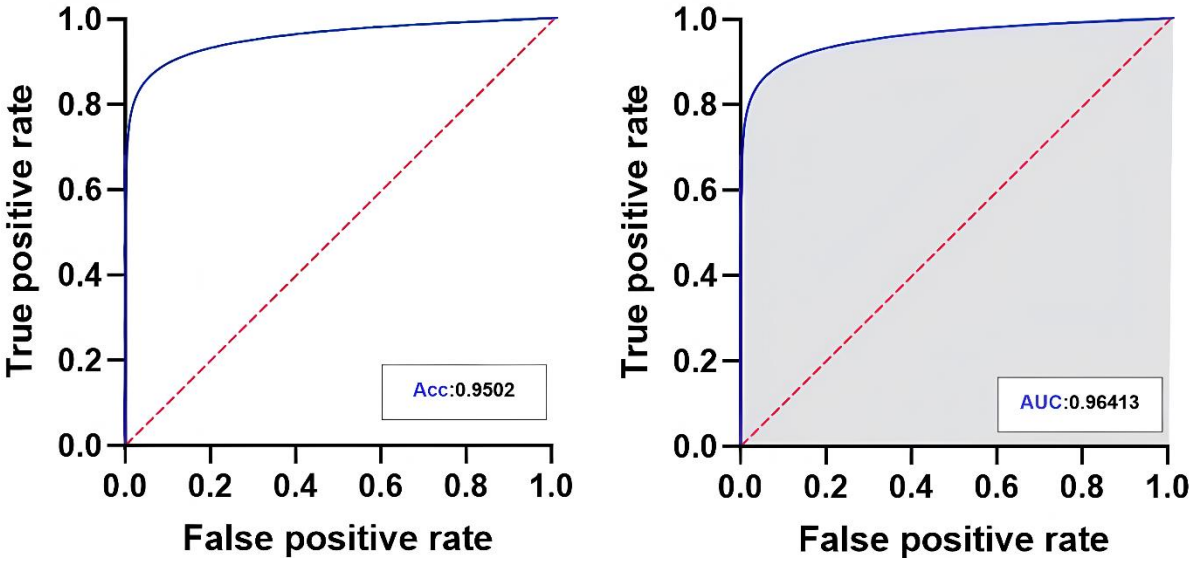


Рис. 4.10. ROC-кривая и AUC идентификация руки

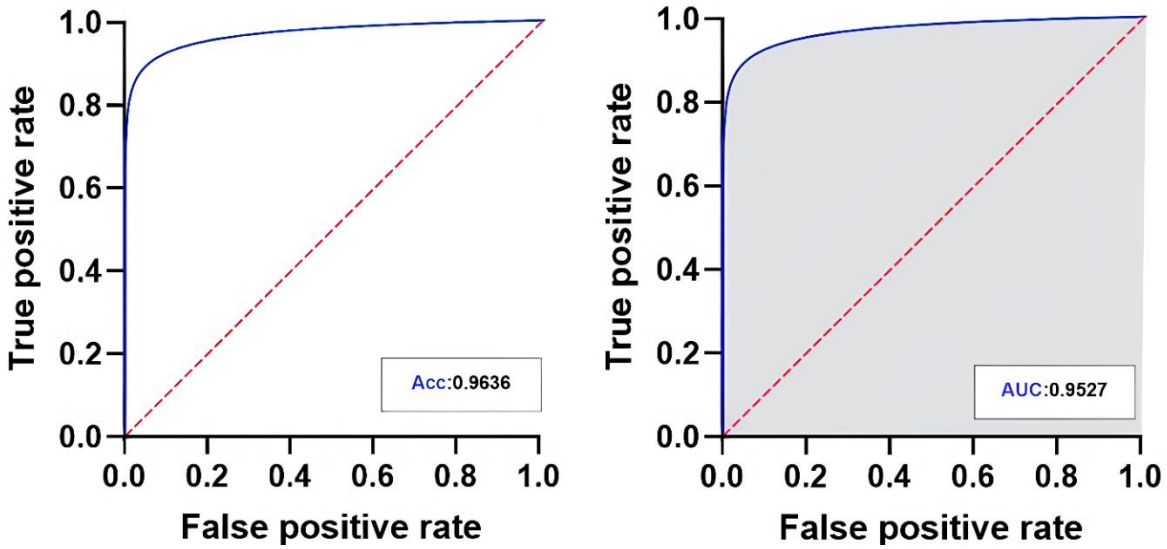


Рис. 4.11. ROC-кривая и AUC динамики мыши

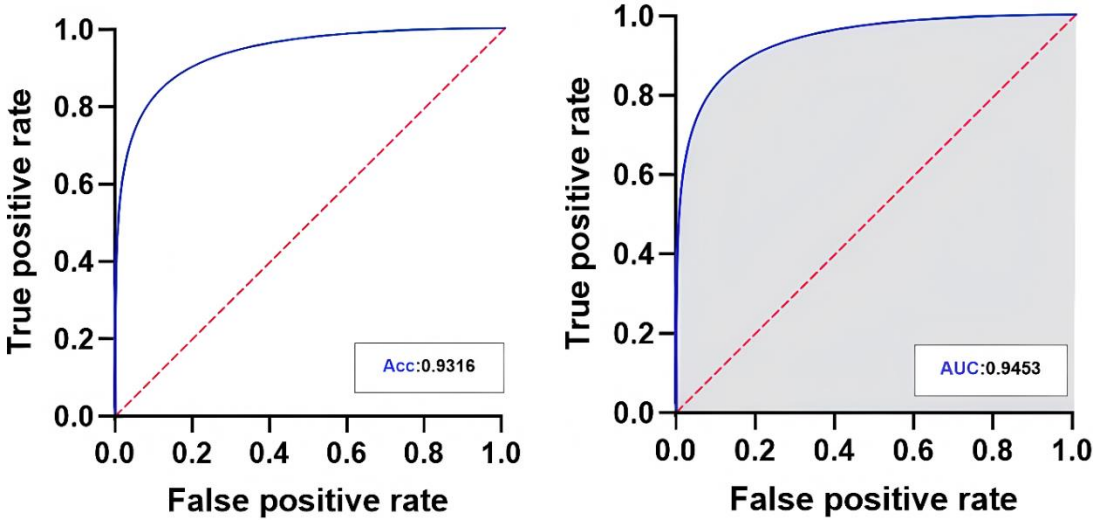


Рис. 4.12. ROC-кривая и AUC OTP

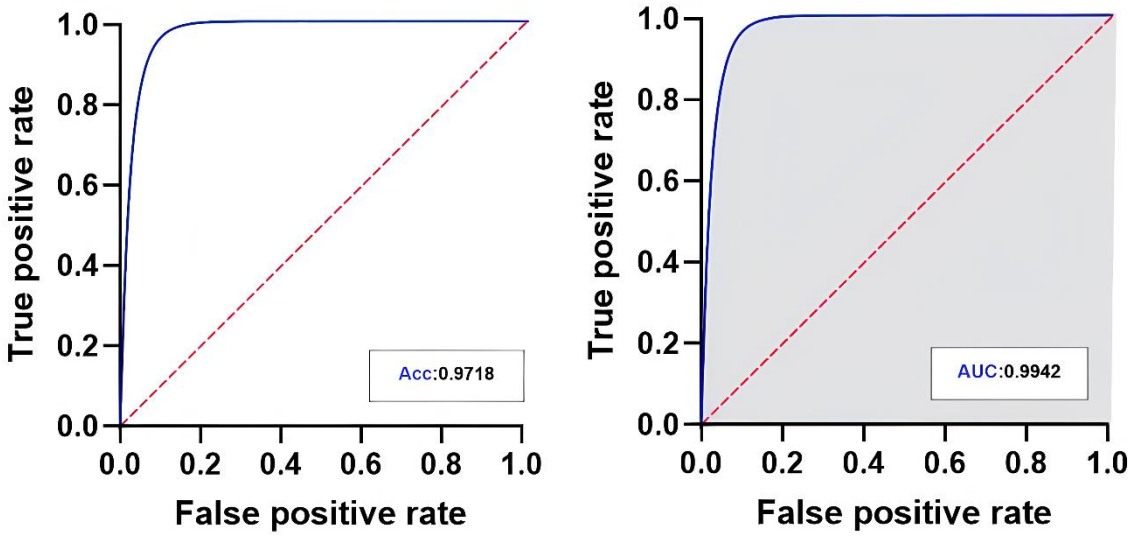


Рис. 4.13. ROC-кривая и AUC непрерывной аутентификации

Выводы по 4-й главе

В данном разделе проводилось исследование и разработка методов обработки данных, характеризующих динамику работы пользователей с мышью на основе непрерывной аутентификации.

В ходе проведенного исследования были достигнуты следующие результаты:

- Проведен исследовательский анализ особенностей непрерывной аутентификации и ее важности для защиты веб-приложений, а также разрыва между ней и статической аутентификацией;
- Разработана система, способная отслеживать поведение и движения пользователя при использовании мыши в веб-приложении;
- Предложен подход к повышению точности перемещения мыши, основанный на разделении веб-страницы на квадраты, где каждый квадрат представляет собой клавишу на клавиатуре, с целью определения пройденного векторного расстояния. Предлагается четыре движения мыши на веб-странице, каждое из которых представляет: Манхэттенское расстояние, Евклидово расстояние, расстояние Минковского и векторное расстояние;
- В используемом подходе было предложено вычислять пороговое значение на этапе проверки, а не на этапе обучения, как в предыдущих

исследованиях, чтобы снизить показатели ложного отвержения и ложного принятия.

• Предложен подход, который вычисляет пороговое значение периодически и многократно при каждой остановке движения с использованием расстояния Левенштейна. В отличие от предыдущих методов, процесс проверки выполняется многократно в пределах страницы и не основан на определенном времени или количестве движений мыши. Повышено качество аутентификации пользователей в среднем на 2% по сравнению с используемой в существующих работах стандартизацией признаков.

• Была проведена серия экспериментов, в результате которых было подтверждено высокое качество работы предложенных алгоритмов. В результате использования предложенной комбинации алгоритмов достигнуто качество распознавания порядка 0.9718 ROC AUC, что значительно превосходит качество аутентификации пользователей при использовании методов, рассмотренных в существующих научных работах.

ЗАКЛЮЧЕНИЕ

При постановке задачи были рассмотрены основные преимущества поведенческой биометрической аутентификации, ее высокая востребованность на рынке и ее важность для непрерывной аутентификации из-за не трения с пользователем и тот факт, что она не требует дополнительного оборудования. а также перечислены наиболее проблемы, существующие во всех системах аутентификации на основе Динамические нажатия клавиш и мыши. Анализ современных биометрических систем аутентификации и их точности был проведен на (ROC AUC). Недостатки и угрозы, с которыми сталкиваются современные системы, представлены в технологии аутентификации и выборе несовместимого порогового значения для пользователя в долгосрочной перспективе, в дополнение к использованию нейронных сетей и глубокого обучения на основе прогнозов обнаружение аномалий, которые могут быть неверным.

Поставленная в диссертационном исследовании цель по обеспечению защиты от угроз безопасности в системах веб-приложений и по защите пользовательских данных от взлома, путем создания статической и непрерывной системы многофакторной биометрическо-поведенческой аутентификацией не трения с пользователем, достигнута.

Основные научные результаты диссертационной работы состоят в следующем:

1. Разработана высокоточная система двухфакторной аутентификации (2FA), через анализ поведенческой модели для определения пользователя на основе динамики нажатия клавиш и мыши для аутентификации:
   – Разработана модель аутентификации на основе динамики нажатия для выявления аномалий путем алгоритма на основе объединения трех расстояний: Манхэттенского расстояния, Евклидова расстояния и расстояния Чебышева, чтобы вычислить по теореме Пифагора угол прямоугольного треугольника, прилежащего к гипотенузе, как индивидуальное пороговое значение для пользователей, с целью

уменьшения значения частоты ложного отклонения и принятия. Позволяет извлечь характеристики нажатых клавиш от всех случаев использования клавиатуры, включая не валидные к вводу пароля.

– Разработана модель аутентификации на основе динамики мыши, благодаря алгоритму использования расстояния Минковского, которое рассчитывается через кривую четверти круга, и Манхэттенского расстояния, которое находится через площадь четверти круга и длину дуги четверти круга. Исходя из полученных данных высчитывается пороговое значение для последующей аутентификации пользователя. Тестирование зависит от длины дуги, рассчитанной по расстоянию Минковского.
– Разработана модель идентификации количества использованных при печати рук (1 или 2), с применением законов движения кинематики, Предложен подход разделения клавиатуры на 8 частей для облегчения Рассчитывание расстояния между клавишами.

2. Разработана система многофакторной аутентификации (MFA) пользователей и субъектов доступа для веб-приложения. Предложен подход генерации случайного слова от существующего пароля на основе его биометрических данных с помощью расстояние Жаккара, которое рассчитывает сходство между случайным словом и самим паролем для принятия решения, будут проходить последующие измерения через Манхэттенское или Евклидово расстояние. Отправка сообщение со случайным паролем происходит через библиотеку PHPMalier. Система позволяет объединить фактор знания, владения с фактором свойства, связав систему OTP с поведенческо-биометрической системой, позволяет добавить уровень защиты и затруднить взлом кода OTP.
3. Разработана система непрерывной аутентификации пользователей на основе динамики мыши. Предложен подход разделения веб-страниц на сектора, с расстоянием между ними 19мм. Каждый сектор является кодом клавиши по системе (ASCII), подразделяется четыре типа движения

мыши, каждое из которых представляет: Перемещение из одной точки в другую по прямой линии, Перемещение из одной точки в другую длинным прямоугольным движением, Движение из одной точки в другую по кривой и Движение из одной точки в другую зигзагообразными способом. Каждое из движений представляют соответствующие метрики Манхэттенское расстояние, Евклидово расстояние, расстояние Минковского и векторное расстояние с использованием расстояния Левенштейна, которое рассчитывает отличия между образованными строками на этапе обучения и тестирования при каждом моменте затишья. Благодаря этому вычисления порогового значения на этапе проверки, вместо этапа обучения, чтобы снизить показатели ложного отвержения и принятия. быстро выявлять аномалий с каждым моментом затишья.

Перспективными направлениями дальнейших исследований являются

- Разработка алгоритма и подхода, позволяющего определить тип используемой клавиатуры и мыши.
- Разработка подхода, основанного на выявлении самочувствия пользователя, для определения степени влияния окружающей среды и расчета величины биометрических изменений.

СПИСОК ЛИТЕРАТУРЫ

1. Md M.H., Shamima S.N., Marjan A., Rafita H., Fabiha.N.D. Broken Authentication and Session Management Vulnerability: A Case Study Of Web Application / M.H. Md, S.N. Shamima, A. Marjan, H. Rafita,N.D. Fabiha// International Journal of Simulation: Systems, Science & Technology. – 2018. – Vol. 19. – №. 2. – pp. 1-11.
2. Deepa G., Thilagam P.S. Securing web applications from injection and logic vulnerabilities: Approaches and challenges/ G. Deepa, P.S. Thilagam// Information and Software Technology. – 2016. – Vol. 74. – №. 1. – pp. 160-180.
3. Raheem A.S, Veeragandham S., Gillela K. A survey on Major Problems and solutions of Top 10 Web Application Security / A.S. Raheem, S. Veeragandham, K. Gillela // International Journal of Advanced Science and Technology. – 2020. – Vol. 29. – №. 3. – pp. 01-07.
4. Lakh Y., Nyemkova E., Piskozub A., Yanishevskyi V. Investigation of the Broken Authentication Vulnerability in Web Applications / Y. Lakh, E. Nyemkova, A. Piskozub, V. Yanishevskyi // International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). – 2021. – Vol. 2. – №. 1. – pp. 928-931
5. Dipankar D., Arunava R., Abhijit N. Advances in User Authentication / D. Dipankar, R. Arunava, N. Abhijit – Memphis. – 2017. – 369 c
6. Papathanasaki M., Maglaras L., Nick Ayres. Modern Authentication Methods: A Comprehensive Survey/ M. Papathanasaki, L. Maglaras, N. Ayres// intechopen journals AI, Computer Science and Robotics Technology. – 2022. – Vol. 7. – №. 2. – pp. 1-14
7. Tirfe T, Anand V. K. Multi-Factor Authentication: . A survey on trends of two-factor authentication / A. Ometov, S. Bezzateev, N. Mäkitalo, T. Mikkonen, Y. Koucheryavy // Contemporary Issues in Communication, Cloud and Big Data Analytics: Proceedings of CCB 2020. – Springer Singapore. – 2022. – Vol. 4.

– №. 3. – pp. 285-296

8. Chanda K. Nyemkova E., Piskozub A., Yanishevskyi V. Password Security: An Analysis of Password Strengths and Vulnerabilities / K. Chanda // International Journal of Computer Network and Information Security. – 2017. – Vol. 8. – №. 7. – pp. 23-30
9. Misini E. Biometric Authentication/E. Misini// Computer Security and Reliability. – 2022. – Vol. 2. – №. 1. – pp. 1-7
10. Ometov A, Bezzateev S., Mäkitalo N., Mikkonen T., Koucheryavy Y. Multi-Factor Authentication: A Survey / A. Ometov, S. Bezzateev, N. Mäkitalo, T. Mikkonen, Y. Koucheryavy // Cryptography. – 2018. – Vol. 2. – №. 1. – pp. 1-31
11. Bhattacharyya D, Ranjan R., Alisherov F., Choi M. Biometric Authentication: A Review / D. Bhattacharyya, R. Ranjan, F. Alisherov, M. Choi // Advanced Research in Dynamical and Control Systems. – 2009. – Vol. 2. – №. 03. – pp. 13-27
12. Ricanek K. Choi M. Beyond Recognition: The Promise of Biometric Analytics / K. Ricanek, M. Choi // Computer. – 2014. – Vol. 47. – №. 09. – pp. 87-89
13. Cyber Crime & Security. Statista researchers [Электронный ресурс] // Annual number of data compromises and individuals impacted in the United States from 2005 to 2023. USA. – 2023. Режим доступа: https://www.statista.com/statistics/273550/data-breaches-recorded-in-the-united-states-by-number-of-breaches-and-records-exposed/.
14. Yadav D, Gupta D., Singh D., Kumar D., Sharma U. Vulnerabilities and Security of Web Applications / D. Yadav, D. Gupta, D. Singh, D. Kumar, U .Sharma // International Conference on Computing Communication and Automation (ICCCA). IEEE. – 2018. – pp. 1-05
15. Web Security Academy.PortSwigger [Электронный ресурс] // Authentication vulnerabilities. USA. – 2024. Режим доступа: https://portswigger.net/web-security/authentication

16. Ndiaye Y, Barais O., Blouin A., Bouabdallah A., Aillery N. Requirements for preventing logic flaws in the authentication procedure of web applications / Y. Ndiaye, O. Barais, A. Blouin, A. Bouabdallah, N. Aillery // Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing. – 2019. – pp. 1628-1620
17. Naz A, Sarwar M., Kaleem M., Bouabdallah A., Mushtaq M.A, Rashid S. A comprehensive survey on social engineering-based attacks on social networks / A. Naz, O. Sarwar, M. Kaleem, A. Bouabdallah, M.A. Mushtaq, S. Rashid // ADVANCED AND APPLIED SCIENCES. – 2024. – pp. 139-154
18. Authentication. Strongdm [Электронный ресурс] // 11 Common Authentication Vulnerabilities You Need to Know. canada. – 2024. Режим доступа: https://portswigger.net/web-security/authentication
19. Wang X, Yan Z., Zhang R., Zhang P. Attacks and defenses in user authentication systems: A survey. Network and Computer Applications / X. Wang, Z. Yan, Z. Zhang, R. Zhang, P. Zhang // Journal of Network and Computer Applications. – 2021. – Vol. 188. – №. 103080. – pp. 1-21
20. Amuthadevi C, Srivastava S., Khatoria R., Sangwan V. A Study on Web Application Vulnerabilities to and an optimal Security Architecture. Network and Computer Applications / X. Wang, Z. Yan, Z. Zhang, R. Zhang, P. Zhang // arXiv preprint arXiv:2204.07107. – 2021. – Vol. 188. – №. 103080. – pp. 1-06
21. Kumar Y, Satyanarayana A. S., Kumar A., Sharma V. Risks and Threats to Web Applications and Their Preventions: A Theoretical Study on Vital Risks and Threats / Y. Kumar, A. S. Satyanarayana, A. Kumar, V. Sharma // International Journal of Scientific Research in Computer Science Engineering and Information Technology. – 2021. – Vol. 07. – №. 2. – pp. 432-438
22. Behavioral Biometrics Market. ResearchAndMarkets [Электронный ресурс] // Behavioral Biometrics Market by Solution, Type, Organization Size, Deployment, Application, Vertical - Global Forecast 2025-2030. USA. – 2024. Режим доступа:

https://www.researchandmarkets.com/report/behavioral-characteristics?srsltid=AfmBOooO1nujv1W8uvFx91tf5_RKTpl-YMDclaMMM1gNxAWrBAE7zrF9

23. Behavioral Biometrics Market. AlliedMarketResearch [Электронный ресурс] // Behavioral Biometrics Market Outlook: 2025. – 2018. Режим доступа: https://www.alliedmarketresearch.com/behavioral-biometrics-market
24. Typingdna company [Электронный ресурс] // Режим доступа:https://www.typingdna.com
25. Biocatch company [Электронный ресурс] // Режим доступа: https://www.biocatch.com/
26. BehavioSec company [Электронный ресурс] // Режим доступа: https://risk.lexisnexis.com/products/behaviosec
27. Secuve company [Электронный ресурс] // Режим доступа: - https://www.secuve.com/eng/
28. Plurilock company [Электронный ресурс] // Режим доступа: - https://plurilock.com/behavioral-biometrics-guide/2-a-brief-history-of-behavioral-biometrics/
29. NoPassword Cybersecurity company [Электронный ресурс] // Режим доступа: https://cybersecurity-excellence-awards.com/candidates/nopassword/
30. Федеральный закон. КонтурНорматив [Электронный ресурс] // российская федерация федеральный закон об осуществлении идентификации и (или) аутентификации физических лиц с использованием биометрических персональных данных, о внесении изменений в отдельные законодательные акты российской федерации и признании утратившими силу отдельных положений законодательных актов российской федерации. – 2022. Режим доступа:https://normativ.kontur.ru/document?moduleId=1&documentId=439254&ysclid=m4n45iglzd847009460

31. Marana A. N, Falguera F. P. S, Falguera J. R., Jain A. K. Biometrics for Human Identification / A. N. Marana, F. P. S. Falguera, J. R. Falguera, A. K. Jain // RITA. – 2006. – Vol. 13. – №. 2. – pp. 103-130
32. Ahmed A. A. E, Traore I., Ahmed A. Digital Fingerprinting Based on Keystroke Dynamics / A. A. Ahmed, I. Traore, A. Ahmed // In HAISA. – 2008. – C. 94-104
33. Salman O. A, Hameed S. M. Maxion R. User Authentication via Mouse Dynamics/ C. Shen, Z. Cai, X. Guan// Sensors. – 2018. – C. 963-968
34. Pusara M, Brodley C. E. User Re-Authentication via Mouse Movements / M. Pusara, C. E. Brodley // In ACM Workshop Visualization and Data Mining for Computer Security. – 2004. – C. 1-8.
35. Mondal S, Bours P. Combining Keystroke and Mouse Dynamics for Continuous User Authentication and Identification/ S. Mondal, P. Bours // international conference on identity, security and behavior analysis (ISBA). – 2016. – C. 1-8.
36. Kim J., Kim H., Kang P. Keystroke dynamics-based user authentication using freely typed text based on user-adaptive feature extraction and novelty detection/ J. Kim, H. Kim, P. Kang //Applied Soft Computing. – 2018. – Vol. 62. – №. 6. – C. 1077-1087.
37. Zhou Q, Yang Y., Hong F., Feng Y., Guo Z. User identification and authentication using keystroke dynamics with acoustic signal/ Q. Zhou, Y. Yang, F. Hong, Y. Feng, Z. Guo //Applied Soft Computing. – 2016. – C. 445-449.
38. Zheng N, Paloski A., Wang H. An Efficient User Verification System via Mouse Movements/ N. Zheng, A. Paloski, H. Wang // In Proceedings of the 18th ACM conference on Computer and communications security. – 2011. – C. 139-150.
39. Kaixin W, Hongri L., Bailing W., Shujie H., Jia S. A User Authentication and Identification Model Based on Mouse Dynamics / W. Kaixin, L. Hongri, W. Bailing, H. Shujie, S. Jia // Sensors. – 2017. – C. 1-6.

40. Quraishi S. J, Bedi S. S. On Mouse Dynamics as Continuous User Authentication / S. J. Quraishi, S. S Bedi // International Journal of Scientific & Technology Research. – 2012. – Vol. 8. – №. 10. – pp. 1-12.
41. Shen C, Cai Z., Guan X., Maxion R. User Authentication Through Mouse Dynamics/ C. Shen, Z. Cai, X. Guan// Sensors. – 2012. – Vol. 8. – №. 1. – C. 16-30
42. Kim J., Kang P. Recurrent neural network-based user authentication for freely typed keystroke data / J. Kim, P. Kang // arXiv preprint arXiv:1806.06190. – 2018. – Vol. 27. – №. 8. – pp. 861-874
43. Kasprowski P, Borowska Z., Harezlak K. Biometric Identification Based on Keystroke Dynamics / P. Kasprowski, Z. Borowska, K. Harezlak // Sensors. – 2022. – Vol. 22. – №. 9. – C. 1-24.
44. Fawcett T. An introduction to ROC analysis / T. Fawcett // Pattern Recognition Letters. – 2006. – Vol. 27. – №. 8. – pp. 861-874
45. Shen C, Cai Z., Guan X. Continuous Authentication for Mouse Dynamics: A Pattern-Growth Approach / C. Shen, Z. Cai, X. Guan// Sensors. – 2012. – C. 1-12.
46. Deng Y., Zhong Y. User Authentication via Mouse Dynamics/ Y. Deng, Y. Zhong // Recent Advances in User Authentication Using Keystroke Dynamics Biometrics, GCSR. – 2018. – Vol. 2. – №. 1. – C. 23-40
47. Killourhy K. S., Maxion R. A. Comparing Anomaly-Detection Algorithms for Keystroke Dynamics / K. S. Killourhy, R. A. Maxion // In IEEE/IFIP International Conference on Dependable Systems Networks. – 2009.– C. 125-134
48. Opinion Smith.D . InfoSecurity Magazine [Электронный ресурс] // How Secure Is Behavioral Biometrics? – 2019. Режим доступа: https://www.infosecurity-magazine.com/opinions/secure-behavioral-biometrics/
49. Behavioral Biometrics. OneSpan[Электронный ресурс] // What is behavioral biometrics?. – 2024. Режим доступа:

https://www.onespan.com/topics/behavioral-biometrics

50. Finnegan O. L., White J. W., Armstrong B., Adams E. L. The utility of behavioral biometrics in user authentication and demographic characteristic detection: a scoping review / O. L. Finnegan, J. W. White, B. Rosenberger, E. L. Adams // Systematic Reviews. – 2024. – Vol. 13. – №. 1. – C. 1-17
51. Telo J. ANALYZING THE EFFECTIVENESS OF BEHAVIORAL BIOMETRICS IN AUTHENTICATION: A COMPREHENSIVE REVIEW / J. Telo // Journal of Sustainable Technology and Infrastructure for Developing Countrie. – 2019. – Vol. 2. – №. 1. – C. 19-36
52. Gaikwad J., Kulkarni B., Phadol N., Sarukte S. User Authentication using Keystroke Dynamics / J. Gaikwad, B. Kulkarni, N. Phadol, S. Sarukte // Global Research and Development Journal for Engineering. – 2018. – Vol. 3. – №. 6. – C. 58-66
53. Balagani S. P., Woodard D. L. Biometric Authentication and Identification using Keystroke Dynamics: A Survey/ S. P Balagani, D. L Woodard // Journal of Pattern recognition research. – 2012. – Vol. 7. – №. 1. – C. 116-139
54. Seeger M., Bours B. How to Comprehensively Describe a Biometric Update Mechanisms for Keystroke Dynamics / M. Seeger, B. Bours // Third International Workshop on Security and Communication Networks (IWSCN). – 2011. – C. 59-65
55. Hassan S., Selim M., Zayed H. User Authentication with Adaptive Keystroke Dynamics / S. Hassan, M. Selim, H. Zayed // ternational Journal of Computer Science Issues. – 2013. – Vol. 10. – №. 4. – C. 127-135
56. Balagani K. S., Phoha V. V., Ray A., Phoha S. On the Discriminability of Keystroke Feature Vectors Used in Fixed Text Keystroke Authentication / K. S Balagani, V. V Phoha, A. Ray, S. Phoha // Pattern Recognition Letter. – 2011. – Vol. 32. – №. 7. – C. 1070-1080
57. Obaidat M.S., Traore I., Woungang I. Biometric-Based Physical and Cybersecurity Systems / M. S Obaidat, I. Traore, I. Woungang // Cham:

Springer International Publishing. – 2019. – Vol. 7. – №. 1. – C. 116-139

58. Killourhy K., Maxion R. A. Comparing anomaly-detection algorithms for keystroke dynamics / K. Killourhy, R. A Maxion // IEEE/IFIP International Conference on Dependable Systems & Networks. – 2009. – Vol. 7. – №. 1. – C. 125-134
59. Giot R., Dorizzi B., Rosenberger C. A review on the public benchmark databases for static keystroke dynamics / R. Giot, B. Dorizzi, C.A Rosenberger // Computers & Security. – 2015. – Vol. 55. – №. 1. – C. 46-61
60. Leggett J., Williams G., Usnick M., Longnecker M. Dynamic identity verification via keystroke characteristics / J. Leggett, G. Williams, M. Usnick, M. Longnecker // International Journal of Man-Machine Studies. – 1991. – Vol. 35. – №. 6. – C. 859 -870
61. Shadman R., Wahab A. A., Manno M., Lukaszewski M., Hou D., Hussain F. Keystroke Dynamics: Concepts, Techniques, and Applications / R. Shadman, A. A. Wahab, M. Manno, M. Lukaszewski, D. Hou, F. Hussain // arXiv preprint arXiv:2303.04605. – 2023. – Vol. 10. – №. 4. – C. 1-35
62. Yang S., Xu G., Meng H., Wang M. Progressive neighbors pursuit for radar images classification / S. Yang , G. Xu , H. Meng, M. Wang // Applied Soft Computing.– 2021. – Vol. 109. – №. 4. – C. 1-11
63. Coghetto R. Chebyshev Distance / R. Coghetto // Formalized Mathematics.– 2016. – Vol. 24. – №. 2. – C. 121-141
64. The free encyclopedia. Wikipedia[Электронный ресурс] // Chebyshev distance. – 2024. Режим доступа: https://en.wikipedia.org/wiki/Chebyshev_distance
65. Teia L. Extended Pythagoras Theorem using Triangles, and its Applications to Engineering / L. Teia // The Journal of Open Engineering. – 2021. – Vol. 24. – №. 2. – C. 1-29
66. https://en.wikipedia.org/wiki/Square#:~:text=All%20four%20internal%20angles%20of,is%20equal%20to%2090%C2%B0.
67. Dantcheva A., Velardo C., D'Angelo A., Dugelay J. L. Bag of soft biometrics

for person identification: New trends and challenges / A. Dantcheva, C. Velardo, A. D'Angelo, M. Wang, J. L. Dugelay // Multimedia Tools and Applications.– 2011. – Vol. 51. – №. 2. – C. 739-777

68. Rhodes H. T. F. Alphonse Bertillon, Father of Scientific Detection / H. T. F. Rhodes – New York. – 1956. – 83 c
69. Jain A. K., Dass S. C., Nandakumar K. Soft biometric traits for personal recognition systems/ A. K. Jain, S. C. Dass, K. // Nandakumar International conference on biometric authentication. – Berlin, Heidelberg : Springer Berlin Heidelberg .– 2004. – Vol. 51. – №. 2. – C. 731-738
70. Golovanov A.L. Development of an authentication system based on keyboard handwriting based on free texts. In: Mathematical and Computer Modeling / A. L. Golovanov // Collection of Materials of the XI International Scientific Conference Dedicated to the Memory of V.A. Romankov .– 2024. – Vol. 67. – №. 3. – C. 236-237
71. Yamali D.D. Research of systems for user identification based on the analysis of keyboard handwriting / M. M. Satybaldieva // Scientific Aspect.– 2024. – Vol. 14. – №. 5. – C. 1897-1903
72. Yamali D.D. Revolution in authentication through keyboard handwriting / D. D. Yamali // Research Center "Technical Innovations".– 2024. – Vol. 23. – №. 1. – C. 114-119
73. Polous K.I. Comparative analysis of biometric authentication methods Society. Youth. Society / K. I. Polous // Modern Science, Technology and Innovation .– 2021. – Vol. 20. – №. 1. – C. 61-63
74. Larionov M.Y. Prospects for the development of biometric identification and authentication of personality / M. Y. Larionov // Modern Science, Technology and Innovation .– 2021. – Vol. 42. – №. 1. – C. 897-902
75. Resnikoff J. The paradox of automation: QWERTY and the neuter keyboard / J. Resnikoff // Labor. – 2021. – Vol. 4. – №. 1. – C. 9-39
76. Yakovlev V.A., Skachkova V.V. Automatic Selection of Graphical Materials for Authentication System Based on a Graphical Password / V. A. Yakovlev,

V. V. Skachkova // Computer Systems.. – 2015. – Vol. 1. – №. 2. – C. 64-73

77. Janakiraman R., Sim T. Automatic Selection of Graphical Materials for Authentication System Based on a Graphical Password / R. Janakiraman, T. Sim // In Advances in Biometrics: International Conference .– 2007. – C. 584-593
78. Yousef M. A. A. A. Biometric and behavioral authentication and soft biometrics using keystroke and mouse dynamics / M. A. A. A. Yousef // АПИНО 2023.– 2023. – C. 70-75
79. Mondal S., Bours P. Continuous authentication using mouse dynamics / S. Mondal, P. Bours // International conference of the BIOSIG special interest group (BIOSIG) . – 2013. – C. 1-12
80. Bours P. Continuous authentication using mouse dynamics / P. Bours // Information Security Technical Report . – 2012. – Vol. 17. – №. 1-2. – C. 36-43
81. Krasov A. V., Alyotum Y., Ushakov I. A., Maksimov V. V., Arkhipov A. V. User authentication and identification using biometric keystroke dynamics based on the “Manhattan and Euclidean distance”/ A. V. Krasov, Y. Alyotum,I. A. Ushakov, V. V. Maksimov, A. V. Arkhipov // Vestnik of St. Petersburg State University of Technology and Design. – 2023. – Vol. 4. – №. 1. – C. 49-56
82. Idrus S. Z. S., Cherrier E., Rosenberger C., Bours P. Soft biometrics for keystroke dynamics: Profiling individuals while typing passwords / S. Z. S. Idrus, E. Cherrier,I. C. Rosenberger, P. Bours // Computers & Security.. – 2014. – Vol. 45. – №. 1. – C. 147-155
83. Rmisheva Y. D., Omelchenko T. A. Separate results of the application of the software authentication tool by keystroke dynamics / Y. D. Rmisheva, T. A. Omelchenko // NBI Technologies. – 2023. – Vol. 17. – №. 1. – C. 11-16
84. Batskikh A. V., Drovnikova I. G., Rogozin E. A. On the issue of using a new information technology related to additional authentication of access subjects using keyboard handwriting / A. V. Batskikh, I. G. Drovnikova // The

Bulletin of Voronezh. – 2020. – Vol. 2. – №. 1. – C. 21-33

85. Rajarajeswari S., Karthik K. N., Divyasri. K., Anvith, Singhal. R. Keystroke Dynamics-Based Analysis and Classification of Hand Posture Using Machine Learning Techniques / S. Rajarajeswari, K. N. Karthik, K. Divyasri, Anvith, R. Singhal // International Conference on Data Science and Network Engineering.. – 2024.– C. 57-69
86. Ahmed A. A. E., Traore I. Anomaly intrusion detection based on biometrics / A. A. E. Ahmed, I. Traore // Proceedings from the sixth annual IEEE SMC information assurance workshop. – 2005. – C. 452-453
87. Mondal S., Bours P. Continuous Authentication using Behavioural Biometrics/ S. Mondal, P. Bours // Collaborative European Research Conference (CERC) – 2013. – C. 1-12
88. Katerina T., Nicolaos P. Mouse behavioral patterns and keystroke dynamics in End-User Development: What can they tell us about users' behavioral attributes?.Computers in Human Behavior / T. Katerina, P. Nicolaos // Computers in Human Behavior. – 2018. – Vol. 83. – №. 1. – C. 288-305
89. Ahmed A. A. E., Traore I. A New Biometrics Technology based on Mouse Dynamics / A. A. E. Ahmed, I. Traore // IEEE Transactions on dependable and secure computing. – 2007.– C. 165-179
90. Kaixin W., Hongri L., Bailing W., Shujie H., Jia S. A User Authentication and Identification Model Based on Mouse Dynamics. Association for Computing Machinery / W. Kaixin, L. Hongri, W. Bailing, H. Shujie, S. Jia // Proceedings of the 6th International Conference on Information Engineering. – 2017. – C. 1-6
91. BALAGANESH P., SONIYA. A A Survey Of Authentication Based On Mouse Behaviours / P. BALAGANESH, A. SONIYA . International Journal of Advanced Information Science and Technology (IJAIST). – 2014. – Vol. 22. – №. 22. – C. 42-45
92. Hinbarji Z., Albatal R., Gurrin C. Dynamic User Authentication Based on Mouse Movements Curve / Z. Hinbarji, R. Albatal, C. Gurrin // MultiMedia

Modeling: 21st International Conference. – 2015. – Vol. 61. – №. 6. – C. 111-122

93. Gaikwad J., Kulkarni B., Phadol N., Sarukte S. Exploring visual representations of computer mouse movements for bot detection using deep learning approaches / Z. Gaikwad, R. Kulkarni, C. Phadol, S. Sarukte // Expert Systems With Applications. – 2023. – Vol. 229. – №. 22. – C. 2450-2469.
94. Коржик, В. И. Основы криптографии / В. И. Коржик, В. А. Яковлев. – Санкт-Петербург : Общество с ограниченной ответственностью "Издательский центр "Интермедия", 2016. – 296 с. – ISBN 978-5-89160-097-3. – EDN WEQWMN.
95. Abdurakhimov B., Boykuziyev I., Abdurazzokov J. Encryption systems and the history of their development //Scientific Collection «InterConf+». – 2022. – Vol. 95. №. 18. – C. 768-776.
96. Singh K. J., Manimegalai R. Evolution of encryption techniques and data security mechanisms/ K. J. Singh, R. Manimegalai //World Applied Sciences Journal. – 2015. – Vol. 33. – №. 10. – C. 1597-1613.
97. Lamport L. A User Authentication and Identification Model Based on Mouse Dynamics. Association for Computing Machinery / L. Lamport // Communications of the ACM. – 1981. – Vol. 24. – №. 11. – C. 770-772
98. Lone S. A., Mir A. H. A stable and secure one-time-password generation mechanism/ S. A. Lone, A. H. A. Mir //Journal of Advanced Research in Dynamical and Control Systems. – 2019. – Vol. 11. – №. 6. – C. 1187-1196.
99. Ma S. et al. An empirical study of sms one-time password authentication in android apps/S. Ma //Proceedings of the 35th annual computer security applications conference. – 2019. – C. 339-354.
100. Almeida L. E. et al. One-Time Passwords: A Literary Review of Different Protocols and Their Applications/ L. E. Almeida //International Conference on Advanced Research in Technologies, Information, Innovation and Sustainability. – Cham : Springer Nature Switzerland, 2023. – C. 205-219.

101. Aravindhan K., Karthiga R. R. One time password: A survey/ K. Aravindhan, R. R . Karthiga //International Journal of Emerging Trends in Engineering and Development. – 2013. – Vol. 1. – №. 3. – С. 613-623.
102. Kumar D. One Time Password Security Security System/ D. Kumar //International Journal for Advance Research and Development. – 2017. – Vol. 2. – №. 6. – С. 60-65.
103. Erdem E., Sandıkkaya M. T. OTPaaS—One time password as a service/ E. Erdem, M. T. Sandıkkaya //IEEE Transactions on Information Forensics and Security. – 2018. – Vol. 14. – №. 3. – С. 743-756.
104. Huang C. Y., Ma S. P., Chen K. T. Using one-time passwords to prevent password phishing attacks / C. Y. Huang, S. P. Ma, K. T. Chen //Journal of Network and Computer Applications. – 2011. – Vol. 34. – №. 4. – С. 1292-1301.
105. Uma M., Padmavathi G. A survey on various cyber attacks and their classification/M. Uma, G. Padmavathi //Int. J. Netw. Secur. – 2013. – Vol. 15. – №. 5. – С. 390-396.
106. Lastdrager E. E. H. Achieving a consensual definition of phishing based on a systematic review of the literature/ E. E. H. Lastdrager //Crime Science. – 2014. – Vol. 3. – С. 1-10.
107. Jakobsson M., Myers S. (ed.). Phishing and countermeasures: understanding the increasing problem of electronic identity theft / M. Jakobsson, S. Myers. – Canda. – 2007. – 700 с
108. APWG. Phishing Activity Trends Report: 3rd Quarter 2024. 2024. Available online:https://docs.apwg.org/reports/apwg_trends_report_q3_2024.pdf?_gl=1*1q3gpys*_ga*MTY2OTIzNjgwNi4xNzM2NzEzMDM2*_ga_55RF0RHXSR*MTczNjcyMjM3MC4yLjEuMTczNjcyMzM1Ny4wLjAuMA.
109. Total costs of phishing. Hoxhunt [Электронный ресурс] // What are the top 10 costs of phishing. – 2023. Режим доступа: https://hoxhunt.com/blog/what-are-the-top-10-costs-of-phishing
110. National cyber security Center Jordan. SafeOnline [Электронный ресурс]

// Tips to Protect Your Business from Phishing Attacks. – 2024. Режим доступа: https://safeonline.jo/AR/ListDetails

111. Newsroom. Stanford University IT [Электронный ресурс] // Launches Phishing Awareness Service. – 2016. Режим доступа: https://uit.stanford.edu/news/university-it-launches-phishing-awareness-service
112. Buza K. Person Identification Based on Keystroke Dynamics: Demo and Open Challenge / k. Buza //CAiSE Forum. – 2016. – Vol. 4. – C. 161-168.
113. Varshney G., Misra M., Atrey P. K. A survey and classification of web phishing detection schemes/ G. Varshney, M. Misra, P. K. Atrey //Security and Communication Networks. – 2016. – Vol. 9. – №. 18. – C. 6266-6284.
114. Masri R., Aldwairi M. Automated malicious advertisement detection using virustotal, urlvoid, and trendmicro/ R. Masri, M. Aldwairi //2017 8th International Conference on Information and Communication Systems (ICICS). – IEEE, 2017. – C. 336-341.
115. Jain A. K., Gupta B. B. Towards detection of phishing websites on client-side using machine learning based approach/ A. K. Jain, B. B. Gupta //Telecommunication Systems. – 2018. – Vol. 68. – C. 687-700.
116. Jain A. K., Gupta B. B. Towards detection of phishing websites on client-side using machine learning based approach/ A. K. Jain, B. B. Gupta //Telecommunication Systems. – 2018. – Vol. 68. – C. 687-700.
117. Mao J. et al. BaitAlarm: detecting phishing sites using similarity in fundamental visual features/J. Mao //2013 5th international conference on intelligent networking and collaborative systems. – IEEE, 2013. – C. 790-795.
118. Tirfe D., Anand V. K. A survey on trends of two-factor authentication/ D. Tirfe , V. K. Anand //Contemporary Issues in Communication, Cloud and Big Data Analytics: Proceedings of CCB 2020. – Springer Singapore, 2022. – C. 285-296.
119. Costigan N. The growing pain of phishing: is biometrics the cure? /N.

Costigan //Biometric Technology Today. – 2016. – Т. 2016. – №. 2. – С. 8-11.

120. Alsultan A., Warwick K. User-friendly free-text keystroke dynamics authentication for practical applications/ A. Alsultan, K. Warwick //2013 IEEE International Conference on Systems, Man, and Cybernetics. – IEEE, 2013. – С. 4658-4663.
121. Basu S., Islam S. K. H. Quantum-attack-resilience OTP-based multi-factor mutual authentication and session key agreement scheme for mobile users/ S. Basu, S. K. H. Islam //Computers and Electrical Engineering. – 2024. – Vol. 119. – С. 1120–1141.
122. Costa L. F. Further generalizations of the Jaccard index/ L. F. Costa //arXiv preprint arXiv:2110.09619. – 2021. – Vol. 4. – С. 1-15.
123. The free encyclopedia. Wikipedia [Электронный ресурс] // Jaccard index. – 2025. Режим доступа: https://en.wikipedia.org/wiki/Jaccard_index
124. Travieso G., Benatti A., Costa L. F. An Analytical Approach to the Jaccard Similarity Index/G. Travieso, A. Benatti, L. F. Costa //arXiv preprint arXiv:2410.16436. – 2024. – Vol. 9. – С. 1-17.
125. Fletcher S. et al. Comparing sets of patterns with the Jaccard index/ S. Fletcher //Australasian Journal of Information Systems. – 2018. – Vol. 119. – С. 1-17
126. Documentation Group. Phpnet [Электронный ресурс] // Explain Function strlen. – 2023. Режим доступа: https://www.php.net/manual/ru/function.strlen.php
127. Documentation Group. Phpnet [Электронный ресурс] // Explain Function str_repeat. – 2025. Режим доступа: https://www.php.net/manual/en/function.str-repeat.php
128. Documentation Group. Phpnet [Электронный ресурс] // Explain Function str_shuffle. – 2025. Режим доступа: https://www.php.net/manual/ru/function.str-shuffle.php
129. Sudana I. M., Qudus N., Prasetyo S. E. Implementation of PHPMailer with

SMTP protocol in the development of web-based e-learning prototype/I. M. Sudana, N. Qudus, S. E. Prasetyo //Journal of physics: Conference series. – IOP Publishing, 2019. – Vol. 1321. – №. 3. – С. 1-8.

130. Ranjith S. A REVIEW BASED ON IMPLEMENTATION OF SMTP STANDARD PROTOCOL USING PHPMAILER CLASS IN DYNAMIC WEB APPLICATIONS OVER NORMAL PHP MAIL FUNCTIONS /S. Ranjith //Journal of the Maharaja Sayajirao University of Baroda, 2019. – Vol. 60. – №. 3. – С. 97-102.
131. Yampolskiy R. V., Govindaraju V. Behavioural biometrics: a survey and classification/ R. V. Yampolskiy, V. Govindaraju //International Journal of Biometrics. – 2008. – Vol. 1. – №. 1. – С. 81-113.
132. Shen C. et al. User authentication through mouse dynamics/ C. Shen //IEEE Transactions on Information Forensics and Security. – 2013. – Vol. 8. – №. 1. – С. 16-30.
133. Antal M., Fejér N. Mouse dynamics based user recognition using deep learning/M. Antal, N. Fejér //Acta Universitatis Sapientiae, Informatica. – 2020. – Vol. 12. – №. 1. – С. 39-50.
134. Antal M., Fejér N. Mouse dynamics based user recognition using deep learning/M. Antal, N. Fejér //Acta Universitatis Sapientiae, Informatica. – 2020. – Vol. 12. – №. 1. – С. 39-50.
135. Jaiswal A. K., Tiwari P., Hossain M. S. Predicting users' behavior using mouse movement information: an information foraging theory perspective/ A. K. Jaiswal, P. Tiwari, M. S. Hossain //Neural Computing and Applications. – 2023. – Vol. 35. – №. 33. – С. 23767-23780.
136. Mason J. et al. An investigation of biometric authentication in the healthcare environment/ J. Mason //Array. – 2020. – Vol. 8. – С. 1-8.
137. Kratky P., Chuda D. Recognition of web users with the aid of biometric user model/ P. Kratky, D. Chuda //Journal of Intelligent Information Systems. – 2018. – Vol. 51. – С. 621-646.
138. Ahmed A. A. E., Traore I. A new biometric technology based on mouse

dynamics/ A. A. E. Ahmed, I. Traore //IEEE Transactions on dependable and secure computing. – 2007. – Vol. 4. – №. 3. – С. 165-179.

139. Strecker S., Van Haaften W., Dave R. An analysis of IoT cyber security driven by machine learning/S. Strecker, W. Van Haaften, R. Dave //Proceedings of International Conference on Communication and Computational Technologies: ICCCT 2021. – Springer Singapore, 2021. – С. 725-753.
140. Ackerson J. M., Dave R., Seliya N. Applications of recurrent neural network for biometric authentication & anomaly detection/J. M. Ackerson, R. Dave, N. Seliya //Information. – 2021. – Vol. 12. – №. 7. – С. 272.
141. Gunn D. J. et al. Touch-based active cloud authentication using traditional machine learning and LSTM on a distributed tensorflow framework/ D. J. Gunn //International Journal of Computational Intelligence and Applications. – 2019. – Vol. 18. – №. 04. – С. 1950022.
142. Wei A., Zhao Y., Cai Z. A deep learning approach to web bot detection using mouse behavioral biometrics/ A. Wei, Y. Zhao, Z. Cai//Biometric Recognition: 14th Chinese Conference, CCBR 2019, Zhuzhou, China, October 12–13, 2019, Proceedings 14. – Springer International Publishing, 2019. – С. 388-395.
143. Siddiqui N., Pryor L., Dave R. User authentication schemes using machine learning methods—a review/N. Siddiqui, L. Pryor, R. Dave //Proceedings of International Conference on Communication and Computational Technologies: ICCCT 2021. – Springer Singapore, 2021. – С. 703-723.
144. Zhang L. et al. Voicelive: A phoneme localization based liveness detection for voice authentication on smartphones/L. Zhang //Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. – 2020. – С. 1080-1091.
145. Wildes R. P. Iris recognition: an emerging biometric technology/R. P. Wildes //Proceedings of the IEEE. – 1997. – Vol. 85. – №. 9. – С. 1348-1363.

146. Prakash A., Krishnaveni R., Dhanalakshmi R. Continuous user authentication using multimodal biometric traits with optimal feature level fusion /A. Prakash, R. Krishnaveni, R. Dhanalakshmi // International Journal of Biomedical Engineering and Technology. – Springer Singapore, 2020. – Vol. 34. – №. 1. – С. 1-19
147. Teh P. S. et al. A survey on touch dynamics authentication in mobile devices/P. S. Teh //Computers & Security. – 2016. – Vol. 59. – С. 210-235.
148. Ayotte B. et al. Group leakage overestimates performance: A case study in keystroke dynamics/B. Ayotte //Proceedings of the ieee/cvf conference on computer vision and pattern recognition. – 2021. – С. 1410-1417.
149. Bhana B., Flowerday S. Passphrase and keystroke dynamics authentication: Usable security/B. Bhana, S. Flowerday //Computers & Security. – 2020. – Vol. 96. – С. 101925.
150. Shen C., Cai Z., Guan X. Continuous authentication for mouse dynamics: A pattern-growth approach/ C. Shen, Z. Cai, X. Guan//IEEE/IFIP international conference on dependable systems and networks (DSN 2012). –2012. – С. 1-12.
151. Antal M., Fejér N., Buza K. SapiMouse: Mouse dynamics-based user authentication using deep feature learning/M. Antal, N. Fejér, K. Buza //2021 IEEE 15th International Symposium on Applied Computational Intelligence and Informatics (SACI). – 2021. – С. 61-66.
152. Jaiswal A. K., Tiwari P., Hossain M. S. Predicting users' behavior using mouse movement information: an information foraging theory perspective/A. K. Jaiswal, P. Tiwari, M. S. Hossain //Neural Computing and Applications. – 2023. – Vol. 35. – №. 33. – С. 23767-23780.
153. Traore I. et al. Combining mouse and keystroke dynamics biometrics for risk-based authentication in web environments/ I. Traore //2012 fourth international conference on digital home. –2012. – С. 138-145.
154. Data Breach Investigations 2024 Report. [Электронный ресурс] // Режим доступа:

https://www.verizon.com/business/engb/resources/reports/204/dbir/2024-dbir-data-breach-investigations-report.pdf.

155. Reports in Cyber Security. the Identity Theft Resource Center [Электронный ресурс] // Report on the rise in user data theft 2024. – 2024. Режим доступа: https://www.idtheftcenter.org/post/q3-2024-data-breach-report-record-year-unlikely/
156. Report to the Nations. ACFE [Электронный ресурс] //Organizations Lost an Average of More Than $1.5M Per Fraud CaseExplain Function str_shuffle. – 2024. Режим доступа: https://www.acfe.com/about-the-acfe/newsroom-for-media/press-releases/press-release-detail?s=2024-Report-to-the-Nations
157. Cybersecurity News. Security Magazine [Электронный ресурс] //Average Business User Has 191 Passwords. – 2017. Режим доступа: https://www.securitymagazine.com/articles/88475-average-business-user-has-191-passwords
158. Growth Report. Pymnts Intelligence [Электронный ресурс] //Consumers Like Biometrics More Than Passwords. – 2023. Режим доступа: https://www.pymnts.com/authentication/2023/consumers-like-biometrics-more-than-passwords/
159. Thomas P. A., Preetha Mathew K. A broad review on non-intrusive active user authentication in biometrics/ P. A. Thomas, K. Preetha Mathew //Journal of Ambient Intelligence and Humanized Computing. – 2023. – Vol. 14. – №. 1. – С. 339-360
160. Zhang K. et al. Joint face detection and alignment using multitask cascaded convolutional networks/K. Zhang //IEEE signal processing letters. – 2016. – Vol. 23. – №. 10. – С. 1499-1503.
161. Smith-Creasey M., Albalooshi F. A., Rajarajan M. Continuous face authentication scheme for mobile devices with tracking and liveness detection/ M. Smith-Creasey, F. A. Albalooshi, M. Rajarajan //Microprocessors and Microsystems. – 2018. – Vol. 63. – С. 147-157.

162. JK J. R. Eye tracking in human-computer interaction and usability research: ready to deliver the promises (section commentary)/J. R. JK //The mind's eye: cognitive and applied aspects of eye movement research. – 2003. – С. 573-605.
163. Cheung W., Vhaduri S. Continuous authentication of wearable device users from heart rate, gait, and breathing data/W. Cheung, S. Vhaduri //2020 8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob). – IEEE, 2020. – С. 587-592.
164. Sayed B. et al. Biometric authentication using mouse gesture dynamics/ B. Sayed //IEEE systems journal. – 2013. – Vol. 7. – №. 2. – С. 262-274.
165. uraishi S. J., Bedi S. S. Secure system of continuous user authentication using mouse dynamics/S. J. Quraishi, S. S. Bedi //2022 3rd International Conference on Intelligent Engineering and Management (ICIEM). – IEEE, 2022. – С. 138-144.
166. Chen L. et al. Continuous authentication based on user interaction behavior/ L. Chen //2019 7th International Symposium on Digital Forensics and Security (ISDFS). – IEEE, 2019. – С. 1-6.
167. Mondal S., Bours P. Continuous authentication in a real world settings/ S. Mondal, P. Bours//2015 eighth international conference on advances in pattern recognition (ICAPR). – IEEE, 2015. – С. 1-6.
168. Carrillo C. M. Continuous biometric authentication for authorized aircraft personnel: A proposed design / C. M. Carrillo – California. – 2003. – 114 с
169. Crouse D. et al. Continuous authentication of mobile user: Fusion of face image and inertial measurement unit data/D. Crouse //2015 International Conference on Biometrics (ICB). –2015. – С. 135-142.
170. Derman E., Salah A. A. Continuous real-time vehicle driver authentication using convolutional neural network based face recognition/E. Derman, A. A. Salah //2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). – 2018. – С. 577-584.
171. Janakiraman R. et al. Using continuous face verification to improve desktop

security/R. Janakiraman //2005 Seventh IEEE Workshops on Applications of Computer. – 2005. – Vol. 1. – C. 501-507.

172. Wang M., Abbass H. A., Hu J. Continuous authentication using EEG and face images for trusted autonomous systems/M. Wang, H. A. Abbass, J. Hu //2016 14th Annual Conference on Privacy, Security and Trust (PST). – 2016. – C. 368-375.
173. Agrafioti F., Bui F. M., Hatzinakos D. Secure telemedicine: Biometrics for remote and continuous patient verification/F. Agrafioti, F. M. Bui, D. Hatzinakos //Journal of Computer Networks and Communications. – 2012. – Vol. 2012. – №. 1. – C. 924791.
174. Flior E., Kowalski K. Continuous biometric user authentication in online examinations/ E. Flior, K. Kowalski//2010 seventh International Conference on information technology: new generations. – 2010. – C. 488-492.
175. Al Abdulwahid A. et al. Continuous and transparent multimodal authentication: reviewing the state of the art/A. Al Abdulwahid //Cluster Computing. – 2016. – Vol. 19. – C. 455-474.
176. Stylios I. C. et al. A review of continuous authentication using behavioral biometrics/ I. C. Stylios //Proceedings of the SouthEast European Design Automation, Computer Engineering, Computer Networks and Social Media Conference. – 2016. – C. 72-79.
177. Mahfouz A., Mahmoud T. M., Eldin A. S. A survey on behavioral biometric authentication on smartphones/A. Mahfouz, T. M. Mahmoud, A. S. Eldin //Journal of information security and applications. – 2017. – Vol. 37. – C. 28-37.
178. Stanić M. Continuous user verification based on behavioral biometrics using mouse dynamics/ M. Stanić //Proceedings of the ITI 2013 35th International Conference on Information Technology Interfaces. – 2013. – C. 251-256.
179. Shen C., Cai Z., Guan X. Continuous authentication for mouse dynamics: A pattern-growth approach/C. Shen, Z. Cai, X. Guan //IEEE/IFIP international conference on dependable systems and networks (DSN 2012). – 2012. – C.

1-12.

180. Fridman L. et al. Multi-modal decision fusion for continuous authentication/ L. Fridman //Computers & Electrical Engineering. – 2015. – Vol. 41. – C. 142-156.
181. Bours P., Mondal S. Performance evaluation of continuous authentication systems/ P. Bours, S. Mondal //Iet Biometrics. – 2015. – Vol. 4. – №. 4. – C. 220-226.
182. Centeno M. P., Guan Y., van Moorsel A. Mobile based continuous authentication using deep features/ M. P. Centeno,Y. Guan, A. van Moorsel //Proceedings of the 2nd international workshop on embedded and mobile deep learning. – 2018. – C. 19-24.
183. Jorquera Valero J. M. et al. Improving the security and QoE in mobile devices through an intelligent and adaptive continuous authentication system/ J. M. Jorquera Valero //Sensors. – 2018. – Vol. 18. – №. 11. – C. 3769.
184. Feher C. et al. User identity verification via mouse dynamics/ C. Feher //Information Sciences. – 2012. – Vol. 201. – C. 19-36.
185. Feng T. et al. Continuous mobile authentication using touchscreen gestures/ T. Feng //2012 IEEE conference on technologies for homeland security (HST). – IEEE, 2012. – C. 451-456.
186. Roy A., Halevi T., Memon N. An HMM-based behavior modeling approach for continuous mobile authentication //2014 IEEE international conference on acoustics, speech and signal processing (ICASSP). – IEEE, 2014. – C. 3789-3793.
187. Fridman L. et al. Active authentication on mobile devices via stylometry, application usage, web browsing, and GPS location/L. Fridman //IEEE Systems Journal. – 2016. – Vol. 11. – №. 2. – C. 513-521.
188. Digital Businesses Combine. TechTarget and Informa [Электронный ресурс] //What is ASCII (American Standard Code for Information Interchange). – 2025. Режим доступа:

https://www.techtarget.com/whatis/definition/ASCII-American-Standard-Code-for-Information-Interchange

189. Szabo F. The Linear Algebra Survival Guide: Illustrated with Mathematica / F. Szabo – Montreal, Canada. – 2015. – 417 с.
190. Dokmanic I. et al. Euclidean distance matrices: essential theory, algorithms, and applications/ I. Dokmanic //IEEE Signal Processing Magazine. – 2015. – Vol. 32. – №. 6. – С. 12-30.
191. Duarte F. S. L. G. et al. Decomposing time series into deterministic and stochastic influences: A survey/ F. S. L. G. Duarte //Digital Signal Processing. – 2019. – Т. 95. – С. 102582.
192. Metcalf L., Casey W. Cybersecurity and Applied Mathematics/ L. Metcalf, W. Casey – USA. – 2016. – 240 с.
193. Awrejcewicz J. Particle Kinematics and an Introduction to the Kinematics of Rigid Bodies/J. Awrejcewicz – USA. – 2012. – 262 с.
194. Haldar R., Mukhopadhyay D. Levenshtein distance technique in dictionary lookup methods: An improved approach //arXiv preprint arXiv:1101.1232. – 2011. – Vol. 17. – №. 2. – С. 1-5.
195. Aldoukali M. B, Elburase E. A. Using Levenshtein Distance Algorithm to Increase Database Search Efficiency and Accuracy/ M. B. Aldoukali, E. A. Elburase // Pure and Applied Sciences. – 2022. – Vol. 6. – №. 10. – С. 15-22.
196. The free encyclopedia. Wikipedia [Электронный ресурс] //Brief biography of Vladimir Levenshtein. – 2024. Режим доступа: https://en.wikipedia.org/wiki/Vladimir_Levenshtein
197. The free encyclopedia. Wikipedia [Электронный ресурс] // Wagner–Fischer algorithm. – 2024. Режим доступа: https://en.wikipedia.org/wiki/Wagner%E2%80%93Fischer_algorithm
198. Nguyễn T. H. Automata Technique for The LCS Problem/T. H. Nguyễn //Journal of Computer Science and Cybernetics. – 2019. – Vol. 35. – №. 1. – С. 21-37.

ПРИЛОЖЕНИЕ 1. Программа ЭВМ

РОССИЙСКАЯ ФЕДЕРАЦИЯ

СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2024664769

Программа по многофакторной аутентификации пользователей на основе биометрических динамических методов

Правообладатель: ***Федеральное государственное бюджетное образовательное учреждение высшего образования «Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича» (RU)***

Авторы: ***Альотум Юсеф Мохаммед Абд Аллх (JO), Пешков Андрей Иванович (RU)***

Заявка № **2024663318**
Дата поступления **11 июня 2024 г.**
Дата государственной регистрации
в Реестре программ для ЭВМ ***24 июня 2024 г.***

Руководитель Федеральной службы
по интеллектуальной собственности

Ю.С. Зубов