

АЛГОРИТМЫ РАСПОЗНАВАНИЯ КЛАВИАТУРНОГО ПОЧЕРКА ДЛЯ СИСТЕМ БИОМЕТРИЧЕСКОЙ АУТЕНТИФИКАЦИИ

Пулко Т.А.¹, Лах А.М.², Румас С.С.³

¹Пулко Татьяна Александровна – кандидат технических наук, доцент,
кафедра защиты информации,
УО «Белорусский государственный университет информатики и радиоэлектроники»,
²Лах Александр Максимович – учащийся,
³Румас Сергей Сергеевич – учащийся,
направления «Информационная безопасность»
УО «Национальный детский технопарк»,
г. Минск, Республика Беларусь

Аннотация: в статье рассматривается актуальность использования методов биометрической идентификации на основе методов распознавание клавиатурного почерка. Приведены результаты исследования существующих систем анализа клавиатурного почерка, рассмотрены алгоритмы распознавания клавиатурного почерка: статистический метод, метод на основе нейронных сетей, метод машинного обучения, гибридный метод. Рассмотрены этапы работы алгоритмов распознавания клавиатурного почерка для систем биометрической аутентификации.

Ключевые слова: клавиатурный почерк, биометрия, аутентификация, информационная безопасность.

KEYBOARD HANDBOOK RECOGNITION ALGORITHMS FOR BIOMETRIC AUTHENTICATION SYSTEMS

Pulko T.A.¹, Lakh A.A.², Rumas S.S.²

¹Pulko Tatyana Aleksandrovna – PhD in Technical sciences, Associate Professor,
DEPARTMENT OF INFORMATION SECURITY,
BELARUSIAN STATE UNIVERSITY OF INFORMATICS AND RADIOELECTRONICS,
²Lakh Alexander Maksimovich – student,
³Rumas Sergey Sergeevich – student
FIELD OF “INFORMATION SECURITY”
NATIONAL CHILDREN'S TECHNOPARK,
MINSK, REPUBLIC OF BELARUS

Abstract: the article discusses the relevance of using biometric identification methods based on keyboard handwriting recognition methods. The results of a study of existing systems for analyzing keyboard handwriting are presented, algorithms for recognizing keyboard handwriting are considered: a statistical method, a method based on neural networks, a machine learning method, a hybrid method. The stages of operation of keyboard handwriting recognition algorithms for biometric authentication systems are considered.

Keywords: keyboard handwriting, biometrics, authentication, information security.

УДК 004.056.53

Биометрическая аутентификация по поведенческим характеристикам предлагает надежное и удобное решение для преодоления этих проблем, связанных с обеспечением безопасности данных и идентификации пользователей в современном цифровом мире. Традиционные методы аутентификации, такие как пароли и токены, часто уязвимы для взлома и фишинга, тогда как поведенческая биометрия анализирует уникальные паттерны поведения пользователей, такие как динамика набора текста, движения мыши, использование сенсорного экрана, схема навигации и прочее, позволяя осуществлять аутентификацию пользователей автоматически на основе их поведения. Эти характеристики сложно подделать или взломать, при этом они остаются постоянными во времени, что делает их надежными идентификаторами, снижающими риск несанкционированного доступа и мошенничества. К изменяющимся паттернам поведения пользователей поведенческая биометрика может адаптироваться, обеспечивая непрерывную безопасность. Следует отметить, что поведенческие характеристики не передаются через электронную почту или текстовые сообщения, что делает их неуязвимыми для популярных в настоящее время атак фишинга. Помимо актуальности, биометрическая аутентификация по поведенческим характеристикам предлагает ряд преимуществ, связанных с аутентификацией без вмешательства пользователя, минимальными затратами на внедрение таких систем и их точностью, которая может достигать более 99%. По мере совершенствования технологий поведенческая биометрика будет играть все более важную роль в создании безопасной и бесшовной цифровой среды [1].

Одним из малоизученных, но перспективных методов биометрической идентификации является распознавание клавиатурного почерка. Этот метод основывается на том, что каждый человек печатает текст уникальным образом, используя индивидуальные движения пальцев и ритм нажатия клавиш.

Алгоритмы распознавания клавиатурного почерка обычно используют следующие этапы работы:

- сбор данных, заключающийся в записи клавиатурных нажатий, включая время нажатия, длительность нажатия и время отпускания клавиши;
- предобработка, в результате которой осуществляется очистка данных от шума и преобразование их в необходимый формат;
- извлечение признаков, например, выделение характерных особенностей нажатий, таких как скорость, сила нажатия, траектория движения пальцев;
- обучение модели, в результате которого создается модель, способная различать клавиатурный почерк разных людей, если это необходимо;
- тестирование для оценки точности работы алгоритма на новых данных [2].

Для распознавания клавиатурного почерка могут использоваться алгоритмы, основанные на различных методах и подходах. Одними из таких методов являются статистические методы, основанные на анализе распределения признаков клавиатурного почерка [3]. При этом может выполняться анализ частоты нажатия клавиш (например, некоторые люди чаще используют "а", а другие - "о"), анализ длительности нажатия (измеряется время, которое человек держит клавишу нажатой; у каждого человека есть уникальный стиль нажатия, что позволяет отличить его от других) или анализ времени между нажатиями (интервалы между нажатиями клавиш также могут быть уникальны для каждого человека). Для статистических методов характерна простота реализации, не большой объем данных, требуемый для обучения. Однако, они не могут обеспечить высокую точность данных, особенно при использовании коротких текстов, а также следует отметить их неустойчивость к изменениям в стиле набора текста, вызванным усталостью, стрессом или внешними факторами [4].

Для обучения моделей распознавания большой популярностью пользуются искусственные нейронные сети. Многослойные перцептроны (MLP) обучаются на больших объемах данных и могут выявлять сложные зависимости между различными параметрами клавиатурного почерка. Сверточные нейронные сети (CNN) используют сверточные слои для извлечения пространственных особенностей из данных, что позволяет лучше учесть временную динамику нажатий клавиш. Рекуррентные нейронные сети (RNN) подходят для обработки последовательных данных, таких как клавиатурный почерк, и могут "запомнить" предыдущие действия пользователя. Безусловно такие алгоритмы для распознавания клавиатурного почерка отличаются высокой точностью, особенно при использовании больших объемов данных для обучения, и способностью адаптироваться к индивидуальным особенностям клавиатурного почерка. При этом, они требуют больших объемов данных для обучения, трудозатратны по времени и могут быть сложными в реализации [5].

Методы машинного обучения используют различные алгоритмы машинного обучения, такие как SVM, kNN, деревья решений и др.

Алгоритм машинного обучения SVM (Support Vector Machines), характеризуется обучением на разделении данных на классы, определяя оптимальную гиперплоскость, которая максимально эффективно разделяет пользователей. Алгоритмы kNN (k-Nearest Neighbors) классифицируют новый набор данных на основе сходства с ближайшими k образцами из обучающей выборки. Также могут применяться деревья принятия решений, которые позволяют создавать древовидную структуру, разделяющую данные на основе определенных критериев, что позволяет классифицировать пользователя по его клавиатурному почерку.

Независимо от типа возможного алгоритма машинного обучения и выбранного подхода для задач распознавания клавиатурного почерка, все они будут относительно просты в реализации и вместе с тем, хорошо справляться с нелинейными зависимостями в данных. Однако, могут быть менее точными, например, чем нейронные сети и будут требовать тщательного подбора параметров для достижения оптимальной точности. Кроме этого, для реализации для задач распознавания клавиатурного почерка с применением машинного обучения требуется больше времени для обучения модели и обработки данных, чем у других методов биометрии, присутствует чувствительность к внешним факторам, а значит такие системы подвержены ошибкам из-за внешних факторов, таких как усталость, стресс, шум и др.

Наилучшим вариантом для систем биометрической аутентификации с распознаванием клавиатурного почерка могут служить гибридные методы, характеризующиеся сочетанием нескольких алгоритмов для повышения точности и устойчивости. Например, использование нейронных сетей для извлечения признаков и SVM для классификации.

Какой алгоритм распознавания клавиатурного почерка для систем биометрической аутентификации бы ни был выбран важно обеспечить в системе следующие функции:

- предварительную обработку данных, чтобы очистить данные от шума, например, от случайных нажатий клавиш;

– выбор наиболее информативных признаков клавиатурного почерка, потому что они существенно влияют на точность алгоритмов;

– учесть размер выборки вследствие того, что для достижения высокой точности требуется большой объем данных для обучения алгоритмов.

В заключение, следует отметить, что выбор оптимального алгоритма зависит от конкретной задачи и требований к точности, скорости и ресурсам. Современные алгоритмы, особенно нейронные сети, способны достигать высокой точности в распознавании клавиатурного почерка, делая этот метод перспективным для систем биометрической аутентификации.

Распознавание клавиатурного почерка в системах безопасности может быть использовано для аутентификация пользователей при входе в системы, приложения и веб-сайты, для контроля доступа к защищенным ресурсам, для мониторинга активности пользователей при обнаружении мошенничества и несанкционированного доступа.

Список литературы/References

1. Пулко Т.А., Лах А.А., Румас С.С. Актуальность биометрической аутентификации пользователей по поведенческим характеристикам // Технические средства защиты информации: тез. докл. XXII Белорусско-Российской науч.-техн. конф., Республика Беларусь, Минск, 12 июня 2024 года) – Минск, 2024. – С.72.
2. Болл Р.М. Руководство по биометрии – Москва: Техносфера, 2007. – 368 с.
3. Аутентификация пользователя по клавиатурному почерку в целях защиты средств вычислительной техники от несанкционированного доступа [Электронный ресурс]. – Режим доступа: <https://web.snauka.ru/issues/2023/11/100987>. – Дата доступа: 12.04.2024.
4. Довгаль В.А. Захват параметров клавиатурного почерка и его особенности / В.А. Довгаль // Информационные системы и технологии в моделировании и управлении: сб. материалов Всерос. науч.-практ. конф., 5–7 июня 2017, г. Симферополь : АРИАЛ, 2017. – С. 230–236.
5. Жашкова Т.В. Нейросетевая идентификация типа личности человека по клавиатурному почерку: статья / Т.В. Жашкова, О.М. Шарунова, Э.Ш. Исянова. – Международный студенческий научный вестник, №3, ч.1, 2015. – 6 с.