УДК 519.6+004.021
DOI: 10.26102/2310-6018/2024.45.2.021

Исследование поведенческой биометрии методами анализа данных и машинного обучения

И.С. Смирнов, А.А. Кочкаров✉

Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

Резюме. В статье показаны возможности применения методов машинного обучения для построения и анализа системы аутентификации на основе динамики нажатий клавиш. В работе обоснована необходимость улучшения многофакторной системы аутентификации. Предложен способ классификации работ поведенческой биометрии для сравнения и использования результатов исследований. Рассмотрены базовые возможности обработки и генерирования динамических и статических признаков динамики нажатий клавиш. Протестированы различные комбинации наборов признаков и выборок обучения, описана лучшая комбинация с равной частотой ошибок (Equal Error Rate (EER)) 4,7%. Итеративный анализ качества системы позволяет установить важность первых символов последовательности ввода, а также нелинейную взаимосвязь степени ранжирования модели и EER. Высокие показатели, достигнутые бустинговой моделью, свидетельствуют о значительном потенциале поведенческой аутентификации для дальнейшего улучшения, развития и применения. Приводится значимость данного метода, его практическая полезность не только в задаче аутентификации, перспективы развития, включая использование нейросетевых методов и анализ динамики данных. Несмотря на достигнутые результаты, отмечается необходимость дальнейшей работы над моделью, включая разработку дополнительных моделей кластеризации, классификации, изменение набора признаков и построение каскада. Подчеркивается важность исследуемой области, способной принести значительный вклад в развитие информационной безопасности и технологий.

Ключевые слова: аутентификация, поведенческая биометрия, динамика нажатий клавиш, классификация, машинное обучение.

Для цитирования: Смирнов И.С., Кочкаров А.А. Исследование поведенческой биометрии методами анализа данных и машинного обучения. *Моделирование, оптимизация и информационные технологии.* 2024;12(2). URL: https://moitvivt.ru/ru/journal/pdf?id=1596 DOI: 10.26102/2310-6018/2024.45.2.021

The study of behavioral biometrics using data analysis and machine learning methods

I.S. Smirnov, A.A. Kochkarov✉

Financial University under the Government of the Russian Federation, Moscow, the Russian Federation

Abstract. The article shows the possibilities of using machine learning methods to build and analyze an authentication system based on the dynamics of keystrokes. The paper substantiates the need to improve the multifactor authentication system. A method of classifying the work of behavioral biometrics for comparison and use of research results is proposed. The basic possibilities of processing and generating dynamic and static signs of the dynamics of keystrokes are considered. Various combinations of feature sets and training samples were tested, and the best combination with an Equal Error Rate (EER) of 4.7% was described. An iterative analysis of the quality of the system allows us to establish the importance of the first characters of the input sequence, as well as the nonlinear relationship between the degree of ranking of the model and EER. The high performance achieved by the boosting model indicates the significant potential of behavioral authentication for further improvement, development and application.

The significance of this method, its practical usefulness not only in the task of authentication, development prospects, including the use of neural network methods and data dynamics analysis are presented. Despite the achieved results, there is a need for further work on the model, including the development of additional clustering, classification models, changing the set of features and building a cascade. The importance of the research area, which can make a significant contribution to the development of information security and technology, is emphasized.

Keywords: authentication, behavioral biometrics, keystroke dynamics, classification, machine learning.

For citation: Smirnov I.S., Kochkarov A.A. *The study of behavioral biometrics using data analysis and machine learning methods, Optimization and Information Technology*. 2024;12(2). URL: https://moitvivt.ru/ru/journal/pdf?id=1596 DOI: 10.26102/2310-6018/2024.45.2.021 (In Russ.).

Введение

Поведенческая биометрия – это совокупность динамических характеристик человека, достаточных для его однозначного определения. К поведенческой биометрии относится «стиль» взаимодействия пользователя с периферийными устройствами. В ранних исследованиях конца XX в. было утверждено, что «цифровой след», способ обращения человека с устройством, уникален, как и обычный стиль письма [1].

Свойство уникальности делает поведенческую биометрию подходящим решением для задачи аутентификации. Аутентификация – проверка информации, предоставленной пользователем, с той, что уже зарегистрирована в базе данных, это действие сравнимо с «наложением данных один к одному». Поведенческая аутентификация используется повсеместно во время прохождения капчи: на основании того, как пользователь взаимодействует с интерфейсом во время ее прохождения, модель определяет, является ли он ботом или нет [2].

Информация является одним из важнейших современных ресурсов [3], а её конфиденциальность – необходимым условием, обеспечиваемым высоким уровнем безопасности системы. Методы аутентификации, направленные на защиту доступа к личным данным, постоянно совершенствуются и дополняют друг друга, обеспечивая практически полную гарантию безопасности при правильном использовании. При этом, из-за человеческого фактора, конфиденциальность информации часто нарушается третьими лицами, не имеющими правомерного доступа.

Модель аутентификации на основе поведенческой биометрии может стать эффективным и удобным способом дополнительной нативной проверки пользователя. Актуальность тематики подтверждается результатами исследований в этой области. Однако нет общепринятых правил, положений и моделей использования поведенческой биометрии для аутентификации, в частности, в русскоязычном сегменте практически отсутствуют работы на данную тематику.

Таким образом, цель исследования – анализ и классификация существующих подходов и решений, модельное и статистическое изучение данных и закономерностей поведенческой биометрии, разработка системы на основе подтвержденных гипотез для аутентификации пользователя. Это позволит не только структурировать знания о текущих достижениях в данной области, создать систему типизации по месту возможного применения результатов исследования, но и продвинуться в создании системы дополнительной проверки личности, что может значительно сократить объем украденных данных.

Поведенческая биометрия и применение технологии непрерывной аутентификации

Построение системы аутентификации, на основе поведенческой биометрии является активной областью исследований, где одной из ключевых задач является выявление закономерностей во взаимодействии пользователей с устройствами, подбор и разработка различных моделей для решения этой задачи.

Существуют два основных типа аутентификации – статическая и непрерывная. Большинство исследований посвящено первому типу, так как непрерывная аутентификация – в значительной степени прикладная задача, ее решение полностью зависит от места применения результатов работы. Статическая аутентификация представляет собой одноразовую проверку пользователя с целью подтверждения его подлинности.

Научные работы классифицируются по периферийному устройству, с которым взаимодействует пользователь. Это может включать в себя физическую клавиатуру, манипулятор мышь, мобильный телефон или комбинацию устройств. Данное исследование затрагивает только аутентификацию на основе динамики нажатий клавиш – физическую клавиатуру.

Все подобные исследования зависят от опыта, с помощью которого был собран набор данных. Их сбор основан на степени экспериментального контроля. Его основные параметры – свобода выбора устройства: бренд и модель устройства могут быть заданы определенным набором [4], устройство может быть свободным, но зарегистрированным [5], а может быть произвольным и не находиться в метаданных датасета. Фиксированный или произвольный текст, как и устройство, может быть строго задан единым для всех участников или для каждого по отдельности [6], или может быть произвольным выбором самого пользователя (при сборе данных из повседневной жизни). Периодичность – для работы с главным недостатком поведенческой аутентификации, изменчивостью, данные по пользователю собираются не разово, а с заданным интервалом, в разных работах от пяти дней до четырех недель.

Вместе с этим в научных работах подчеркивается важность группы определяющих характеристик набора данных, а следовательно, и для эксперимента в целом. При широком формате исследования с публичным принципом сбора данных за счет объема выборки обеспечивается возрастное и региональное многообразие, иначе высокая дисперсия демографии создается искусственно.

Основное понятие в наборе данных для классификации – уровень целевой переменной (далее target rate) – обозначает долю положительных наблюдений среди всех наблюдений. Данные с очень низким показателем являются «проблемными» для обычных методов классификации, так как имеет высокую склонность к переобучению, что в условиях изменчивости поведения человека многократно усложняет выбор или разработку архитектуры, а также подбор гиперпараметров. Надежность тестов модели аутентификации зависит от полноты отложенной выборки, что включает в себя и повторение ввода одного и того же предложения разными пользователями, и общее количество зарегистрированных примеров. Недостаток исследований с отсутствием «массовости» – высокий target rate, что зачастую не соответствует реальности, а тем более критическим ситуациям.

С практической стороны и разработчик, и пользователь заинтересованы в том, чтобы количество символов, необходимых для успешной аутентификации было минимально достижимым. В исследовании [7] вводимый текст является PIN-кодом, соответственно, его размер устанавливается в привычных практических пределах от 4 до 8. В среднем, для устойчивого извлечения признаков из последовательности, размер

текста, необходимого для классификации, варьируется от 15 до 50 символов, в некоторых экспериментах пользователи вводили текст длиной более 500 слов.

При сборе данных выявляется наличие образцов-самозванцев. Волонтеры получают текст, который вводили другие пользователи (например, логин и пароль) и производят «попытку взлома». Регистрируются биометрические показатели [8] и на их основе создается набор данных, содержащий негативные наблюдения, на таком наборе данных тестируется устойчивость модели. В некоторых исследованиях [9] пользователи в режиме реального времени «атакуют» чужие учетные записи. Также негативных наблюдений может не быть в контексте полного сходства вводимого текста, в таких случаях сравнивается паттерн печати текста как такового.

Для решения задачи аутентификации с помощью поведенческой биометрии принято использовать три различных способа, описывающих эталонный «профиль» пользователя. Базовое решение включает в себя применение различных статистических методов – на основе рассчитанных признаков n-графов, комбинаций времени последовательного нажатия n клавиш, вычисляется расстояние между векторами истинного «профиля» и попытки аутентификации с помощью расстояния Махаланобиса, евклидового, манхэттенского и других. Такие способы проигрывают моделям машинного обучения, так как сильно ограничены в объемах выборки для обучения, а при больших размерах тестовых данных теряют в точности. Модели машинного обучения – наивный байес, SVM, K-Means, Random Forest, градиентный бустинг – не требуют высоких затрат по времени и ресурсам для обучения, способны объяснить взаимосвязь принятых моделью решений и используемых признаков, но при этом не видят рекуррентной связи в данных. Поэтому для последовательностей принято использовать модели глубокого обучения: CNN, RNN, GRU, BRNN, LSTM. На основе этих архитектур достигнуты наилучшие текущие результаты в исследуемой тематике [10].

Для отражения производительности модели аутентификации принято использовать коэффициент ложного принятия (False Acceptance Rate) – вероятность того, что модель аутентификации пропустит пользователя, не владеющего учетной записью, и коэффициент ложного отказа (False Rejection Rate), частота ложных отклонений – это вероятность того, что модель аутентификации ошибается и не пропускает в систему истинного пользователя. Равная частота ошибок (Equal Error Rate) – значение этой метрики устанавливается в равновероятном значении ошибок ложного принятия и отказа. Также используется стандартная метрика для моделей классификации – инвариантность к порогу классификации и масштабу предсказаний (ROC-AUC).

Для повышения качества системы некоторые исследователи делают акцент не на архитектуре модели, а на проверке различных гипотез, но при разработке чаще используются датасеты малого масштаба в плане уникального количества испытуемых. В исследовании [11] авторы утверждают, что самыми важными символами в последовательности являются первые и последние. Объясняется это тем, что перед тем, как начать писать текст, человек думает о том, что он собирается написать, визуализирует набор текста в виде действий, совершает меньше ошибок. В середине последовательности растет количество пауз, ошибок и, соответственно, дисперсия временных интервалов нажатий клавиш.

Авторы исследования [12] предлагают использовать нечеткую логику, хотя эта область математики является сильно неформальной. В работе [6] предлагается использовать как метрику размер выборки, необходимой для обучения. Эта метрика действительно важна – она отражает жизнеспособность модели в «боевых» условиях.

Для глубокого обучения в работе [13] предлагается не просто каскад из нескольких моделей, а концептуально иной подход к построению дополнительной

модели. Авторы разработали к основной модели еще одну, но данными для обучения для нее служат динамика нажатий, зарегистрированная на фиксированном тексте для всех пользователей. Такой подход позволил на порядок улучшить метрики, но количество волонтеров, участвовавших в эксперименте, слишком мало для аналогичного утверждения в больших масштабах.

Авторы работы [14] при анализе набора данных столкнулись с сильной изменчивостью дисперсии и предложили классифицировать пользователей на три класса: малая ошибка ERR, средняя и высокая. А далее перешли к построению классификатора и акцентировали внимании на ранжирующей способности модели. Предсказание модели классификации помогало основной модели аутентификации лучше подбирать пороги решений, что повысило ERR в каждом бакете.

В настоящем исследовании используется набор данных «136M Keystrokes Dataset» [5], который содержит регистрацию нажатий клавиш от 168 960 пользователей по 15 примеров полуфиксированных предложений для каждого. На сайтах, где в течение минуты предлагается вводить несвязные слова, размещалась реклама «альтернативного научного теста», и каждый желающий мог пройти этот тест. Базовая длина предложения варьируется от 15 до 70 символов. Дополнительно, после ввода, пользователям нужно было пройти анкетирование и указать свои метаданные.

Данные представлены в табличном виде – одна строка соответствует одной нажатой клавише. Посчитаны основные признаки задачи динамики нажатий клавиш: полное время нажатия и «времена полета» (flight time), для каждого из которых созданы описательные признаки с помощью агрегаторов – средние значения, стандартные отклонения, медианы, первый и третий квартили. Схема расчета описательных признаков продемонстрирована на Рисунке 1.

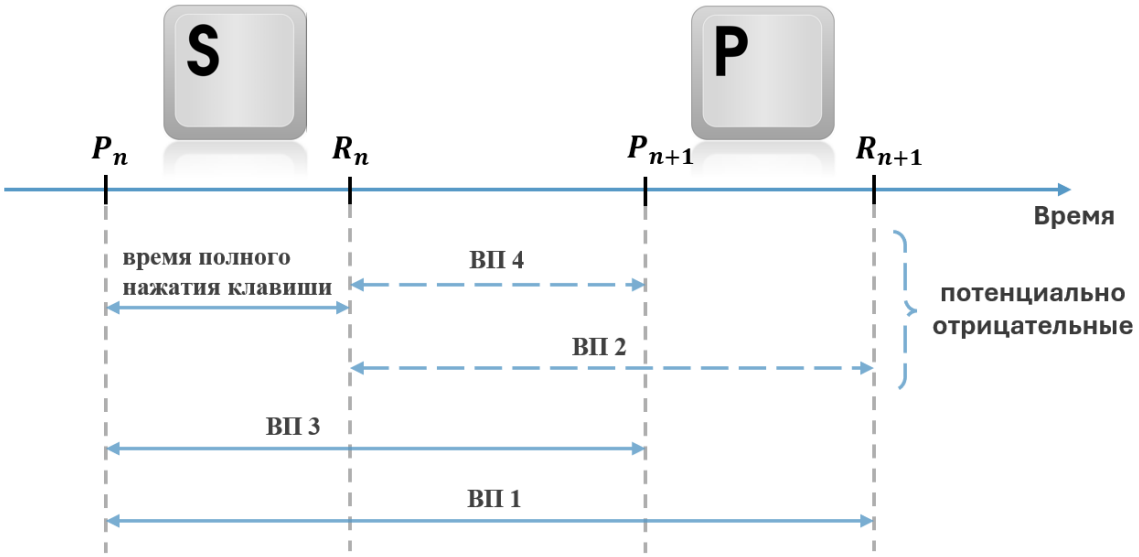

Рисунок 1 – Способ вычисления основных признаков
Figure 1 – The method of calculating the main features

Самые часто употребляемые клавиши – «пробел» и «е» – составляют более 23% от всех клавиш при печати, поэтому дополнительно сгенерированы признаки относительно двух клавиш. Однако, стиль набора «е» имеет дисперсию значительно выше, чем «пробел», медиана стандартного отклонения сильно смещена влево. Это связано с высокой зависимостью от биграммов, в которых встречается клавиша.

Для добавления динамики последовательность разбивается на три подпоследовательности – первые и последние пять символов и все, что находятся между ними. Так как минимальная длина предложения равна 15, то относительно всей выборки ограничение в пять символов – максимально возможное для сохранения стабильности статистики. На Рисунке 2 продемонстрировано отличие между распределениями среднего значения и стандартного отклонения полного времени нажатия клавиши. Медиана на втором графике смещена вправо у последовательности первых пяти символов, что подтверждает гипотезу о различии в почерке внутри одного предложения.

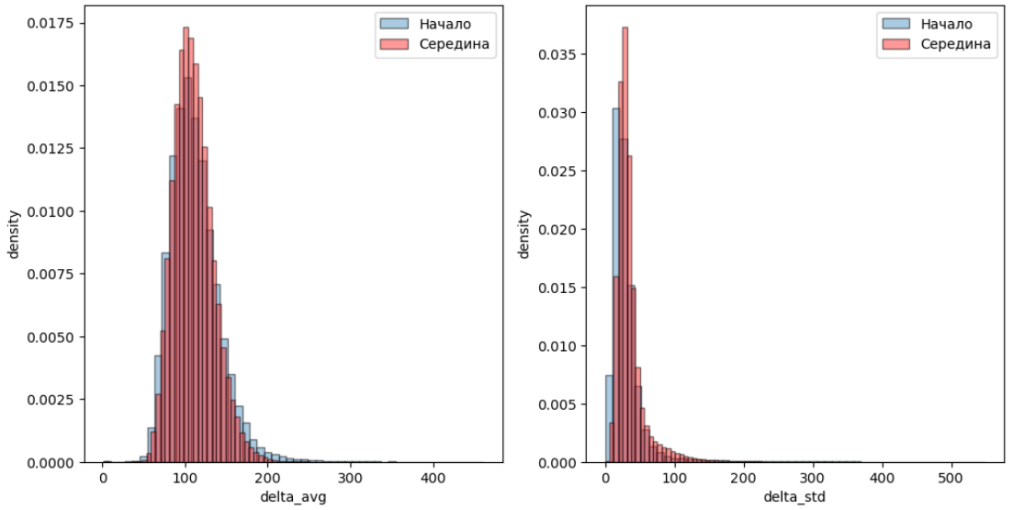

Рисунок 2 – Распределение delta на двух срезах
Figure 2 – Delta distribution on two slices

В работе [15] авторы устанавливают ограничение для признаков в 3000 мс. При аналогичном ограничении признаки имеют визуально нормальное или логнормальное распределение, но остаются сильные выбросы. За порогом трех стандартных отклонений находится 12% всех значений, что противоречит закону о распределении Гаусса-Лапласа, поэтому в работе используется ограничение 99 перцентиля относительно каждого признака.

Для построения модели поведенческой аутентификации используется градиентный бустинг, который не требует одинаковой размерности входных данных. Этот способ позволяет однозначно интерпретировать взаимосвязи и результаты за счет важности признаков, имеет высокую скорость обучения, что необходимо, так как для каждого пользователя обучается отдельная модель классификации, что позволяет проводить множественные широкие тесты. Выполнение этих критериев позволяет эффективно использовать модель для анализа динамики нажатий клавиш.

Применение градиентного бустинга для поведенческой аутентификации

Для проверки гипотез и тестирования наборов признаков и архитектур определен оптимальный размер выборки для проведения экспериментов, итеративно подобраны гиперпараметры.

Граница принятия решений выбирается отдельно для каждого пользователя и его модели. Выбор границы осуществляется шестью различными способами: среднее гармоническое точности (precision) и полноты (recall), оптимальное значение для максимального ROC-AUC, наибольшее среднее FAR и FRR для нуля, одной, двух и трех ошибок ложного отказа. Из-за ограниченного количества положительных примеров

метрика FRR всегда заметно дискретна внутри каждой модели – тестами покрываются модели с количеством таргетов от 2 до 11.

Построено базовое решение на основных признаках для всей выборки, относительно которого проводилось тестирование. Были аккумулятивно использованы 15 самых часто встречаемых клавиш, разбиение последовательности на подвыборки и различные комбинации. Усложнение вариативности признаков не дает улучшения качества из-за переобучения – естественное следствие низкого target rate для модели машинного обучения.

Лучшие средние результаты показала модель с 11 положительными примерами и дополнительным набором признаков по клавише «пробел». В Таблице 1 приведены результаты тестирования модели с 21 признаком для различных границ принятия решения и количества таргетов на обучении.

Таблица 1 – Метрики качества с признаками по пробелу
Table 1 – Quality metrics with features by space

train samples	EER f1	EER 0	EER 1	EER 2	EER 3	EER m
2	40,3%	18,8%	17,7%	18,6%	20,5%	14,4%
3	40,5%	16,4%	15,4%	16,6%	18,8%	12,0%
4	39,8%	14,4%	13,2%	14,9%	17,7%	10,2%
5	39,0%	12,5%	12,1%	14,4%	18,0%	9,0%
6	38,9%	11,2%	11,3%	14,7%	19,0%	8,1%
7	37,5%	9,7%	11,3%	15,3%	20,6%	7,5%
8	36,0%	8,3%	11,0%	16,5%	22,7%	6,6%
9	35,5%	6,8%	11,4%	18,5%	26,2%	5,7%
10	32,7%	6,3%	12,6%	21,6%	31,1%	5,5%
11	30,9%	5,1%	14,6%	26,2%	38,3%	4,7%

В зависимости от вектора исследования могут быть выбраны различные критерии определения положительного результата. При ориентире на безопасность системы и максимизацию метрики FAR следует работать со сбалансированным выбором границы – модель с EER f1 для девяти таргетов достигает FAR равный 0,39%, что является значительным улучшением по сравнению с другими вариантами, но при этом приводит к отклонению 7 из 10 попыток авторизации реального пользователя.

Метрика EER 0 отражает максимальное удобство системы, с обеспеченным гарантированным доступом реального пользователя. Вся ошибка концентрируется в ложном принятии (FAR).

Возможно достижение лучших относительных результатов с минимальным объемом обучающей выборки – EER 1 достигает оптимальной точности с 6 таргетами. Большая часть исследований направлена на достижение максимальной точности равной частоты ошибок – выбор границы EER m.

Подтверждена гипотеза о важности способа, которым пользователь набирает первые символы по сравнению со всеми остальными. Комбинация признаков базового решения с динамическим подходом привела к наилучшей точности. Важно отметить, что топ-5 наиболее значимых признаков включают в себя только время полного нажатия клавиши.

Усложнение архитектуры признаков на подвыборках ухудшает качество модели, что связано с небольшой длиной подпоследовательности. Для использования локальных признаков необходимо построение более сложной модели или использование глубокого обучения.

Между степенью упорядоченности и главной метрикой аутентификации найдена нелинейная зависимость – для более слабых моделей классификации EER будет кратно слабее, чем показатель ROC AUC, что продемонстрировано на Рисунке 3. Итоговая точность модели сильно падает из-за пользователей, которых модель затрудняется определить. Это подтверждается сравнением медианного (0,965) и среднего (0,948) значений ROC-AUC.

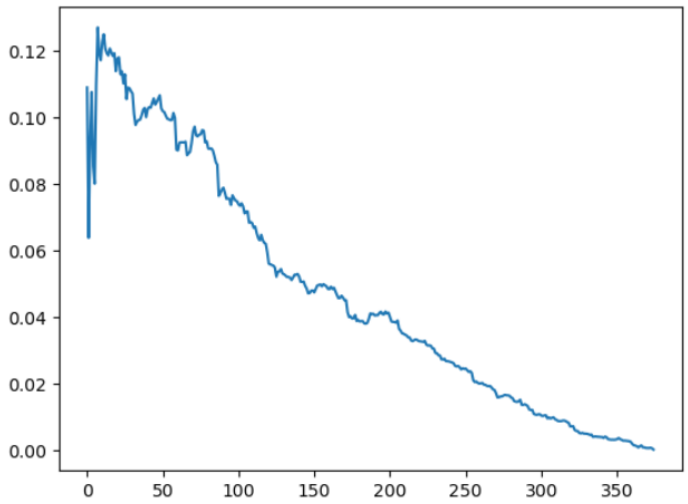

Рисунок 3 – Относительная разница упорядоченных ROC-AUC и EER
Figure 3 – Relative difference of ordered ROC-AUC and EER

Таким образом, решение на основе градиентного бустинга способно достигнуть показателей, превосходящих средние результаты среди подобных исследований. Выдвинуты утверждения относительно взаимосвязи признаков и метрик модели.

Заключение

Построение системы аутентификации на основе поведенческой биометрии, в частности динамики нажатий клавиш, представляет собой эффективный и актуальный метод проверки подлинности пользователя.

Полученные результаты не являются предельно достижимыми. Возможно уменьшение разновидности признаков для работы с высокой дисперсией биграмм, объединение статистик по часто используемым клавишам в наборы по 2–3, 4–7 клавиш в частоты употребления, аккумулятивное разделение последовательности, добавление классификации и усложнение архитектуры системы до каскада.

Этот подход имеет большое практическое значение, поскольку позволяет автоматизировать и улучшить процесс многофакторной аутентификации. Модельное выявление несанкционированного доступа может быть удобным средством дополнительной защиты для пользователя, так как он может и не догадываться о постоянном мониторинге.

Этот метод также может быть полезен для исследования различных изменений в психологическом состоянии человека. Результаты исследования могут быть использованы в смежных работах, направленных на анализ настроения и обнаружения депрессии или, например, для предсказания возраста или выявления болезней на основе динамики нажатий клавиш.

Таким образом, развитие и систематизация исследования динамики нажатий клавиш для аутентификации пользователя представляет собой перспективную и важную область, которая может принести значительный вклад в область информационной

безопасности и технологий. Реализация предложенных перспектив позволит повысить эффективность и точность модели, что окажет влияние на развитие методов поведенческой аутентификации.

СПИСОК ИСТОЧНИКОВ / REFERENCES

1. Obaidat M.S. A verification methodology for computer systems users. In: *SAC '95: 1995 Symposium on Applied Computing, 26-28 February 1995, Nashville, USA*. New York: Association for Computing Machinery; 1995. P. 258–262. https://doi.org/10.1145/315891.315976
2. Roshanbin N., Miller J. Enhancing CAPTCHA Security Using Interactivity, Dynamism, and Mouse Movement Patterns. *International Journal of Systems and Service-Oriented Engineering*. 2016;6(1):17–36. https://doi.org/10.4018/ijssoe.2016010102
3. Chang L., Shi F., Taghizadeh-Hesary F., Saydaliev H.B. Information and communication technologies development and the resource curse. *Resources Policy*. 2023;80. https://doi.org/10.1016/j.resourpol.2022.103123
4. Saini B.S., Kaur N., Bhatia K.S. Keystroke Dynamics for Mobile Phones: A Survey. *Indian Journal of Science and Technology*. 2016;9(6). https://doi.org/10.17485/ijst/2016/v9i6/82084
5. Dhakal V., Feit A.M., Kristensson P.O., Oulasvirta A. Observations on Typing from 136 Million Keystrokes. In: *CHI '18: CHI Conference on Human Factors in Computing Systems, 21-26 April 2018, Montreal, QC, Canada*. New York: Association for Computing Machinery; 2018. https://doi.org/10.1145/3173574.3174220
6. Lee H., Hwang J.Y., Kim D.I., Lee S., Lee S.-H., Shin J.S. Understanding Keystroke Dynamics for Smartphone Users Authentication and Keystroke Dynamics on Smartphones Built-In Motion Sensors. *Security and Communication Networks*. 2018;2018. https://doi.org/10.1155/2018/2567463
7. Georgiev M., Eberz S., Martinovic I. Techniques for Continuous Touch-Based Authentication. In: *ISPEC 2022: The 17th International Conference on Information Security Practice and Experience, 23-25 November 2022, Taipei, Taiwan*. Cham: Springer; 2022. P. 409–431. https://doi.org/10.1007/978-3-031-21280-2_23
8. Abernethy M., Rai S. Applying Feature Selection to Reduce Variability in Keystroke Dynamics Data for Authentication Systems. In: *13th Australian Information Warfare and Security Conference, 3-5 December 2012, Perth, Western Australia*. Perth: SRI Security Research Institute, Edith Cowan University; 2012. P. 17–23. https://doi.org/10.4225/75/57a841bcbefaf
9. Messerman A., Mustafić T., Camtepe S.A., Albayrak S. Continuous and non-intrusive identity verification in real-time environments based on free-text keystroke dynamics. In: *2011 International Joint Conference on Biometrics (IJCB), 11-13 October 2011, Washington, DC, USA*. IEEE; 2011. https://doi.org/10.1109/ijcb.2011.6117552
10. Stragapede G. et al. IEEE BigData 2023 Keystroke Verification Challenge (KVC). In: *2023 IEEE International Conference on Big Data (BigData), 15-18 December 2023, Sorrento, Italy*. IEEE; 2024. P. 6092–6100. https://doi.org/10.1109/BigData59044.2023.10386557
11. Usman A.K., Shah M.H. Strengthening E-Banking security using Keystroke Dynamics. *Journal of Internet Banking and Commerce*. 2013;18(3). URL: https://www.icommercecentral.com/open-access/strengthening-ebanking-security-using-keystroke-dynamics.php?aid=38267

12. Shah A., Shah P., Shah H., Bhadane C. Strengthening user Authentication using Keystroke Dynamics. *International Journal of Engineering Research & Technology.* 2015;4(11).
13. Stragapede G., Delgado-Santos P., Tolosana R., Vera-Rodriguez R., Guest R., Morales A. TypeFormer: Transformers for Mobile Keystroke Biometrics. URL: https://doi.org/10.48550/arXiv.2212.13075 (Accessed 10th April 2024).
14. Özbek M.E. A flexible approach for biometric menagerie on user classification of keystroke data. *Journal of Electrical Engineering*. 2023;74(1). https://doi.org/10.2478/jee-2023-0003
15. Teh P.S., Teoh A.B.J., Yue S. A Survey of Keystroke Dynamics Biometrics. *The Scientific World Journal*. 2013;2013. https://doi.org/10.1155/2013/408280
16. Гузаиров М.Б., Исмагилова А.С., Лушников Н.Д. Аутентификация пользователей информационной системы по изображению лица. *Моделирование, оптимизация и информационные технологии*. 2023;11(4). https://doi.org/10.26102/2310-6018/2023.43.4.017
Guzairov M.B., Ismagilova A.S., Lushnikov N.D. Authentication of information system users by facial image. *Modelirovanie, optimizatsiya i informatsionnye tekhnologii = Modeling, Optimization and Information Technology.* 2023;11(4). (In Russ.). https://doi.org/10.26102/2310-6018/2023.43.4.017

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Смирнов Илья Сергеевич, специалист по интеллектуальному анализу данных, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация.
e-mail: 202591@edu.fa.ru
ORCID: https://orcid.org/0009-0000-2982-1441

Ilya S. Smirnov, Data Mining Specialist, Financial University under the Government of the Russian Federation, Moscow, the Russian Federation.

Кочкаров Азрет Ахматович, доктор технических наук степень, доцент, профессор кафедры анализа данных и машинного обучения, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация.
e-mail: AAKochkarov@fa.ru
ORCID: https://orcid.org/0000-0002-3232-5331

Azret A. Kochkarov, Doctor of Technical Sciences, assistant professor, professor of the data analisys and machine learning department, Financial University under the Government of the Russian Federation, Moscow, the Russian Federation.

Статья поступила в редакцию 31.05.2024; одобрена после рецензирования 10.06.2024; принята к публикации 19.06.2024.

The article was submitted 31.05.2024; approved after reviewing 10.06.2024; accepted for publication 19.06.2024.