УДК 519.6+004.056
DOI: 10.26102/2310-6018/2026.53.2.015

Применение методов искусственного интеллекта для анализа поведенческой биометрии человека в обеспечении безопасности сложных информационных систем

О.В. Шелестова✉, А.А. Кочкаров

Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация

Резюме. В статье рассматривается применение методов и технологий искусственного интеллекта для анализа поведенческой биометрии человека в задачах обеспечения безопасности сложных информационных систем. Актуальность исследования обусловлена ограничениями традиционных механизмов аутентификации, ориентированных преимущественно на начальный этап пользовательской сессии и недостаточно эффективных при выявлении подмены пользователя в процессе взаимодействия с системой. В качестве альтернативного подхода предлагается использование поведенческих характеристик пользователя для непрерывной оценки доверия к текущей сессии. В работе проведен анализ обезличенных данных ввода текста на мобильном устройстве, отражающих временные и структурные особенности взаимодействия пользователя с интерфейсом. Показано, что совокупность таких характеристик позволяет выявлять устойчивые поведенческие закономерности, пригодные для профилирования пользователей. С применением методов снижения размерности и кластерного анализа выделены типовые поведенческие профили, отличающиеся по стилю и ритму ввода, а также характеру исправлений. Установлено, что принадлежность к кластеру сохраняется на протяжении нескольких сессий при допустимой вариативности отдельных признаков. Предложен риск-ориентированный подход к оценке отклонений поведения, основанный на сопоставлении текущих поведенческих признаков с типовым кластерным профилем. Результаты исследования подтверждают целесообразность использования кластерных поведенческих профилей в системах риск-ориентированного управления доступом и могут быть использованы при проектировании и развитии механизмов непрерывной аутентификации в сложных информационных системах.

Ключевые слова: поведенческая биометрия, информационная безопасность, искусственный интеллект, машинное обучение, кластерный анализ, непрерывная аутентификация, анализ пользовательского поведения.

Для цитирования: Шелестова О.В., Кочкаров А.А. Применение методов искусственного интеллекта для анализа поведенческой биометрии человека в обеспечении безопасности сложных информационных систем. *Моделирование, оптимизация и информационные технологии.* 2026;14(2). URL: https://moitvivt.ru/ru/journal/article?id=2201 DOI: 10.26102/2310-6018/2026.53.2.015

Application of artificial intelligence methods to analyze human behavioral biometrics in ensuring the security of complex information systems

O.V. Shelestova✉, A.A. Kochkarov

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

Abstract. This article examines the application of artificial intelligence methods and technologies to analyzing human behavioral biometrics in the security of complex information systems. The relevance of the study stems from the limitations of traditional authentication mechanisms, which focus primarily on the initial stage of a user session and are ineffective in detecting user impersonation during interaction with the system. An alternative approach is proposed, using user behavioral characteristics to continuously assess trust in the current session. The paper analyzes anonymized text input data on a mobile device, reflecting the temporal and structural features of user interaction with the interface. It is shown that the combination of such characteristics allows for the identification of stable behavioral patterns suitable for user profiling. Using dimensionality reduction and cluster analysis methods, typical behavioral profiles are identified, differing in input style and rhythm, as well as the nature of corrections. Cluster membership is established to be maintained across multiple sessions with acceptable variability in individual characteristics. A risk-based approach to assessing behavioral deviations is proposed, based on comparing current behavioral indicators with a typical cluster profile. The study's results confirm the feasibility of using cluster behavioral profiles in risk-based access control systems and can be used in the design and development of continuous authentication mechanisms in complex information systems.

Keywords: behavioral biometrics, information security, artificial intelligence, machine learning, cluster analysis, continuous authentication, user behavior analysis.

For citation: Shelestova O.V., Kochkarov A.A. Application of artificial intelligence methods to analyze human behavioral biometrics in ensuring the security of complex information systems. *Modeling, Optimization and Information Technology*. 2026;14(2). (In Russ.). URL: https://moitvivt.ru/ru/journal/article?id=2201 DOI: 10.26102/2310-6018/2026.53.2.015

Введение

В условиях активной цифровизации различных сфер деятельности человека и роста удаленных форм взаимодействия пользователей с информационными системами возрастает значимость надежных механизмов аутентификации и контроля доступа. Современные цифровые сервисы все чаще функционируют в распределенной среде, где пользователь получает доступ к критически важным данным и функциям вне защищенного периметра организации. При этом традиционные методы аутентификации, основанные на знании или владении (пароли, одноразовые коды), в большинстве случаев обеспечивают проверку личности только на этапе входа в систему и не позволяют своевременно выявлять ситуацию подмены пользователя в рамках уже активной сессии.

Одним из перспективных направлений повышения уровня безопасности является применение поведенческой биометрии – подхода, основанного на анализе индивидуальных характеристик взаимодействия человека с цифровым интерфейсом [1]. В отличие от физиологических биометрических признаков, поведенческие характеристики формируются в процессе работы пользователя с системой и отражают особенности моторики, когнитивных стратегий и привычек взаимодействия. К таким характеристикам относятся темп и ритм ввода текста, структура ошибок и исправлений, длительности пауз, использование вспомогательных функций интерфейса и другие динамические параметры поведения [2].

Особый интерес представляет использование поведенческой биометрии в рамках концепции непрерывной аутентификации, при которой контроль подлинности пользователя осуществляется на протяжении всей сессии, а не только в момент входа [3]. Такой подход позволяет повысить устойчивость информационных систем к внутренним угрозам, компрометации учетных данных и несанкционированному использованию рабочих сессий. Вместе с тем практическая реализация непрерывной аутентификации сталкивается с рядом методологических и технических сложностей, связанных с

высокой вариативностью поведенческих данных, зависимостью поведения от контекста и необходимостью интерпретации получаемых результатов.

Современные методы и технологии искусственного интеллекта позволяют эффективно работать с многомерными и нестационарными поведенческими данными, выявляя устойчивые закономерности и типовые стратегии взаимодействия пользователей с системой [4, 5]. Однако значительная часть существующих исследований ориентирована либо на задачу индивидуальной идентификации, либо на анализ поведения в лабораторных условиях, тогда как в прикладных системах информационной безопасности требуется формирование интерпретируемых поведенческих профилей, пригодных для использования в риск-ориентированных механизмах контроля доступа. В этой связи актуальной является задача исследования возможностей применения методов искусственного интеллекта для анализа поведенческой биометрии и выделения типовых поведенческих профилей пользователей, отражающих устойчивые различия в стилях взаимодействия с интерфейсом. Использование таких профилей позволяет перейти к более гибкой модели оценки доверия, учитывающей степень отклонения текущего поведения от ожидаемого.

Целью настоящей статьи является исследование методов анализа поведенческой биометрии человека на основе данных мобильного ввода текста с применением технологий искусственного интеллекта, а также оценка возможности использования кластерных поведенческих профилей в задачах повышения безопасности сложных информационных систем.

Материалы и методы

Эмпирической основой исследования послужил обезличенный открытый набор данных мобильного ввода текста, включающий сведения о более чем 37 тыс. участников и более чем 560 тыс. сессий ввода. Сбор данных проводился в ходе выполнения задания по вводу текста, при котором участникам предлагалось воспроизвести отображаемый на экране текст. Используемый браузерный тест был разработан в рамках университетского исследования и размещен на серверной инфраструктуре образовательного учреждения. Участие в эксперименте носило добровольный характер: пользователи проходили тестирование через публичный веб-ресурс typingtest.com компании TypingMaster Inc., предоставляющей сервисы оценки и тренировки навыков печати. Полученная в ходе эксперимента информация была обезличена и не содержит персональных идентификаторов, что исключает возможность установления личности. Данные представлены на двух уровнях: агрегированные сведения об участниках и массив сессионных поведенческих метрик, отражающих особенности взаимодействия пользователя с мобильной клавиатурой [6].

В качестве исходных переменных использовались числовые поведенческие признаки, характеризующие процесс ввода текста. Для обеспечения интерпретируемости и последующего анализа метрики рассматривались с учетом их функционального назначения. Отдельно анализировались показатели, преимущественно влияющие на темп и ритм ввода (например, скорость печати, межклавишные интервалы и параметры пауз), показатели точности и корректирующих действий (ошибка на символ, частота Backspace и связанные метрики), а также признаки использования функций мобильной клавиатуры, включая механизмы автокоррекции и непрерывного ввода (Swype). Данный подход позволил связать численные характеристики с содержательными аспектами поведения и корректнее интерпретировать дальнейшие результаты статистического и кластерного анализа.

Подготовка данных ограничивалась минимально необходимыми операциями, обеспечивающими корректность применения методов анализа, без использования дополнительных преобразований, влияющих на смысловую интерпретацию признаков. Выполнялась проверка корректности значений и типов данных, исключение неполных наблюдений, а также приведение признаков к сопоставимому масштабу. Поскольку используемые методы (снижение размерности и кластеризация) чувствительны к масштабу переменных, применялась стандартизация признаков, при которой значения преобразуются относительно среднего и стандартного отклонения по выборке.

Для выявления взаимосвязей между поведенческими метриками и оценки избыточности признаков был проведен корреляционный анализ. Для каждой пары признаков вычислялся коэффициент линейной корреляции Пирсона, после чего анализировалась матрица корреляций. В рамках интерпретации выделялись пары показателей с высокой по модулю корреляцией как потенциально избыточные, а также комплементарные зависимости умеренной силы, отражающие связь различных компонент поведения (темп, точность, использование функций интерфейса). Результаты корреляционного анализа использовались для содержательного понимания структуры признакового пространства и для обоснования набора метрик, применяемых при формировании поведенческих профилей.

Сегментация пользователей по совокупности поведенческих признаков выполнялась методом неконтролируемого обучения – кластеризацией KMeans [7]. Перед кластеризацией применялось снижение размерности методом главных компонент (PCA), используемое в качестве инструмента анализа структуры данных и качественной проверки различимости групп в пониженном признаковом пространстве. Выбор числа кластеров осуществлялся на основе сочетания количественного и качественного критериев. В контексте количественного критерия использовался метод локтя, основанный на анализе внутрикластерной суммы квадратов отклонений при переборе нескольких значений числа кластеров. Для качественного критерия использовалась визуальная оценка интерпретируемости сегментации в проекции главных компонент, поскольку для поведенческих данных характерно естественное перекрытие групп и важна не геометрическая разделимость, а наличие устойчивых областей концентрации, соответствующих различным стратегиям поведения. Итоговое значение числа кластеров выбиралось как компромисс между снижением внутрикластерной дисперсии и интерпретируемостью профилей.

Кластеризация выполнялась на стандартизированных признаках с использованием алгоритма KMeans с инициализацией центров методом k-means++ и фиксацией начального состояния генератора случайных чисел для обеспечения воспроизводимости результатов. Для повышения устойчивости разбиения алгоритм запускался с несколькими инициализациями, после чего выбиралось решение с минимальной внутрикластерной дисперсией. По результатам кластеризации каждому наблюдению присваивалась метка кластера. Для интерпретации полученных групп рассчитывались средние значения признаков внутри каждого кластера и формировались профили кластеров, отражающие различия типовых стратегий ввода по ключевым аспектам поведения, включая темп, точность и характер исправлений, а также использование функций интерфейса.

Для дополнительного анализа различимости кластеров и структуры совместных распределений признаков использовалась визуализация парных зависимостей: диаграммы рассеяния по парам выбранных ключевых метрик с цветовым разделением по меткам кластеров и одномерные распределения по диагонали. Такой анализ позволяет оценивать зоны наибольшей плотности для каждой группы и характер перекрытия

кластеров в разных проекциях признакового пространства, что особенно важно для поведенческих данных, где различение стратегий обычно проявляется не в отдельных метриках, а в их сочетаниях.

Результаты

Результаты исследования отражают выявленные закономерности в поведенческих характеристиках пользователей при вводе текста на мобильных устройствах и демонстрируют возможности построения типовых поведенческих профилей на основе методов неконтролируемого машинного обучения. Для обеспечения интерпретируемости дальнейших выводов исходное множество поведенческих метрик было предварительно рассмотрено не как разрозненные показатели, а как совокупность параметров, отражающих различные компоненты взаимодействия с мобильной клавиатурой [8]. В рамках анализа выделялись признаки, преимущественно описывающие темп и ритм ввода (например, скорость печати, межклавишные интервалы, параметры пауз), признаки точности и коррекции набора (ошибка на символ, частота Backspace, дополнительные нажатия), а также показатели, характеризующие степень использования функций мобильного интерфейса (в частности, автокоррекция и непрерывный ввод Swype). Такое представление признакового пространства позволило связать численные значения метрик с содержательными аспектами поведения пользователя и корректнее интерпретировать взаимосвязи между признаками на последующих этапах.

Следующим шагом был выполнен корреляционный анализ, целью которого являлось выявление статистически устойчивых связей между метриками, а также обнаружение признаков с высокой степенью избыточности. Результаты корреляционного анализа (Рисунок 1) демонстрируют наличие сильных положительных корреляций между рядом показателей, отражающих схожие стороны поведения. Так, высокая взаимосвязь наблюдается между длиной пользовательского ввода и количеством нажатий клавиш, что ожидаемо, поскольку увеличение объема вводимого текста сопровождается ростом числа действий ввода. Аналогично, выраженная связь между частотой использования Backspace и ошибкой на символ отражает согласованность метрик, описывающих ошибки и исправления: при увеличении ошибочности, как правило, возрастает интенсивность коррекционных действий. Подобные зависимости важны с прикладной точки зрения, поскольку позволяют обосновать сокращение признакового пространства без потери содержательной информации и снизить влияние мультиколлинеарности на результаты моделирования.

Наряду с избыточными взаимосвязями выявлены комплементарные зависимости, отражающие взаимодействие различных компонентов поведения. Отрицательная корреляция между потенциальной длиной ввода и ошибкой на символ интерпретируется как тенденция более уверенного и точного набора у пользователей, способных поддерживать более длинные последовательности ввода. Такие связи не являются тривиальными, поскольку связывают характеристики темпа и структуры ввода с качеством и стратегией самоконтроля. Дополнительный интерес представляют зависимости между ошибочностью и количеством клавиш на символ, которые отражают «стоимость» ошибок в поведении: при росте доли ошибок обычно увеличивается число дополнительных действий, необходимых для достижения корректного результата. В совокупности корреляционная картина подтверждает, что поведение при вводе текста формируется не одним параметром, а комбинацией моторных и когнитивных факторов, а также особенностей использования вспомогательных механизмов интерфейса.

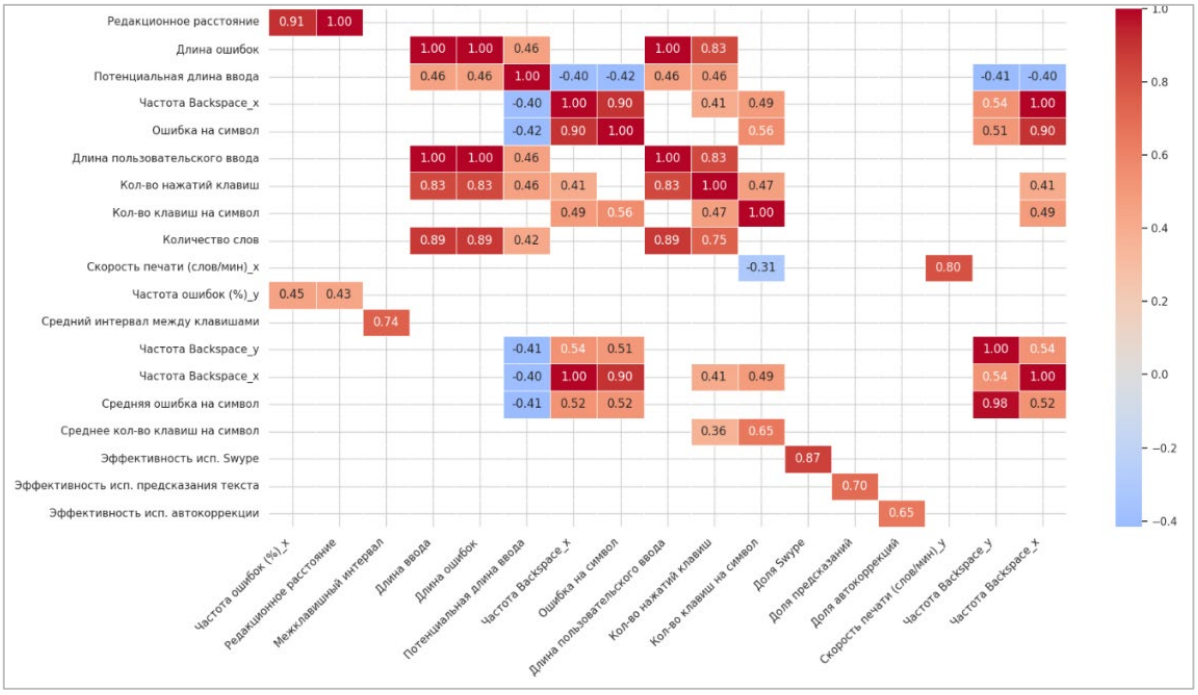

Рисунок 1 – Корреляционный анализ признаков с фильтрацией по высокой взаимосвязи
Figure 1 – Correlation analysis of features with filtering by high correlation

Для перехода от анализа отдельных метрик к формированию типовых профилей был выполнен кластерный анализ пользователей с применением алгоритма KMeans. Существенным этапом являлось обоснование числа кластеров, поскольку чрезмерная детализация приводит к снижению интерпретируемости и нестабильным профилям, тогда как слишком грубая сегментация может скрывать реальные различия в стратегиях поведения. В качестве количественного критерия использовался метод локтя на основе внутрикластерной дисперсии (Рисунок 2). Кривая демонстрирует характерное замедление снижения дисперсии при увеличении числа кластеров, что указывает на наличие ограниченного числа устойчивых поведенческих режимов в выборке и снижение эффективности дальнейшего дробления.

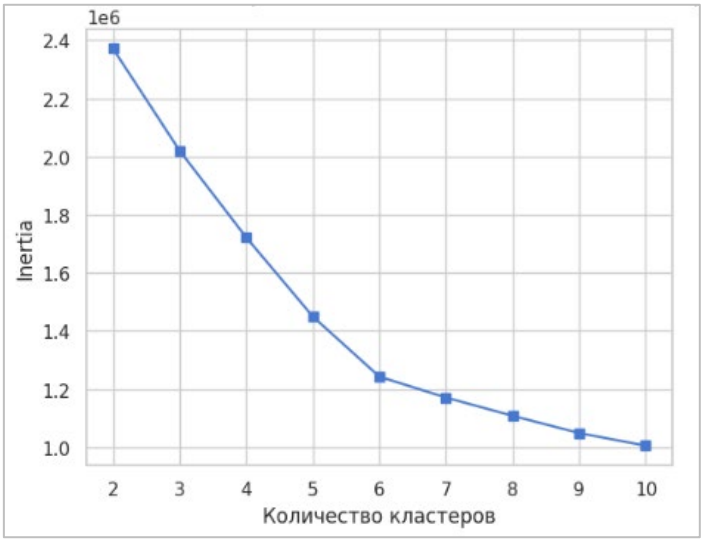

Рисунок 2 – Метод локтя для определения оптимального количества кластеров
Figure 2 – Elbow method for determining the optimal number of clusters

Дополнительно была использована визуальная проверка структуры данных в пониженном пространстве признаков с применением метода главных компонент.

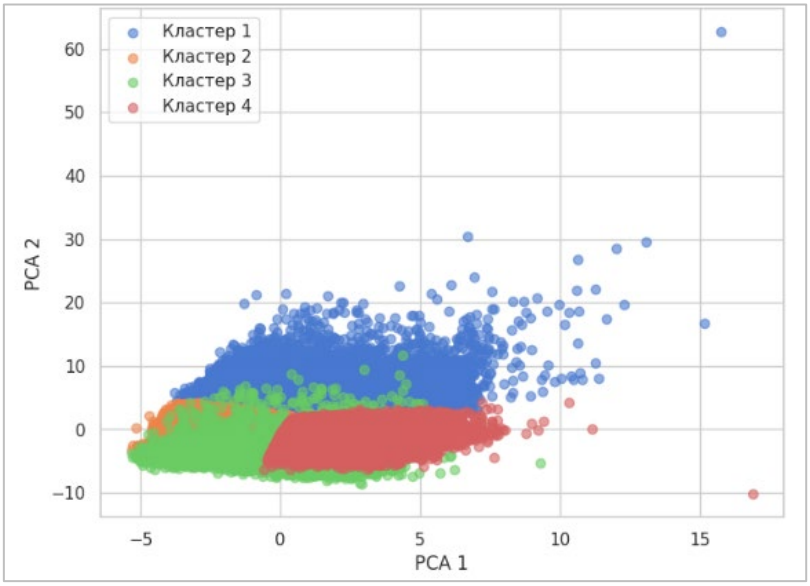

Рисунок 3 – Снижение размерности методом PCA (4 кластера)
Figure 3 – Dimensionality reduction by PCA (4 clusters)

Этот этап применялся не для формального анализа геометрической разделимости кластеров, а для качественной оценки выбранного значения k с точки зрения интерпретируемости поведенческих профилей. Визуальный анализ показал, что при увеличении числа кластеров (Рисунок 3) наблюдается существенное наложение групп, что затрудняет содержательную интерпретацию и снижает практическую ценность сегментации. Выбор $k = 3$ оказался компромиссным решением, обеспечивающим достаточную структурированность поведенческих профилей при сохранении реалистичного перекрытия, типичного для поведенческих данных. Результаты кластеризации при $k = 3$ представлены в пространстве главных компонент на Рисунке 4.

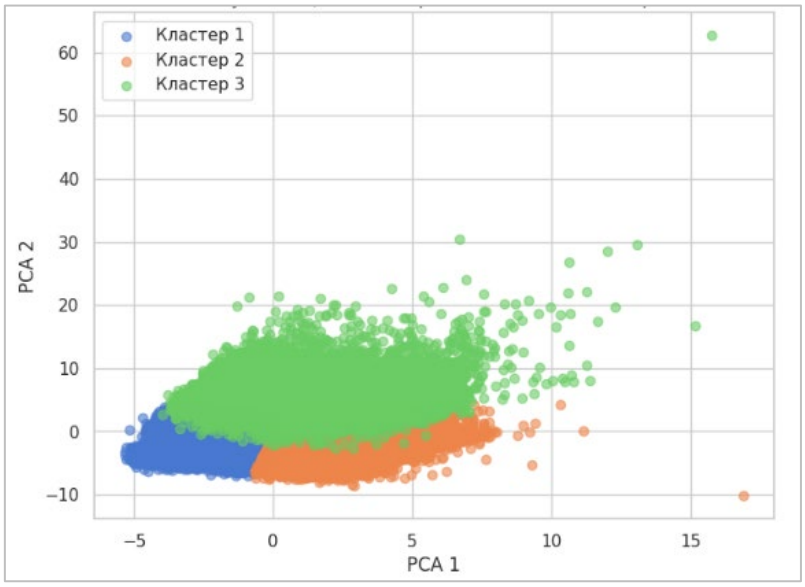

Рисунок 4 – Снижение размерности методом PCA (3 кластера)
Figure 4 – Dimensionality reduction by PCA (3 clusters)

Частичное пересечение кластеров в данном случае является ожидаемым и методологически корректным: поведенческие признаки могут изменяться в зависимости от контекста ввода, уровня концентрации, привычек использования клавиатуры и иных факторов, не приводя при этом к полному изменению стратегии поведения. Поэтому при интерпретации кластеров следует учитывать не столько границы, сколько области наибольшей плотности точек, отражающие доминирующие режимы взаимодействия пользователей с интерфейсом.

Для содержательной интерпретации кластеров была построена тепловая карта средних значений признаков по каждой группе пользователей (Рисунок 5).

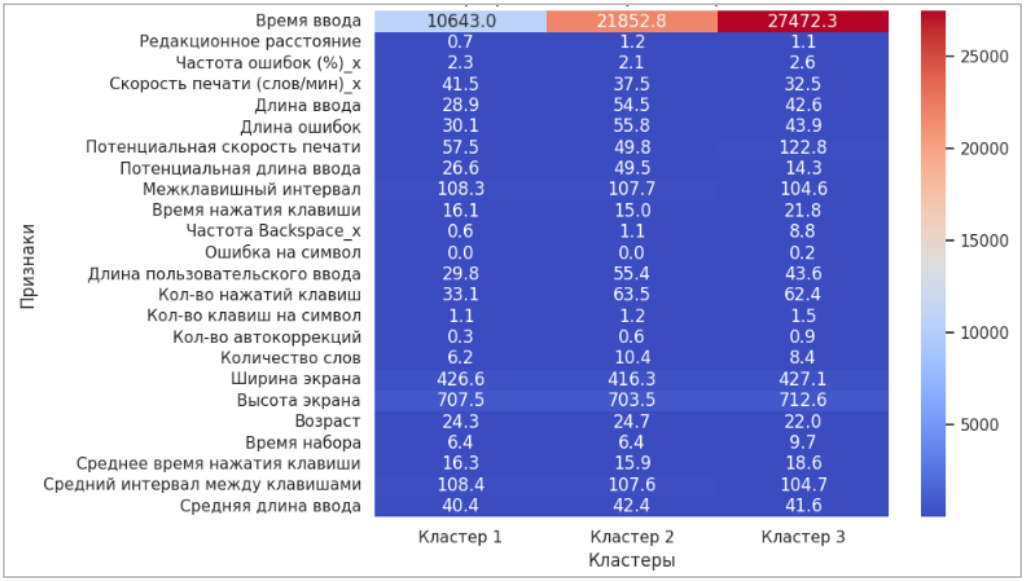

Рисунок 5 – Профили кластеров по признакам
Figure 5 – Cluster profiles by feature

Первый кластер характеризуется сравнительно меньшей длиной пользовательского ввода при более выраженных межклавишных интервалах и низкой ошибочности. Такой профиль согласуется с аккуратным, фрагментированным стилем ввода, когда пользователь вводит текст короткими последовательностями, делая паузы между действиями, и при этом стремится минимизировать число исправлений. Второй кластер демонстрирует более сбалансированную стратегию: умеренные значения скорости и ошибок сочетаются со средней длиной пользовательского ввода, что указывает на способность поддерживать длительное текстовое взаимодействие при сохранении стабильного качества набора. Третий кластер отличается более высоким темпом нажатий при повышенной частоте ошибок и активном использовании исправлений. Такой профиль соответствует импульсивному характеру ввода, при котором быстрые действия сопровождаются корректирующими операциями, а поведение приобретает повышенную вариативность и нестабильность по ряду метрик.

Для уточнения различий между кластерами и проверки согласованности интерпретации был выполнен анализ парных распределений ключевых признаков с цветовым разделением по кластерам, представленный на Рисунке 6.

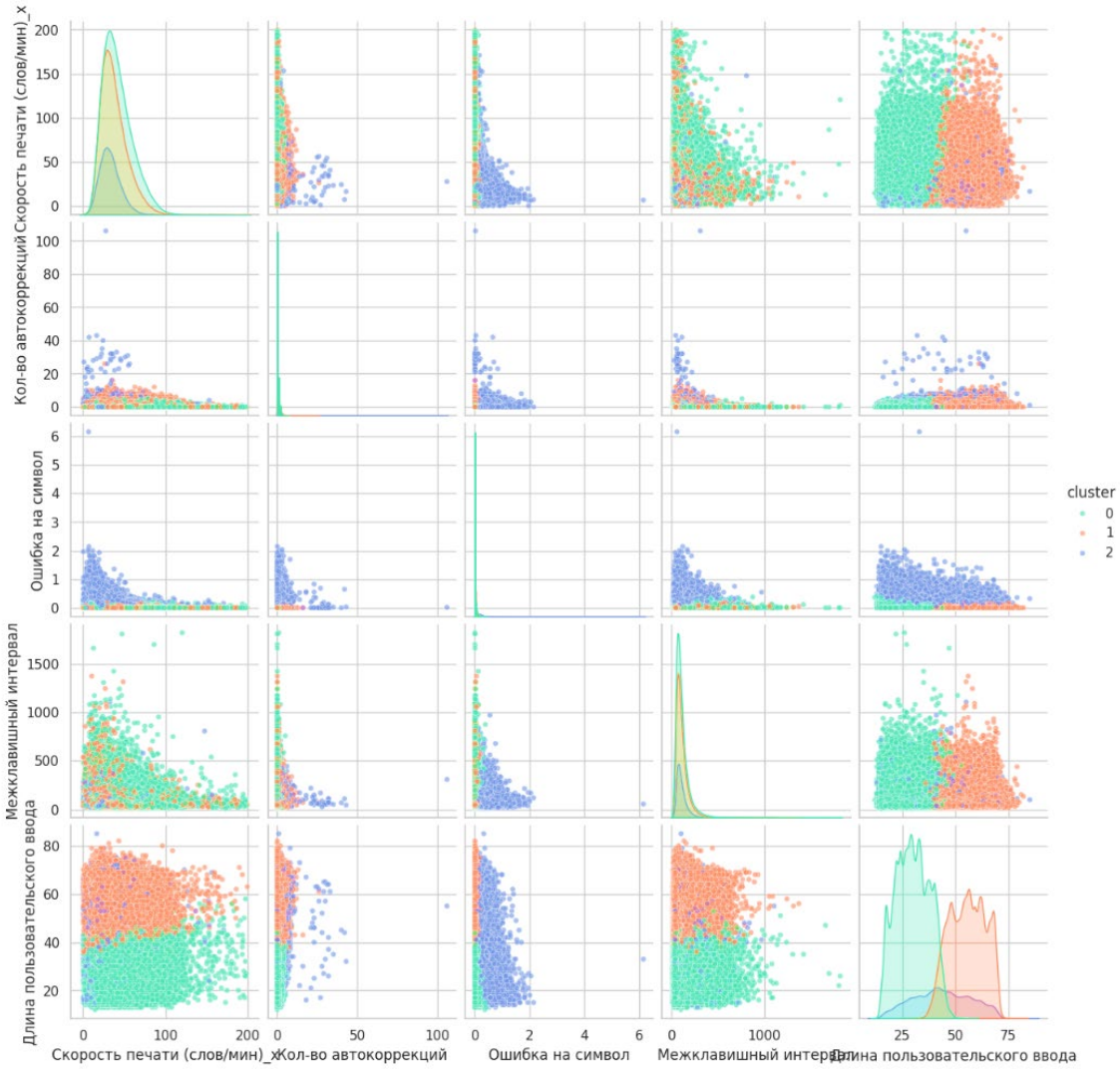

Рисунок 6 – Рассеивание и плотность признаков
Figure 6 – Dispersion and density of features

Визуальное представление данных на Рисунке 6 позволяет оценить совместную структуру признаков и выявить области концентрации значений, характерные для каждой группы, даже при наличии перекрытия. По ряду проекций наблюдается различие «центров плотности» кластеров: второй кластер чаще локализуется в области более длинного непрерывного ввода при умеренной ошибочности, первый – в области коротких фрагментов текста с относительно низкой ошибкой, третий – демонстрирует более широкий разброс значений и повышенную ошибочность при сопоставимой длине ввода, что согласуется с реактивной и вариативной стратегией поведения. Анализ парных распределений показывает, что различия между поведенческими профилями проявляются преимущественно в сочетаниях признаков, а не в отдельных метриках.

Обсуждение

Полученные результаты кластерного анализа подтверждают возможность выявления типовых стратегий поведения пользователей при вводе текста на мобильных устройствах. Несмотря на выраженную вариативность индивидуальных характеристик, в исследовании были обнаружены устойчивые различия в стилях взаимодействия с интерфейсом, отражающие особенности темпа ввода, точности и характера коррекций.

Это свидетельствует о наличии обобщенных поведенческих паттернов, которые могут использоваться для профилирования пользователей в задачах информационной безопасности.

При интерпретации результатов принципиально важно учитывать, что поведенческая биометрия отражает не статические свойства личности, а динамическое состояние пользователя. В отличие от физиологических биометрических признаков, поведенческие характеристики формируются под воздействием как устойчивых индивидуальных факторов, так и временных психофизиологических состояний. К таким состояниям относятся усталость, эмоциональное напряжение, уровень когнитивной нагрузки и стресс, которые способны заметно изменять параметры взаимодействия с интерфейсом [9]. Анализ распределений признаков показывает, что переход пользователя между кластерами в отдельных сессиях может быть обусловлен не изменением его базовой стратегии поведения, а временным смещением характеристик ввода. Например, при повышенной утомляемости наблюдается увеличение межклавишных интервалов, рост количества пауз и снижение длины непрерывного пользовательского ввода. В условиях эмоционального напряжения, напротив, возможно ускорение темпа нажатий, рост числа ошибок и увеличение частоты исправлений [10]. Такие изменения могут приводить к временному приближению поведения пользователя к характеристикам другого кластера. Подобная динамика не должна рассматриваться как недостаток модели, а напротив, отражает естественную природу поведенческих данных. Поведение человека не является строго воспроизводимым и может варьироваться в зависимости от времени суток, продолжительности работы, сложности выполняемых задач и внешних факторов. Поэтому кластерная принадлежность в контексте поведенческой биометрии должна интерпретироваться как область наиболее вероятного поведения пользователя в нормальном состоянии.

Выделенные в ходе исследования кластеры могут рассматриваться как типовые поведенческие режимы. Первый кластер соответствует аккуратному и замедленному стилю взаимодействия, который может усиливаться при утомлении или снижении концентрации внимания. Второй кластер отражает наиболее устойчивое и сбалансированное состояние, характерное для нормального рабочего режима пользователя. Третий кластер ассоциируется с реактивным стилем ввода, который может усиливаться при эмоциональном напряжении, спешке или повышенной когнитивной нагрузке. Таким образом, кластеры отражают не только различия между пользователями, но и возможные функциональные состояния одного и того же пользователя в разные моменты времени. В системах непрерывной аутентификации недопустимо трактовать любое отклонение поведения как признак компрометации учетной записи. Повышение частоты ошибок или изменение темпа ввода может быть следствием усталости или стресса, а не вмешательства третьего лица. В этой связи результаты исследования подтверждают целесообразность использования риск-ориентированного подхода, при котором поведенческие отклонения оцениваются с учетом допустимого диапазона вариативности, а не по принципу строгого соответствия эталону.

Использование кластерных профилей позволяет учитывать такие изменения более корректно. Вместо сравнения текущей сессии с фиксированным шаблоном поведение сопоставляется с типовой стратегией, допускающей вариации внутри кластера. Это снижает чувствительность системы к кратковременным изменениям психоэмоционального состояния пользователя и способствует уменьшению числа ложных отказов, что особенно важно в условиях длительной работы с системой. Дополнительным преимуществом является возможность интерпретации динамики поведения как сигнала контекста, а не только безопасности [11]. Постепенное смещение

характеристик в сторону кластеров с повышенной вариативностью и ошибочностью может указывать на нарастание утомления пользователя. Такие изменения могут использоваться не для блокировки доступа, а для адаптивного усиления контроля только при выполнении критически значимых операций. В то же время исследование имеет ряд ограничений: используемые данные не содержат явных меток состояния пользователя, таких как уровень усталости или эмоционального напряжения, что не позволяет напрямую связать кластерную принадлежность с конкретными психофизиологическими причинами. Это определяет направления дальнейших исследований, связанные с интеграцией контекстных признаков, временных окон анализа и адаптивных механизмов обновления поведенческого профиля.

В прикладных системах поведенческая биометрия должна интерпретироваться с учетом сценариев доступа и уровня потенциального ущерба. Для операций с высокой критичностью (изменение персональных данных, финансовые транзакции или доступ к конфиденциальной информации) допустимо применение более строгих порогов и дополнительных проверок, поскольку последствия компрометации существенно превышают возможные неудобства пользователя. В то же время для низкорисковых действий целесообразно ослабление требований, что позволяет снизить число ложных срабатываний и избежать формирования ощущения нестабильности работы системы [12].

Поведенческая биометрия должна рассматриваться как динамическая система оценки доверия, чувствительная к состоянию пользователя. Выделенные кластерные профили отражают типовые режимы поведения и могут служить основой для построения интеллектуальных механизмов непрерывной аутентификации, обеспечивающих баланс между безопасностью, устойчивостью и удобством эксплуатации сложных информационных систем.

Подход к внедрению поведенческой биометрии в архитектуру информационной системы основывается на представлении поведения пользователя как динамического процесса, характеристики которого могут использоваться для оценки уровня доверия к текущей сессии. На Рисунке 7 представлена схема, отражающая последовательность этапов обработки поведенческих данных и принятия решений в системе управления доступом.

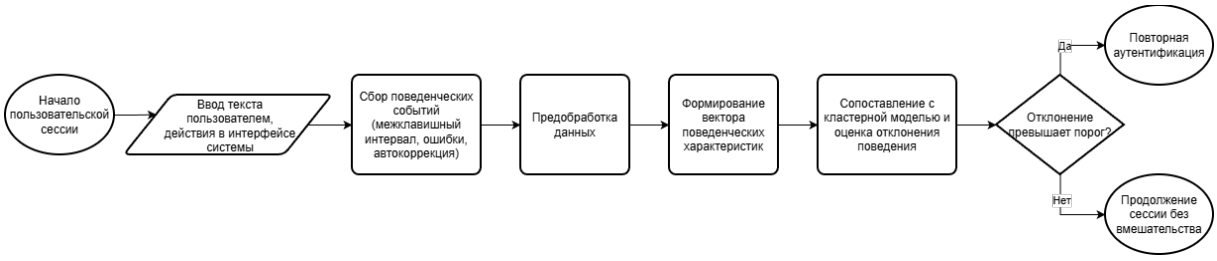

Рисунок 7 – Схема методики анализа поведенческой биометрии
Figure 7 – Scheme of behavioral biometry analysis methodology

Модуль поведенческой биометрии может быть интегрирован в информационную систему и функционировать в фоновом режиме в процессе пользовательской сессии. В ходе взаимодействия пользователя с интерфейсом регистрируются поведенческие события, на основе которых после предварительной обработки формируется вектор поведенческих характеристик. Полученные признаки сопоставляются с типовыми поведенческими профилями, сформированными в результате кластерного анализа, что позволяет оценить степень отклонения текущего поведения от ожидаемого паттерна. На основании рассчитанной оценки риска система принимает решение о дальнейшем ходе

сессии: при превышении установленного порогового значения может быть инициирована повторная аутентификация, тогда как при отсутствии значимых отклонений взаимодействие продолжается без вмешательства со стороны системы.

Заключение

В настоящей статье было проведено исследование возможностей применения методов и технологий искусственного интеллекта для анализа поведенческой биометрии человека на основе данных мобильного ввода текста. Актуальность работы обусловлена ограничениями традиционных механизмов аутентификации, ориентированных преимущественно на начальный этап пользовательской сессии и недостаточно эффективных при выявлении подмены пользователя в процессе взаимодействия с информационной системой.

В рамках исследования был выполнен анализ обезличенного набора поведенческих данных, включающего характеристики ввода текста на мобильных устройствах. Проведены этапы предварительной обработки данных, корреляционного анализа признаков и снижения размерности, что позволило сформировать интерпретируемое признаковое пространство, отражающее ключевые аспекты пользовательского поведения. С применением методов неконтролируемого машинного обучения был выполнен кластерный анализ пользователей, в результате которого выделены типовые поведенческие профили, различающиеся по темпу ввода, точности и характеру коррекций. Полученные результаты показали, что даже при высокой вариативности индивидуального поведения возможно выявление устойчивых стратегий взаимодействия с интерфейсом, сохраняющихся при допустимых изменениях психофизиологического состояния пользователя. Особое внимание в работе уделено интерпретации кластерных профилей с учетом факторов усталости и эмоционального напряжения. Показано, что временные изменения поведения не должны рассматриваться исключительно как аномалии, а требуют контекстной оценки. Это обосновывает целесообразность использования риск-ориентированного подхода, при котором поведенческие характеристики применяются не для строгой идентификации, а для динамической оценки доверия к текущей пользовательской сессии.

Практическая значимость полученных результатов заключается в возможности использования кластерных поведенческих профилей при проектировании подсистем непрерывной аутентификации и контроля доступа в сложных информационных системах. Предложенный подход позволяет снизить чувствительность к кратковременным колебаниям поведения, уменьшить вероятность ложных отказов и обеспечить баланс между уровнем безопасности и удобством пользователя.

Перспективными направлениями дальнейших исследований являются расширение набора анализируемых поведенческих признаков, учет временной структуры последовательностей ввода, интеграция контекстных факторов и разработка адаптивных механизмов обновления поведенческих профилей. Это позволит повысить устойчивость и практическую применимость интеллектуальных систем аутентификации в условиях реальной эксплуатации.

СПИСОК ИСТОЧНИКОВ / REFERENCES

1. Кочегурова Е.А., Затеев Р.П. Скрытый мониторинг пользователя в дистанционной образовательной системе на основе клавиатурной динамики. *Программирование*. 2022;(6):31–45. https://doi.org/10.31857/S0132347422060048

Kochegurova E.A., Zateev R.P. Hidden Monitoring Based on Keystroke Dynamics in Online Examination System. *Programming and Computer Software.* 2022;48(6):385–398. https://doi.org/10.1134/s0361768822060044

2. Yaacob M.N., Idrus S.Z.S., Ali W.N.A.W., et al. A Review on Feature Extraction in Keystroke Dynamics. *Journal of Physics: Conference Series.* 2020;1529(2). https://doi.org/10.1088/1742-6596/1529/2/022088
3. Смирнов И.С., Кочкаров А.А. Исследование поведенческой биометрии методами анализа данных и машинного обучения. *Моделирование, оптимизация и информационные технологии.* 2024;12(2). https://doi.org/10.26102/2310-6018/2024.45.2.021

   Smirnov I.S., Kochkarov A.A. The study of behavioral biometrics using data analysis and machine learning methods. *Modeling, Optimization and Information Technology.* 2024;12(2). (In Russ.). https://doi.org/10.26102/2310-6018/2024.45.2.021
4. Sağbaş E.A., Ballı S. Machine learning-based novel continuous authentication system using soft keyboard typing behavior and motion sensor data. *Neural Computing & Applications.* 2024;36(10):5433–5445. https://doi.org/10.1007/s00521-023-09360-9
5. Stragapede G., Delgado-Santos P., Tolosana R., et al. TypeFormer: transformers for mobile keystroke biometrics. *Neural Computing & Applications.* 2024;36:18531–18545. https://doi.org/10.1007/s00521-024-10140-2
6. Palin K., Feit A.M., Kim S., Kristensson P.O., Oulasvirta A. How do people type on mobile devices?: Observations from a study with 37,000 volunteers. In: *MobileHCI '19: Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services, 01–04 October 2019, Taipei, Taiwan.* New York: Association for Computing Machinery; 2019. https://doi.org/10.1145/3338286.3340120
7. Gautam N., Kumar N. Customer segmentation using k-means clustering for developing sustainable marketing strategies. *Business Informatics*. 2022;16(1):72–82. https://doi.org/10.17323/2587-814X.2022.1.72.82
8. Dias T., Vitorino J., Maia E., Sousa O., Praça I. KeyRecs: A keystroke dynamics and typing pattern recognition dataset. *Data in Brief*. 2023;50. https://doi.org/10.1016/j.dib.2023.109509
9. Wetherell M.A., Lau Sh.-H., Maxion R.A. The effect of socially evaluated multitasking stress on typing rhythms. *Psychophysiology*. 2023;60(8). https://doi.org/10.1111/psyp.14293
10. Tahir M., Halim Z., Waqas M., Sukhia K.N., Tu Sh. Emotion detection using convolutional neural network and long short-term memory: a deep multimodal framework. *Multimedia Tools and Applications*. 2023;83:53497–53530. https://doi.org/10.1007/s11042-023-17653-3
11. Lis K., Niewiadomska-Szynkiewicz E., Dziewulska K. Siamese Neural Network for Keystroke Dynamics-Based Authentication on Partial Passwords. *Sensors*. 2023;23(15). https://doi.org/10.3390/s23156685
12. Wahab A.A., Hou D., Schuckers S., Barbir A. Utilizing Keystroke Dynamics as Additional Security Measure to Protect Account Recovery Mechanism. In: *Proceedings of the 7th International Conference on Information Systems Security and Privacy: Volume 1, 11–13 February 2021, Virtual Event.* SciTePress; 2021. P. 33–42. https://doi.org/10.5220/0010191200330042

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Шелестова Ольга Владимировна, специалист по интеллектуальному анализу данных, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация.
e-mail: 235271@edu.fa.ru
ORCID: 0009-0005-5831-5229

Olga V. Shelestova, Data Mining Specialist, Financial University under the Government of the Russian Federation, Moscow, the Russian Federation.

Кочкаров Азрет Ахматович, доктор технических наук, доцент, профессор кафедры искусственного интеллекта, Финансовый университет при Правительстве Российской Федерации, Москва, Российская Федерация.
e-mail: AAKochkarov@fa.ru
ORCID: 0000-0002-3232-5331

Azret A. Kochkarov, Doctor of Engineering Sciences, Docent, Professor at the Artificial Intelligence Department, Financial University under the Government of the Russian Federation, Moscow, the Russian Federation.

Статья поступила в редакцию 28.01.2026; одобрена после рецензирования 24.02.2026; принята к публикации 26.02.2026.

The article was submitted 28.01.2026; approved after reviewing 24.02.2026; accepted for publication 26.02.2026.