pISSN 2073-2872
eISSN 2311-875X

Угрозы и безопасность

АУТЕНТИФИКАЦИЯ ПОЛЬЗОВАТЕЛЕЙ ПО КЛАВИАТУРНОМУ ПОЧЕРКУ ПРИ ИСПОЛЬЗОВАНИИ СИСТЕМ АВТОМАТИЧЕСКОГО ПРОКТОРИНГА*

Айрат Ринатович АБЗАЛОВ [a,•], Алексей Владимирович ЖИГАНОВ [b], Рузиля Радисовна САМИГУЛЛИНА [c]

[a] старший преподаватель кафедры систем информационной безопасности,
Казанский национальный исследовательский технический университет им. А.Н. Туполева – КАИ,
Казань, Российская Федерация
abzalov@land.ru
ORCID: отсутствует
SPIN-код: 8422-7027

[b] студент Казанского национального исследовательского технического университета им. А.Н. Туполева – КАИ,
Казань, Российская Федерация
Lesha.159@uandex.ru
ORCID: отсутствует
SPIN-код: отсутствует

[c] студентка Казанского национального исследовательского технического университета им. А.Н. Туполева – КАИ,
Казань, Российская Федерация
Samigullina0304@mail.ru
ORCID: отсутствует
SPIN-код: отсутствует

[•] Ответственный автор

История статьи:
Рег. № 677/2019
Получена 23.10.2019
Получена в доработанном виде 14.11.2019
Одобрена 06.12.2019
Доступна онлайн 16.03.2020

УДК 519.876.5
JEL: I21, Y90

Ключевые слова:
клавиатурный почерк, биометрия, аутентификация, автоматический прокторинг, кластеризация

Аннотация

Предмет. Массовые онлайн-курсы и системы тестирования становятся неотъемлемой частью систем дистанционного обучения, которые все чаще основаны на системах автоматического прокторинга, обладающих уязвимостью аутентификации пользователей.

Цели. Исследование методов аутентификации пользователей по клавиатурному почерку. Разработка наиболее точного и эффективного метода аутентификации пользователей по клавиатурному почерку.

Методология. В статье предложена трехуровневая модель повышения точности аутентификации пользователей по клавиатурному почерку, позволяющая повысить эффективность аутентификации не только в системах автоматического прокторинга, но и в сложных критически важных объектах.

Результаты. Проведен эксперимент, в результате которого достоверность аутентификации пользователей составила 97,5%. Уже на статистическом уровне трехуровневой модели сократилось число нелегитимных пользователей в системе.

Выводы. На основе результатов исследования возможно (и планируется в дальнейшем) разработать метод, который будет функционировать на уровне логического сравнения в целях повышения точности. Он будет предназначен для создания усовершенствованного эталона, где бы учитывались особенности каждого пользователя, а также некоторые девиации, связанные с эмоциональными состояниями пользователей. Это могло бы способствовать разработке систем непрерывной аутентификации пользователей для мониторинга их эмоционального состояния при работе на критических объектах.

Для цитирования: Абзалов А.Р., Жиганов А.В., Самигуллина Р.Р. Аутентификация пользователей по клавиатурному почерку при использовании систем автоматического прокторинга // Национальные интересы: приоритеты и безопасность. – 2020. – Т. 16, № 3. – С. 582 – 596.
https://doi.org/10.24891/ni.16.3.582

Введение

Задачи идентификации пользователей при использовании информационных систем, в том числе систем дистанционного обучения, достаточно изучены, но остаются актуальными. Существующие решения основаны на ограниченном количестве методов и средств: электронных замках, аутентификации, технологиях биометрической идентификации. В последнее время широкое распространение получили методы биометрической идентификации пользователей, одним из которых является метод идентификации по клавиатурному почерку (прокторинг), то есть по характеру набора на клавиатуре произвольного текста или произвольной парольной фразы.

Принцип идентификации пользователей по клавиатурному почерку заключается в возможности проведения анализа временных характеристик нажатий клавиш при вводе парольной фразы. При многократном вводе одной и той же фразы подготовленный пользователь обычно осуществляет большую часть манипуляций с клавиатурой на бессознательном уровне, что и порождает эффект клавиатурного почерка, то есть пользователь формирует автоматический стереотип действий при вводе пароля. Контролируемыми параметрами клавиатурного ввода являются время нажатия каждой клавиши из пароля и временные интервалы между нажатием соседних клавиш. В работах[1] [1–7] показаны результаты исследований, в которых точность идентификации пользователей по клавиатурному почерку составляет более 97%. Существуют также и исследования зарубежных авторов. Так, например, Е. Флиором и К. Ковальским была разработана система идентификации пользователей по клавиатурному почерку, в которой из восьми тестируемых пользователей только трое были идентифицированы с вероятностью распознавания 37,5%. В работе Н. Д'Лима и Дж. Миталла описано, что только один из восьми пользователей был распознан, вероятность распознавания – 12,5%.

Современные системы автоматического прокторинга позволяют аутентифицировать пользователей, но в большинстве случаев – по логину и паролю. Метод аутентификации пользователей по клавиатурному почерку не требует никакого специального оборудования, но обладает недостатком – низкой точностью. В настоящей работе предлагается трехуровневая модель повышения точности аутентификации пользователей по клавиатурному почерку, позволяющая повысить эффективность аутентификации не только в системах автоматического прокторинга, но и в сложных критически важных объектах.

Обзор существующих решений в области аутентификации пользователей при использовании систем автоматического прокторинга

Автоматический прокторинг. Центр тестирования – это физическое место, где пользователи первоначально проходят аутентифкацию, а затем текущую, промежуточную или итоговую аттестацию с использованием специальных технических средств. Он позволяет осуществлять компьютерное тестирование, при котором местоположение и действия пользователя контролируются с помощью современных средств наблюдения[2] [8].

[*] Выражаем признательность сотрудникам кафедры систем информационной безопасности Казанского национального исследовательского технического университета им. А.Н. Туполева к.т.н., доценту Рифату Рашатовичу ШАРИПОВУ и к.т.н., доценту Марине Владимировне ТУМБИНСКОЙ за оказанное содействие и поддержку в процессе работы над статьей.

[1] *Петровский В.И., Тумбинская М.В.* История и перспективы развития системы защиты информации на предприятиях России: материалы III международной конференции «История вычислительной техники и ее программного обеспечения в России и странах бывшего СССР: история и перспективы. SoRuCom – 2014». Казань: Изд-во Казанского государственного технического университета им. А.Н. Туполева, 2015. С. 143–146; *Sharipov R.R., Tumbinskaya M.V., Abzalov A.R.* Analysis of Users Keyboard Handwriting Based on Gaussian Reference Signals. International Russian Automation Conference. RusAutoCon, 2019.

[2] *Томилин А.Н., Крайнева И.А., Трегубов В.М., Тумбинская М.В.* Вопросы истории естествознания и техники: материалы III международной конференции «История вычислительной техники и ее программного обеспечения в России и странах бывшего СССР: история и перспективы. SoRuCom – 2014». Казань: Изд-во Казанского государственного технического университета им. А.Н. Туполева, 2015. 173–180.

Автоматизированные системы контроля онлайн тестирования позволяют выявлять подмены личности в процессе тестирования (системы прокторинга, например ProctorEdu, Экзамус и др.). Подобные системы, подразумевающие наличие человека – проктора, который отслеживает процесс тестирования в режиме реального времени и выявляет нарушения, верифицируют пользователей в два этапа: 1) система сверяет пользователя перед камерой с фотографией, которую он загружал при подаче заявки на онлайн тестирование, 2) система фотографирует его непосредственно перед онлайн тестированием и сверяет его с эталоном. Однако системы прокторинга имеют ряд недостатков:

– их производительность зависит от числа прокторов, которые будут находиться онлайн или просматривать видеозаписи;

– задержка по времени при оценивании результатов тестирования (в случае асинхронного прокторинга нужно время на просмотр видеозаписей);

– сложность организации работы большого числа прокторов;

– недобросовестность или необязательность проктора, которые могут сказаться на качестве процесса.

Данное решение эффективно, но требует материальных затрат на приобретение технических средств, найма административного персонала, затрат на аренду помещения для организации центра тестирования. Решение не позволяет работать дистанционно и дает возможность «подтасовывать» результаты тестирования.

Физиологическая биометрия. Решение применяется для аутентификации пользователей на основе физиологических данных (отпечатков пальцев, структуры лица, радужной оболочки и др.). Данное решение доказало надежность только в части идентификации пользователей по распознаванию радужной оболочки глаз[3] [9–11], другие способы идентификации на основе физиологических данных не подтвердили своей надежности (см. работы Е. Флиорома и К. Ковальского, а также Н. Д'Лима и Дж. Миталла).

Поведенческая биометрия. Данное решение позволяет идентифицировать пользователя на основе его уникальных характеристик поведения (походки, голосового ритма, почерка, клавиатурного почерка, характера подписи и т.д.). В работах Е. Флиорома и К. Ковальского, а также Р.Р. Шарипова, А.С. Катасева [12] предложено использовать коэффициент корреляции для сравнения вновь напечатанных данных с данными, хранящимися в базе данных:

$$r = \frac{\sum_{i}^{n} (r_i \cdot t_i)}{\sqrt{\sum_{i=1}^{n} k_i^2 \cdot \sum_{i=1}^{n} t_i^2}}, \qquad (1)$$

где k_i – вектор длины ($i = \overline{1,n}$), отражающий интервал времени (разницу между отпусканием и последующим нажатием клавиши) между нажатиями клавиш в эталонной подписи; t_i – вектор длины ($i = \overline{1,n}$), в котором хранятся временные интервалы между нажатиями клавиш в пробной подписи. Данное решение неэффективно при работе с большими объемами данных.

Авторы работ [5, 11, 13], учитывая скорость набора текста, частоту нажатия клавиш, движения мыши и времени задержки нажатия клавиш, предложили модель на основе нейронной сети [14–16]. Результаты исследований показали, что пользователи были успешно идентифицированы в случаях, когда сходство между сохраненным шаблоном и текущими данными было более 90%. Модель на основе нейронной сети содержала модуль дополнения, который обновлял шаблон

[3] *Sharipov R.R., Tumbinskaya M.V., Abzalov A.R.* Analysis of Users Keyboard Handwriting Based on Gaussian Reference Signals. International Russian Automation Conference. RusAutoCon, 2019.

каждый раз, когда пользователь успешно проходил идентификацию. Решение неэффективно для небольших наборов данных из-за сложности обобщения.

В настоящей работе предлагается трехуровневая модель повышения точности аутентификации пользователей, состоящая из трех этапов:

1) аутентификация пользователей на основе сравнения текущих параметров пользователя с эталонными (формируемыми на этапе регистрации в системе);

2) аутентификация пользователей на основе соотнесения параметров пользователя с классами (легальные и нелегальные – «чужие»);

3) обновление параметров и критериев класса легального пользователя на основе метода рекуррентного вычисления.

Трехуровневая модель повышения точности идентификации пользователей по клавиатурному почерку

Для аутентификации пользователей при использовании онлайн информационных систем предлагается трехуровневая модель аутентификации. На *рис. 1* представлена схема трехуровневой модели системы аутентификации пользователей.

Перед началом использования системы автоматического прокторинга пользователю предлагается пройти авторизацию в системе (ввод логина, пароля и других параметров). После успешной авторизации пользователь может пройти аутентификацию, при этом для подтверждения аутентификации измеряются клавиатурные параметры пользователя, формируются эталоны и выполняется сравнение параметров пользователей по критерию сравнения центров распределения двух совокупностей при допущении, что распределение параметров клавиатурного почерка подчинены нормальному гауссовскому закону (стандартное отклонение) и выполняется соответствие на критерий χ^2 («хи-квадрат»). Вторым уровнем является решение классовой задачи для сравнения текущих данных с данными классов пользователей в случае, если на первом уровне пользователь не прошел аутентификацию.

На третьем уровне происходит обновление данных пользователей на основе процесса рекуррентности.

На *рис. 2* представлена структурная схема детализации трехуровневой модели аутентификации пользователей.

Опишем более подробно каждый уровень разработанной модели.

Статистический уровень. На этом уровне измеряются временные параметры клавиатурного почерка пользователей и формируются эталоны пользователей в виде измерения средних значений и среднеквадратических отклонений (СКО). Далее происходит анализ стандартных отклонений между текущими входными данными и эталонными значениями. Стандартное отклонение временных параметров динамики работы пользователей на клавиатуре вычисляется по формуле

$$S = \sqrt{\frac{1}{N}\sum_{i=1}^{n}\left(x_i - \overline{x_j}\right)^2}, \qquad (2)$$

где S – стандартное отклонение; N – размер выборки из временных параметров; x_i – текущие значение выборки; $\overline{x_j}$ – среднее значение параметров.

Вычисленное значение S сравнивается с сохраненным стандартным отклонением, которое рассчитывается на основе ранее собранных данных конкретного пользователя. Если разница между этими двумя значениями незначительна (в данном случае меньше 3), пользователь принимается, в противном случае переходим к проверке критерия χ^2.

Следующий процесс включает в себя выполнение теста χ^2, который используется для подтверждения или опровержения гипотезы. В предлагаемой системе в качестве

гипотезы для теста принимается предположение «Пользователь легальный». Данный тест предполагает одну степень свободы, поскольку возможны только два возможных результата: пользователь легальный или же нелегальный. Степень свободы используется для выбора критического значения в статистической таблице[4].

Критическое значение выбирается в зависимости от желаемого уровня точности. Для предложенной системы был выбран уровень достоверности 97,5%, который означает вероятность того, что решение о принятии или отклонении гипотезы является верным. Чтобы доказать, верна ли гипотеза, необходимо использовать уже сохраненное среднее значение атрибута в качестве ожидаемого значения и среднее значение входного атрибута в качестве наблюдаемого значения. Цель теста – выяснить, является ли разница между наблюдаемым средним (входным) и ожидаемым средним (сохраненным) результатом случайности или других факторов. Далее представлена формула, где χ^2 – это значение критерия «хи-квадрат»; $\overline{x_j}$ – среднее значение временного параметра клавиатурного почерка (наблюдаемое значение); L_i – среднее значение, сохраненное в эталоне пользователя (ожидаемое значение)[5]:

$$\chi^2 = \sum \frac{(\overline{x_j} - L_i)^2}{L_i}. \quad (3)$$

Значение χ^2 должно быть меньше или равно критическому значению, выбранному из статистической таблицы [2], чтобы принять гипотезу. Если значение χ^2 меньше или равно критическому значению, пользователь принимается с вероятностью 97,5%, в противном случае пользователь отклоняется и переходит на следующий уровень модели аутентификации.

Уровень классификации. На этом уровне определяется набор классифицируемых признаков наборных характеристик пользователей в составлении априорного словаря классов. Основное в данной задаче – это выбор надлежащего принципа классификации этих характеристик. Последнее определяется требованиями, предъявляемыми к системе распознавания, которые, в свою очередь, зависят от того, какие решения могут приниматься системой управления по результатам распознавания неизвестных характеристик клавиатурного почерка. Далее осуществляют определение параметров классов путем разбиения характеристик почерка на классы $L_1, \ldots L_m$. Требуется выделить в пространстве наборных характеристик области S_i (i = 1, …, m), эквивалентные классам, то есть характеризуемые следующей необходимой зависимостью: если характеристики клавиатурного почерка, имеющие параметры $x^0_1, \ldots, x^0_N$, относятся к классу L_i, то представляющее его в пространстве наборных характеристик значение принадлежит области S_i. Пример разбиения классов для представленной задачи представлен на *рис. 3*.

Таким образом можно определить решающую границу между областями S_i, соответствующих классам L_i,

На *рис. 3* показано разбиение двумерного пространства наборных характеристик клавиатурного почерка на области S_1, S_2, соответствующих классам L_1, L_2.

В области S_1 находятся параметры клавиатурного почерка первого класса легальных пользователей, а в области S_2 – параметры случайных незарегистрированных пользователей. Здесь 1 – решающая граница.

Задав разделяющие функции, можно проверить выполнение следующих условий:

Если $F_1(x_1, x_2) > F_2(x_1, x_2)$ – то параметры относятся к легальному пользователю

[4] *Берман Г.Н.* Сборник задач по курсу математического анализа. М.: Наука, 1977. 416 с.

[5] Там же.

системы, а если $F_1(x_1, x_2) < F_2(x_1, x_2)$, то текущий пользователь (проходящий процедуру идентификации) распознается как «чужой» и игнорируется системой.

В противном случае, если же пользователь опознан как «свой», происходит обновление параметров $\overline{x_j}$ и S, сохраненных как эталонные значения, в соответствии с рекуррентными отношениями вида:

$$\overline{x_j} = \frac{1}{N}\sum_{i=1}^{N} x_i ; \quad (4)$$

$$S = \frac{1}{N-1}\cdot\sum_{i=1}^{N} |x_i - \overline{x_j}| , \quad (5)$$

где $\overline{x_j}$ – среднее значение параметров пользователя, N – количество наблюдений, x_i – текущее значение параметров пользователя, S – среднеквадратичное отклонение (обновленное).

Экспериментальная часть

Для проверки работоспособности трехуровневой модели распознавания были проведены экспериментальные исследования и численно-параметрические расчеты. Для эксперимента была выбрана группа из восьми студентов Института компьютерных технологий и защиты информации (КНИТУ-КАИ), владеющих десятипальцевым слепым набором слов на клавиатуре. Сначала им предложили войти на специально разработанный сайт, авторизоваться, в процессе чего формировались эталоны данных пользователей по результатам их работы на клавиатуре. Студенты должны были указать свои ФИО, направление подготовки или специализацию, год обучения, а затем они перенаправлялись на другую страницу, где ими осуществлялся ввод предложенных на экране готовых текстов. Эти тексты представляли собой некие, заранее подготовленные предложения, которые были не известны тестируемой группе.

Далее осуществлялся сбор введенных текущих данных тестируемой группой на стороне клиента из веб-формы. Когда студент набирал символы на клавиатуре, программа-обработчик, разработанная на языке JavaScript, обрабатывала каждое нажатие и отпускание клавиш. Для этого использовались стандартные встроенные функции: *Keydown* – событие при нажатии клавиши, и *Keyup* – событие при отпускании клавиши (*рис 4*).

При нажатии клавиши программа фиксировала первую временную метку, а при отпускании клавиши – ставила вторую временную метку, после этого вычислялась разность времени между вторым и первым событиями, что и являлось временем удержания клавиши. Кроме этого, используя встроенные функции, программа определяла код нажатой клавиши и ее символьное обозначение. Затем JavaScript-функция отправляла код клавиши, имя клавиши, время задержки клавиши и временное значение отпускания клавиши на сервер, используя при этом AJAX-запрос для дальнейшей обработки этих данных.

Следующим этапом осуществлялся сбор данных на стороне сервера. Веб-сервер был реализован на основе Apache HTTP Server 2.2 в локальной вычислительной сети на базе операционной системы Linux Mint. Обработка данных на сервере проходила на распространенном скриптовом языке PHP. При поступлении данных от студента на стороне сервера скрипт запрашивал базу данных для проверки, реализованной на MySQL, для проверки, не были ли какие-либо предыдущие данные (нажатые клавиши) уже сохранены для текущего пользователя. В случае, если данные уже в базе были, обработчик выбирал последние три и использовал их временные метки для расчета временных интервалов диграфов, тригафов и квадграфов путем разности значений временных меток.

В процессе эксперимента статистические данные были разделены на две части:

1) данные для обработки;

2) эталонные данные.

Был выбран параметр «время удержания клавиш студентом» наиболее информативный в нашем случае, хотя это утверждение может быть неверно при работе с большим объемом данных.

По результатам тестов для каждого студента были рассчитаны и выстроены стандартные девиации текущих и эталонных данных. Результаты для всех восьми тестируемых пользователей представлены в *табл. 1*. Кроме того, был рассчитан критерий χ^2 для каждого тестируемого пользователя. Из данных *табл. 1* видно, что пользователь под № 6 был аутентифицирован путем сравнения стандартных девиаций между его данными и эталоном, потому что разница в девиации была менее 3, что и явилось первым шагом в процессе аутентификации.

Пользователи № 3, 5 и 7 были аутентифицированы с использованием критерия χ^2, поскольку их значения критерия были меньше, чем критическое значение – 12,71, и тогда с уверенностью в 97,5% можно утверждать, что пользователи идентифицировались верно. Результаты расчетов показывают, что половина пользователей прошла проверку подлинности в процессе сравнения статистических значений параметров клавиатурного почерка. Однако следует иметь в виду, что предложенная модель будет совершенствоваться по мере обработки большего числа данных, так как полученные статистические значения для каждого пользователя будут более верными.

Первоначальной целью исследования было – убедиться в способности предложенной модели разграничить доступ для легальных пользователей и мошенников, приспосабливаясь к незначительным изменениям параметров клавиатурного почерка, увеличивая тем самым точность аутентификации. В результате один из тестируемых студентов был подвержен сравнению девиаций, в то время как другие три были аутентифицированы с использованием критерия χ^2. Остальные пользователи не прошли процедуры аутентификации на всех уровнях системы, что может быть обусловлено недостаточным количеством идентификационных данных. В случае работы с большим набором идентификационных данных процент безуспешного распознавания будет сведен к минимуму.

Заключение

Результаты исследований показали, что половина данных была обработаны на статистическом уровне предложенной модели, которая является более точной в вычислительном отношении по сравнению с некоторыми известными решениями.

Проблемы с эффективностью можно решить и оптимизировать при разработке больших систем, а простота алгоритмов и программной реализации разработанной модели позволит сэкономить вычислительную мощность в крупномасштабных системах. Применение критерия χ^2 явилось важным дополнением к статистическому сравнению из-за его способности исследовать статистические вариации, что и отразилось в нашем исследовании, где три тестируемых студента прошли процесс аутентификации после оценки критерия χ^2.

В дальнейшем планируется разработать метод, который будет работать на уровне логического сравнения в целях повышения точности. Он будет предназначен для создания усовершенствованного эталона, в котором бы учитывались особенности каждого пользователя, а также некоторые девиации, связанные с эмоциональными состояниями пользователей. Это способствовало бы разработке систем непрерывной аутентификации пользователей для мониторинга их эмоционального состояния при работе на критических объектах.

Таблица 1
Значение девиации и критерий χ^2

Table 1
The deviation value and the χ^2 test

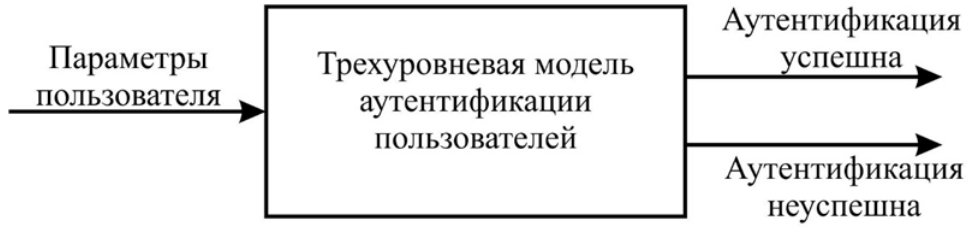
№ испытуемого	Значение девиации (эталонные данные)	Значение девиации (измеренное)	Критерий χ^2
1	86	32	28,7
2	169	41	31,42
3	63	9	5,95
4	62	48	51,41
5	162	35	11,84
6	39	38	13,86
7	42	12	4,82
8	105	118	63,31

Источник: авторская разработка

Source: Authoring

Рисунок 1
Трехуровневая модель аутентификации пользователей

Figure 1
The three-tiered user identification model

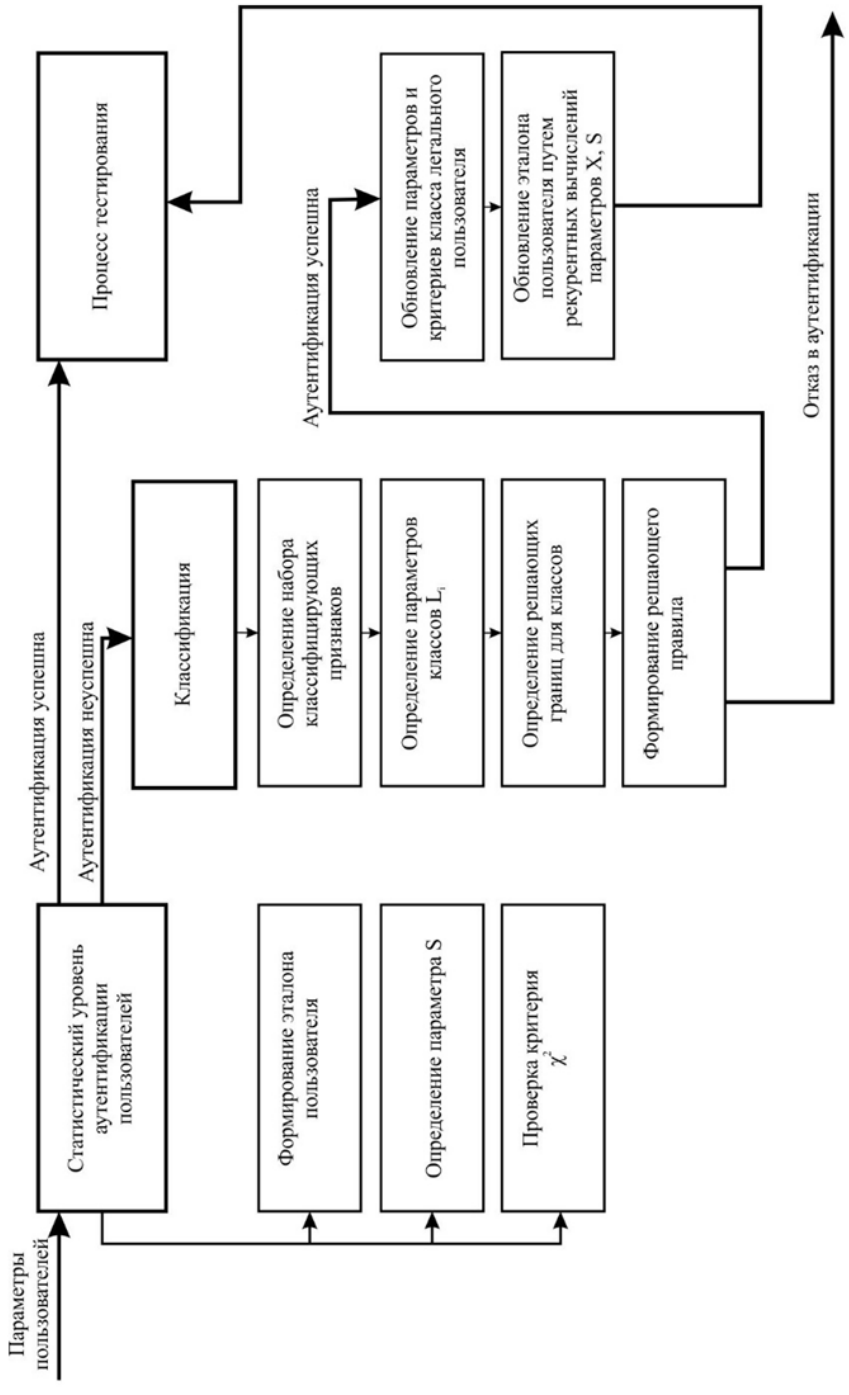


Источник: авторская разработка

Source: Authoring

Рисунок 2
Структурная схема детализации трехуровневой модели аутентификации пользователей

Figure 2
Block diagram of detailing the three-tiered user identification model

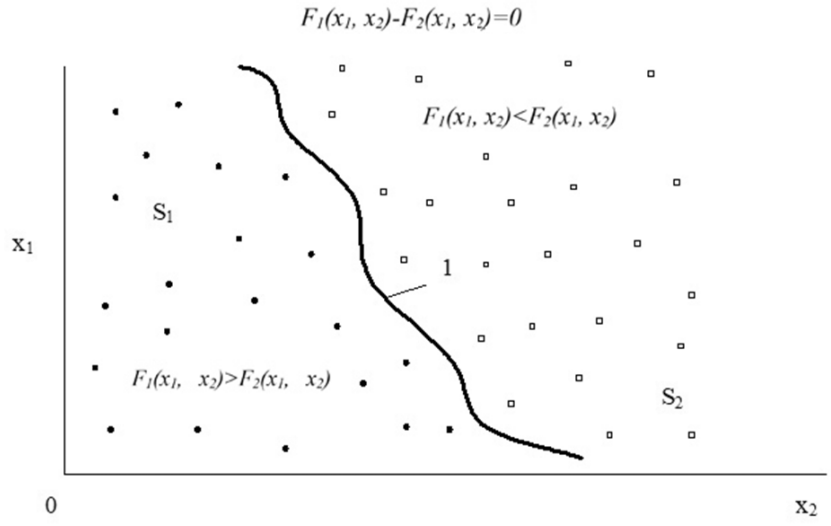


Источник: авторская разработка

Source: Authoring

Рисунок 3
Разбиение классов пользователей

Figure 3
A user classification



Источник: авторская разработка

Source: Authoring

Рисунок 4
Код обработчика событий *Keyup* и *Keydown*

Figure 4
***Keyup* and *Keydown* events handler code**

```
kinput.onkeydown = kinput.onkeyup = kinput.onkeypress = handle;
let lastTime = Date.now();
function handle(e) {
if (form.elements [e.type + 'Ignore'].checked) return;
let text = e.type + ' key=' + e.key + ' code=' + e.code +
(e.shiftKey ? ' shiftKey' : '') +
(e.ctrlKey ? ' ctrlKey' : '') +
(e.altKey ? ' altKey' : '') +
(e.metaKey ? ' metaKey' : '') +
(e.repeat ? ' (repeat)' : '') + "\n";
if (area.value && Date.now() - lastTime > 250) {
area.value += new Array(81).join('-') + '\n';
}
lastTime = Date.now();
area.value += text;
if (form.elements[e.type + 'Stop'].checked) {e.preventDefault();
}
```

Источник: авторская разработка

Source: Authoring

Список литературы

1. *Глова В.И., Катасев А.С., Абзалов А.Р.* Модель локализации искажений и подтверждения авторства исходных текстов программ // Вестник Казанского государственного технического университета им. А.Н. Туполева. 2008. № 3. С. 84–87.

2. *Добровинский Д.С., Ловецкий И.В., Попов М.А.* Прокторинг как инструмент развития дистанционного образования // Научно-техническое и экономическое сотрудничество стран АТР в XXI веке. 2018. Т. 2. С. 27–32.

3. *Клименских М.В., Истомин Д.В., Халфин А.Б., Панченко В.Н.* Обеспечение процедуры дистанционного экзаменационного мероприятия посредством методов идентификации личности студента // Вестник Казанского государственного энергетического университета. 2016. № 3. С. 134–151.

4. *Костюченко Е.Ю., Мещеряков Р.В.* Распознавания пользователя по клавиатурному почерку на фиксированной парольной фразе в компьютерных системах // Известия ЮФУ. Технические науки. 2003. № 4. С. 177–178. URL: https://cyberleninka.ru/article/n/biometricheskie-metody-identifikatsii/viewer

5. *Тумбинская М.В., Баянов Б.И., Рахимов Р.Ж. и др.* Анализ и прогнозирование вредоносного сетевого трафика в облачных сервисах // Бизнес-информатика. 2019. № 1. С. 71–81. URL: https://doi.org/10.17323/1998-0663.2019.1.71.81

6. *Шарипов Р.Р., Катасев А.С., Кирпичников А.П.* Методы анализа клавиатурного почерка пользователей с использованием эталонных гауссовских сигналов // Вестник Технологического университета. 2016. Т. 19. № 13. С. 157–160. URL: https://cyberleninka.ru/article/n/metody-analiza-klaviaturnogo-pocherka-polzovateley-s-ispolzovaniem-etalonnyh-gaussovskih-signalov/viewer

7. *Тумбинская М.В.* Защита информации в социальных сетях от социоинженерных атак злоумышленника // Прикладная информатика. 2017. Т. 12. № 3. С. 88–102.

8. *Шарипов Р.Р., Катасев А.С.* Анализ клавиатурного почерка пользователей инфокоммуникационных систем на основе полигауссового алгоритма // Информация и безопасность. 2016. Т. 19. № 4. С. 587–590.

9. *Тумбинская М.В.* Системный подход к обеспечению защиты от нежелательной информации в социальных сетях // Вопросы кибербезопасности. 2017. № 2. С. 30–44. URL: https://cyberleninka.ru/article/n/sistemnyy-podhod-k-obespecheniyu-zaschity-ot-nezhelatelnoy-informatsii-v-sotsialnyh-setyah/viewer

10. *Тумбинская М.В.* Модель защищенной информационной системы интернет-банкинга // Прикладная информатика. 2015. № 5. С. 62–72. URL: https://cyberleninka.ru/article/n/model-zaschischennoy-informatsionnoy-sistemy-internet-bankinga/viewer

11. *Шарипов Р.Р., Сафиуллин Н.З.* Аппаратурный анализ клавиатурного почерка с использованием эталонных гауссовских сигналов // Вестник Казанского государственного технического университета им. А.Н. Туполева. 2006. № 2. С. 21–23.

12. *Шарипов Р.Р., Катасев А.С.* Система распознавания клавиатурного почерка пользователей на основе полигауссового алгоритма // Вестник Казанского государственного энергетического университета. 2016. № 4. С. 45–59. URL: https://cyberleninka.ru/article/n/sistema-raspoznovaniya-klaviaturnogo-pocherka-polzovateley-na-osnove-poligaussovogo-algoritma/viewer

13. *Тумбинская М.В.* Обеспечение защиты от нежелательной информации в социальных сетях // Вестник Мордовского университета. 2017. Т. 27. № 2. С. 264–288.

14. *Мухаматханов Р.М., Михайлов А.А., Баянов Б.И., Тумбинская М.В.* Классификация DDoS-атак на основе нейросетевой модели // Прикладная информатика. 2019. № 1. С. 96–103.

15. *Кормильцев Н.В., Уваров А.Д., Хаматнуров И.И., Тумбинская М.В.* Анализ защищенности системы безопасности LTE-а от направленного воздействия DDoS-атак // Национальные интересы: приоритеты и безопасность. 2019. Т. 15. Вып. 2. С. 376–392. URL: https://doi.org/10.24891/ni.15.2.376

16. *Александрова Л.А., Тумбинская М.В.* Модель интерактивной обучающей системы // Программные продукты и системы. 2009. № 2. С. 39. URL: https://cyberleninka.ru/article/n/model-interaktivnoy-obuchayuschey-sistemy/viewer

Информация о конфликте интересов

Мы, авторы данной статьи, со всей ответственностью заявляем о частичном и полном отсутствии фактического или потенциального конфликта интересов с какой бы то ни было третьей стороной, который может возникнуть вследствие публикации данной статьи. Настоящее заявление относится к проведению научной работы, сбору и обработке данных, написанию и подготовке статьи, принятию решения о публикации рукописи.

pISSN 2073-2872
eISSN 2311-875X

Threats and Security

USER IDENTIFICATION THROUGH KEYSTROKE DYNAMICS AS PART OF AUTOMATED PROCTORING SYSTEMS

Airat R. ABZALOV [a,•], Aleksei V. ZHIGANOV [b], Ruzilya R. SAMIGULLINA [c]

[a] Kazan National Research Technical University named after A.N. Tupolev – KAI (KNRTU-KAI), Kazan, Republic of Tatarstan, Russian Federation
abzalov@land.ru
ORCID: not available

[b] Kazan National Research Technical University named after A.N. Tupolev – KAI (KNRTU-KAI), Kazan, Republic of Tatarstan, Russian Federation
Lesha.159@yandex.ru
ORCID: not available

[c] Kazan National Research Technical University named after A.N. Tupolev – KAI (KNRTU-KAI), Kazan, Republic of Tatarstan, Russian Federation
Samigullina0304@mail.ru
ORCID: not available

[•] Corresponding author

Article history:
Article No. 677/2019
Received 23 October 2019
Received in revised form
14 November 2019
Accepted 6 December 2019
Available online
16 March 2020

JEL classification: I21, Y90

Keywords: keystroke dynamics, biometrics, authentication, automated proctoring, clustering

Abstract

Subject Popular online courses and testing programs integrate into correspondence education systems, which are more often than not based on automated proctoring. What makes the latter vulnerable is user identification.

Objectives We examine user identification methods through keystroke dynamics and devise a more accurate and effective technique for user identification through keystroke dynamics.

Methods The article sets out a three-tiered model for identifying users more accurately not only in automated proctoring environments, but also in critically sensitive locations.

Results We had an experiment, which showed a 97.5 percent accuracy of user identification. We significantly reduced illegitimate users at the statistical level of the three-tiered model.

Conclusions and Relevance Following the study, it is possible to develop a logic comparison method for higher accuracy. It will serve for creating a more refined model, which would accommodate for distinctions of each user and some deviations of users' emotions. This would contribute to continuous user identification systems to monitor their emotional condition at critical locations.

Please cite this article as: Abzalov A.R., Zhiganov A.V., Samigullina R.R. User Identification Through Keystroke Dynamics as Part of Automated Proctoring Systems. *National Interests: Priorities and Security*, 2020, vol. 16, iss. 3, pp. 582–596.
https://doi.org/10.24891/ni.16.3.582

Acknowledgments

We express our gratitude to the personnel of the Department for Data Security Systems at the Kazan National Research Technical University named after A.N. Tupolev – Associate Professor Rifat R. SHARIPOV, and Candidate of Technical Sciences, Associate Professor, Marina V. TUMBINSKAYA, for their assistance and support when we were preparing the article.

References

1. Glova V.I., Katasev A.S., Abzalov A.R. [Model for localization of distortions and authorship acknowledgement of source program texts]. *Vestnik Kazanskogo gosudarstvennogo tekhnicheskogo universiteta im. A.N. Tupoleva = Vestnik of Kazan State Technical University named after A.N. Tupolev*, 2008, no. 3, pp. 84–87. (In Russ.)

2. Dobrovinskii D.S., Lovetskii I.V., Popov M.A. [Proctoring as a tool of distance education development]. *Nauchno-tekhnicheskoe i ekonomicheskoe sotrudnichestvo stran ATR v XXI veke = Scientific, Technical and Economic Cooperation of the Asia-Pacific Countries in the 21st Century*, 2018, vol. 2, pp. 27–32. (In Russ.)

3. Klimenskikh M.V., Istomin D.V., Khalfin A.B., Panchenko V.N. [Provision of procedure in distance examination by the methods of identification student personality]. *Vestnik Tekhnologicheskogo universiteta = Bulletin of the Kazan National Research Technological University*, 2016, no. 3, pp. 134–151. (In Russ.)

4. Kostyuchenko E.Yu., Meshcheryakov R.V. [User recognition by handwriting from a keyboard on a fixed key phrase in computer systems]. *Izvestiya YuFU. Tekhnicheskie nauki = Izvestiya SfedU. Engineering Sciences*, 2003, no. 4, pp. 177–178. URL: https://cyberleninka.ru/article/n/biometricheskie-metody-identifikatsii/viewer (In Russ.)

5. Tumbinskaya M.V., Bayanov B.I., Rakhimov R.Zh. et al. [Analysis and forecast of undesirable cloud services traffic]. *Biznes-informatika = Business Informatics*, 2019, no. 1, pp. 71–81. (In Russ.) URL: https://doi.org/10.17323/1998-0663.2019.1.71.81

6. Sharipov R.R., Katasev A.S., Kirpichnikov A.P. [Methods for analyzing keystroke dynamics of users through the model Gaussian signals]. *Vestnik Tekhnologicheskogo universiteta = Bulletin of the Kazan National Research Technological University*, 2016, no. 13, pp. 157–160. URL: https://cyberleninka.ru/article/n/metody-analiza-klaviaturnogo-pocherka-polzovateley-s-ispolzovaniem-etalonnyh-gaussovskih-signalov/viewer (In Russ.)

7. Tumbinskaya M.V. [Protection of information in social networks from socio-engineer attacks of the attacker]. *Prikladnaya informatika = Applied Informatics*, 2017, vol. 12, no. 3, pp. 88–102. (In Russ.)

8. Sharipov R.R., Katasev A.S. [Analyzing keystroke dynamics of users in ICT systems through the poly-Gaussian algorithm]. *Informatsiya i bezopasnost' = Information and Security*, 2016, no. 4, pp. 587–590. (In Russ.)

9. Tumbinskaya M.V [System approach to protection against unwanted information in the social networks]. *Voprosy kiberbezopasnosti*, 2017, no. 2, p. 30–44. URL: https://cyberleninka.ru/article/n/sistemnyy-podhod-k-obespecheniyu-zaschity-ot-nezhelatelnoy-informatsii-v-sotsialnyh-setyah/viewer (In Russ.)

10. Tumbinskaya M.V. [Secure information system model of Internet banking]. *Prikladnaya informatika = Applied Informatics*, 2015, no. 5, pp. 62–72. URL: https://cyberleninka.ru/article/n/model-zaschischennoy-informatsionnoy-sistemy-internet-bankinga/viewer (In Russ.)

11. Sharipov R.R., Safiullin N.Z. [Hardware analysis of keyboard script by using the reference Gaussian signals]. *Vestnik Kazanskogo gosudarstvennogo tekhnicheskogo universiteta im. A.N. Tupoleva = Vestnik of Kazan State Technical University named after A.N. Tupolev*, 2006, no. 2, pp. 21–23. (In Russ.)

12. Sharipov R.R., Katasev A.S. [User keyboard handwriting recognition system based on poly Gaussian algorithm]. *Vestnik Kazanskogo gosudarstvennogo energeticheskogo universiteta = Bulletin of Kazan State Power Engineering University*, 2016, no. 4, pp. 45–59. URL: https://cyberleninka.ru/article/n/sistema-raspoznovaniya-klaviaturnogo-pocherka-polzovateley-na-osnove-poligaussovogo-algoritma/viewer (In Russ.)

13. Tumbinskaya M.V. [Providing protection from targeted information in social networks]. *Vestnik Mordovskogo universiteta = Mordovia University Bulletin Journal*, 2017, vol. 27, no. 2, pp. 264–288. (In Russ.)

14. Mukhamatkhanov R.M., Mikhailov A.A., Bayanov B.I., Tumbinskaya M.V. [Classification of DDoS attacks based on the neural network model]. *Prikladnaya informatika = Applied Informatics*, 2019, no. 1, pp. 96–103. (In Russ.)

15. Kormil'tsev N.V., Uvarov A.D., Khamatnurov I.I., Tumbinskaya M.V. [Analysis of the protection of the LT-A security system from the targeted exposure of DDoS attack]. *Natsional'nye interesy: prioritety i bezopasnost' = National Interests: Priorities and Security*, 2019, vol. 15, iss. 2, pp. 376–392. (In Russ.) URL: https://doi.org/10.24891/ni.15.2.376

16. Aleksandrova L.A., Tumbinskaya M.V. [Interactive Learning System Model]. *Programmnye produkty i sistemy = Software and Systems*, 2009, no. 2, p. 39. URL: https://cyberleninka.ru/article/n/model-interaktivnoy-obuchayuschey-sistemy/viewer (In Russ.)

Conflict-of-interest notification

We, the authors of this article, bindingly and explicitly declare of the partial and total lack of actual or potential conflict of interest with any other third party whatsoever, which may arise as a result of the publication of this article. This statement relates to the study, data collection and interpretation, writing and preparation of the article, and the decision to submit the manuscript for publication.