УДК 004.056
DOI: 10.15827/0236-235X.130.266-275
Дата подачи статьи: 03.12.19
2020. Т. 33. № 2. С. 266–275

Моделирование аутентификации пользователей по динамике нажатий клавиш в промышленных автоматизированных системах

М.В. Тумбинская [1], *к.т.н., доцент кафедры «Системы информационной безопасности», tumbinskaya@inbox.ru*
Н.Ф. Асадуллин [1], *студент, nail.asadullin.1997@mail.ru*
Р.Р. Муртазин [1], *студент, ramichhh@gmail.com*

[1] *Казанский национальный исследовательский технический университет, г. Казань, 420111, Россия*

Современные промышленные автоматизированные системы управления используются повсеместно. Сложная архитектура таких систем, требования непрерывности процесса и доступа к сети Интернет делают их легкоуязвимыми для злоумышленников и кибератак. В настоящее время компоненты промышленных автоматизированных систем управления не являются полностью защищенными, следовательно, возникает потребность в их адекватной защите и повышении уровня информационной безопасности.

Как показывает анализ предметной области, 82 % промышленных предприятий не могут противостоять внутреннему нарушителю, который стремится проникнуть в технологическую сеть из корпоративной. Получив доступ к технологическому сегменту сети, он имеет широкие возможности злонамеренного влияния на компоненты автоматизированных систем управления. В 2018 году количество новых уязвимостей в компонентах промышленных автоматизированных систем увеличилось на 30 % по сравнению с 2017 годом. Значительная доля уязвимостей в промышленных автоматизированных системах управления связана с некорректной аутентификацией или избыточными правами доступа пользователей. При этом больше половины уязвимостей могут эксплуатироваться удаленно.

В статье представлен обзор уязвимостей информационной безопасности промышленных автоматизированных систем. На основе анализа предметной области авторы предлагают трехуровневую модель повышения точности аутентификации пользователей по динамике нажатия клавиш, которая позволит повысить уровень безопасности автоматизированных систем управления. Проведено экспериментальное исследование, показавшее высокую способность предложенной модели разграничения доступа для легальных пользователей и злоумышленников с учетом незначительных изменений параметров динамики нажатия клавиш повысить точность аутентификации пользователей. Достоверность аутентификации и пользователей на практике составила 97,5 %.

Ключевые слова: *точность аутентификации, клавиатурный почерк, динамика нажатия клавиш, биометрия, аутентификация, кластеризация.*

Задачи идентификации пользователей при использовании информационных систем, в том числе промышленных автоматизированных систем управления, достаточно изучены, но остаются актуальными. Существующие решения основаны на ограниченном количестве методов и средств: электронных замках, аутентификации, технологиях биометрической идентификации [1, 2]. В последнее время широкое распространение получили методы биометрической идентификации пользователей, одним из которых является метод идентификации по динамике нажатия клавиш, то есть по характеру набора на клавиатуре произвольного текста или произвольной парольной фразы.

Принцип идентификации пользователей по динамике нажатия клавиш 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 при вводе пароля пользователь формирует автоматический стереотип действий. Контролируемыми параметрами клавиатурного ввода являются время нажатия каждой клавиши из пароля и временные интервалы между нажатием соседних клавиш. В работах [3–5] показаны результаты исследований, в которых точность идентификации пользователей по клавиатурному почерку составляет более 97 %. Существуют исследова-

ния и зарубежных авторов. Так, например, в работе [6] описана система идентификации пользователей по динамике нажатия клавиш, и из восьми тестируемых пользователей только три были идентифицированы, вероятность распознавания составила 37,5 %. В работе [7] описано, что только один из восьми пользователей был распознан, вероятность распознавания – 12,5 %.

Современные промышленные автоматизированные системы управления позволяют аутентифицировать пользователей, как правило, по логину и паролю. Метод аутентификации пользователей по динамике нажатия клавиш не требует никакого специального оборудования, но обладает недостатком – низкой точностью. В данной работе предлагается трехуровневая модель повышения точности аутентификации пользователей по динамике нажатия клавиш, позволяющая повысить эффективность аутентификации не только в автоматизированных системах управления, но и в сложных, критически важных и потенциально опасных объектах промышленности.

Обзор существующих решений в области аутентификации пользователей в автоматизированных системах управления

Анализ работ [8, 9] показал, что 82 % промышленных предприятий не могут противостоять внутреннему нарушителю, который стремится проникнуть в технологическую сеть из корпоративной. После получения доступа к технологическому сегменту сети у злоумышленника появляются широкие возможности по злонамеренному влиянию на компоненты автоматизированных систем управления.

По сравнению с 2017 годом в 2018 году количество новых уязвимостей в компонентах промышленных автоматизированных систем увеличилось на 30 %. Как правило, анализ таких систем выявляет несколько уязвимостей. Наиболее часто встречающиеся уязвимости представлены в таблице 1.

С каждым годом наблюдается рост случаев обнаружения уязвимостей в автоматизированных системах управления: 2013 г. – 158 случаев, 2014 г. – 181, 2015 г. – 212, 2016 г. – 115, 2017 г. – 197, 2018 г. – 257.

Значительная доля уязвимостей связана с некорректной аутентификацией или избыточными правами доступа пользователей. При этом больше половины из них (64 %) могут эксплуатироваться удаленно.

В таблице 2 на основе анализа данных из работ [10, 11] представлена информация о распределениях уязвимостей в автоматизированных системах управления по типу.

В настоящее время компоненты промышленных автоматизированных систем управления не являются полностью защищенными, следовательно, возникает потребность в их адекватной защите, повышении уровня информационной безопасности.

Физиологическая биометрия. Решение применяется для аутентификации пользователей на основе физиологических данных (отпечатки пальцев, структура лица, радужная оболочка и др.). Данное решение доказало надежность только в части идентификации пользователей по распознаванию радужной оболочки глаз [5, 12], другие способы идентификации на основе физиологических данных не подтвердили своей надежности [7, 13, 14].

Поведенческая биометрия. Данное решение позволяет идентифицировать пользователя на основе уникальных характеристик его поведения (походка, голосовой ритм, почерк, клавиатурный почерк, характер подписи и т.д.). В работах [13, 15, 16] предложено использовать коэффициент корреляции для сравнения вновь напечатанных данных с данными, хранящимися в БД:

$$r = \sum_i^n (k_i * t_i) / \sqrt{\sum_{i=1}^{n} k_i^2 * \sum_{i=1}^{n} t_i^2}, \quad (1)$$

где k_i – вектор длины ($i = \overline{1,n}$), отражающий интервал времени (разница 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 [5–7], учитывая скорость набора текста, частоту нажатия клавиш, движения мыши и время задержки нажатия клавиш, предложили модель на основе нейронной сети [16, 17]. Результаты исследований показали, что пользователи 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 каждый раз, когда пользователь успешно проходил идентификацию. Решение неэффективно для небольших наборов данных из-за сложности обобщения.

Таблица 1

Уязвимости, выявленные в промышленных автоматизированных системах управления в 2018 году

Table 1

Vulnerabilities identified in industrial automated control systems in 2018

Уязвимость	Рейтинг опасности CVSS	ПО
PT-2018-38: Раскрытие информации	9.8	MGE Galaxy 3000, 4000, 5000, 6000, 9000 MGE EPS 6000, 7000, 8000 MGE Comet UPS MGE Galaxy PW STS (MGE Upsilon)
PT-2018-37: Несанкционированные действия	7.5	
PT-2018-36: Раскрытие информации	5.3	
PT-2018-35: Обход авторизации	10	
PT-2018-31: Внедрение внешних сущностей	5.3	Cisco Secure ACS 5.x
PT-2018-29: Межсайтовое выполнение хранимых сценариев	5.4	
PT-2018-28: Выполнение произвольных команд	9.8	
PT-2018-26: Раскрытие информации	6.1	MatrikonOPC Explorer 5.x
PT-2018-21: Переполнение буфера	7.5	Schneider Electric Modicon Quantum Schneider Electric Modicon Premium Schneider Electric Modicon M340
PT-2018-20: Выполнение произвольных команд	9.8	Schneider Electric Modicon BMXNOR0200
PT-2018-19: Обход авторизации	9.8	Schneider Electric Modicon Quantum Schneider Electric Modicon Premium Schneider Electric Modicon M340 Schneider Electric Modicon BMXNOR0200
PT-2018-17: Раскрытие информации	9.8	
PT-2018-16: Жестко закодированные учетные данные	9.8	
PT-2018-15: Выполнение произвольных команд	8.8	Schneider Electric Modicon Quantum
PT-2018-14: Переполнение буфера	9.0	PHOENIX CONTACT FL SWITCH 3xxx, FL SWITCH 4xxx, FL SWITCH 48xxx
PT-2018-13: Внедрение команд	9.1	
PT-2018-02: Ошибка авторизации	9.8	
PT-2018-10: Подмена запроса на стороне сервера	9.8	Ipswitch WhatsUp Gold 17.x
PT-2018-09: Внедрение команд	9.8	
PT-2018-08: Внедрение SQL-кода	9.8	
PT-2018-07: Выполнение произвольных команд	9.8	
PT-2018-06: Обход аутентификации	9.8	Hirschmann RSR, RS, RSB, MACH100, MACH1000, MACH4000, OCTOPUS, MS
PT-2018-05: Несанкционированное изменение прошивки	7.5	Siemens EN100
PT-2018-04: Раскрытие информации	9.0	Siemens SIPROTEC 4, SIPROTEC Compact
PT-2018-03: Перехват управления	10	Siemens DIGSI 4, EN100

В работе предлагается трехуровневая модель повышения точности аутентификации пользователей по динамике нажатия клавиш.

Перед началом использования промышленной автоматизированной системы управления 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 подчинено нормальному гауссовскому закону (стандартное отклонение) и проверке соответствия на критерий Хи-квадрат (χ^2). На втором уровне решается классовая задача 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 рисунке 1 представлена структурная схема детализации трехуровневой модели аутентификации пользователей.

Опишем более подробно каждый уровень предложенной модели.

Статистический уровень. На этом уровне измеряются временные параметры клавиатурного почерка пользователей и формируются

эталоны пользователей в виде средних значений и среднеквадратических отклонений. Далее происходит анализ стандартных отклонений между текущими входными данными и эталонными значениями. Стандартное отклонение временных параметров динамики работы пользователей на клавиатуре вычисляется по формуле

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

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

Таблица 2

Распределение уязвимостей в автоматизированных системах управления по типу

Table 2

Vulnerability distribution in automated systems management by type

Тип уязвимости	Доля уязвимости, %
Избыточные права и привилегии, недостаточный контроль доступа	11
Неправильная проверка ввода	9
Уязвимости при работе с памятью	8
Выход за пределы назначенного каталога	7
Раскрытие информации	6
Внедрение команд	6
Некорректный контроль доступа	5
Внедрение операторов SQL	5
Межсайтовое выполнение сценариев	4
Некорректный механизм аутентификации	4
Другие	35

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

Следующий уровень – выполнение теста χ^2 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 результата: пользователь легальный или нелегальный. Степень свободы используется для выбора критического значения в статистической таблице [18]. 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. Вычисления проводятся по формуле

$$\chi^2 = \sum \frac{(\overline{x}_j - L_i)^2}{L_i}, \quad (3)$$

где χ^2 – значение критерия Хи-квадрат; $\overline{x}_j$ – наблюдаемое значение (среднее значение временного параметра клавиатурного почерка); L_i – ожидаемое значение (среднее значение, сохраненное в эталоне пользователя) [18].

Значение χ^2 должно быть меньше или равно критическому значению, выбранному из статистической таблицы [18], 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 – 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, на примере работы [19], осуществляем определение параметров классов путем разбиения характеристик почерка на классы L_1, …, L_m. Требуется выделить в пространстве наборных характеристик области S_i, $i = 1, \ldots, m$, эквивалентные классам, то есть, если характеристика клавиатурного почерка, имеющая параметры

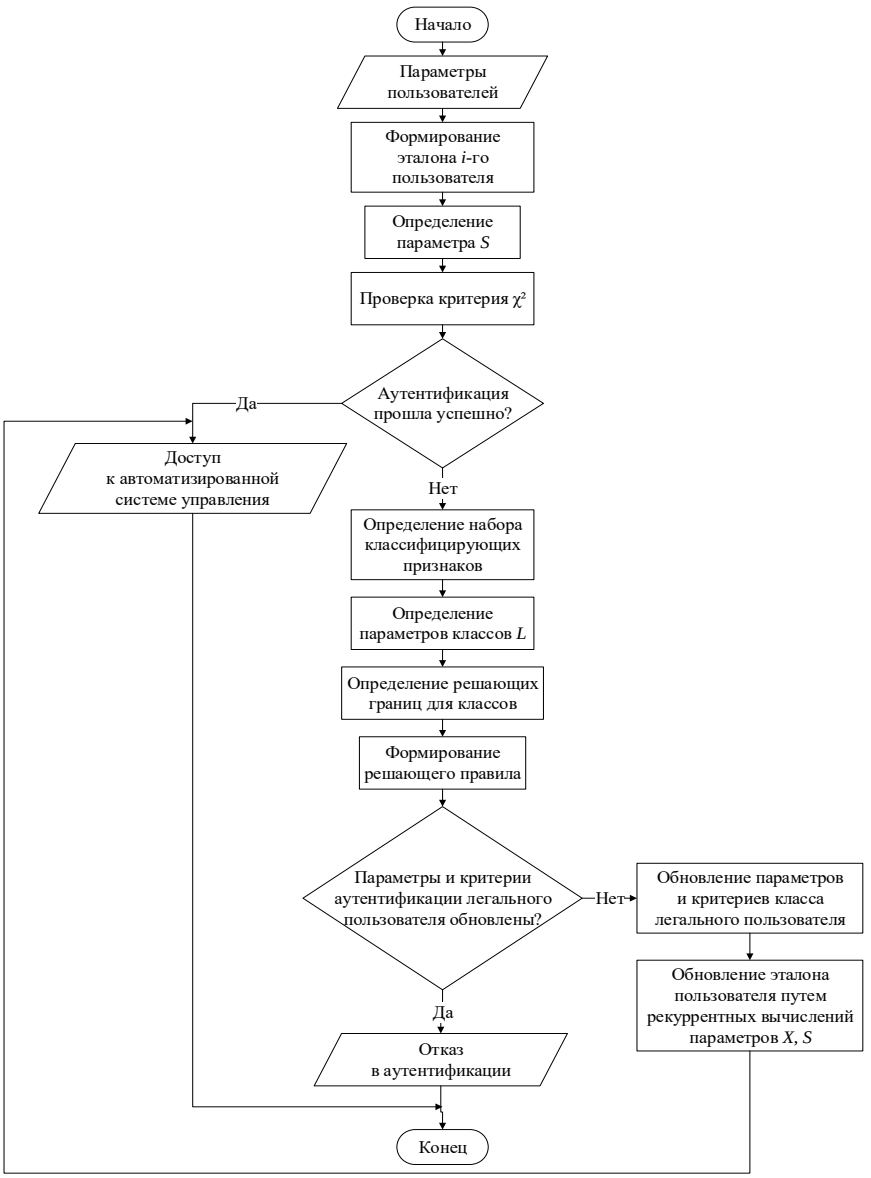

Рис. 1. Структурная схема детализации трехуровневой модели аутентификации пользователей

Fig. 1. Block diagram for detailing a three-level user authentication model

x^0_1, …, x^0_N, относится к классу L_i, то представляющее его в пространстве наборных характеристик значение принадлежит области S_i. Пример разбиения классов для представленной задачи показан на рисунке 2.

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

На рисунке 2 показано разбиение двумерного пространства наборных характеристик клавиатурного почерка на области S_1, S_2, соответствующие 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, 1 – решающая граница.

Задав разделяющие функции, можно проверить выполнение условий.

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

Если пользователь распознался как «свой», происходит обновление параметров $\overline{x}_j$ и S, сохраненных как эталонные значения, в соответ-

ствии с рекуррентными отношениями вида:

$$\bar{x}_j = \frac{1}{N}\sum_{i=1}^{N} x_i \,, \quad (4)$$

где $\bar{x}_j$ – среднее значение параметров пользователя; N – количество наблюдений; x_i – текущее значение параметров пользователя;

$$S = \frac{1}{N-1}\sum_{i=1}^{N}\left|x_i - \bar{x}_j\right| , \quad (5)$$

где S – среднеквадратичное отклонение (обновленное).

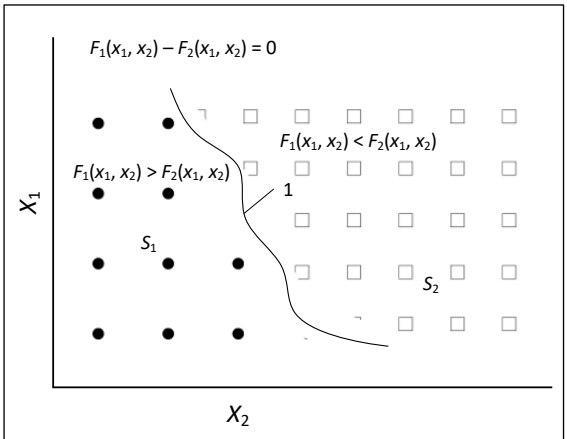

Рис. 2. Разбиение классов пользователей

Fig. 2 Users classes differentiation

Экспериментальная часть

Для проверки работоспособности трехуровневой модели проведены экспериментальные исследования и численно-параметрические расчеты. Для эксперимента была выбрана группа из шести пользователей промышленной автоматизированной системы управления, разработанной корпорацией «Галактика», владеющих десятипальцевым слепым набором слов на клавиатуре. Сначала было предложено войти на специально разработанный web-сайт, авторизоваться, в процессе чего формировались эталоны данных пользователей по результатам их работы на клавиатуре, при этом пользователи должны были указать свои Ф.И.О., направление профессиональной деятельности, должность. Затем они перенаправлялись на другую страницу, где ими осуществлялся ввод предложенных готовых текстов на экране. Эти тексты представляли собой некие заранее подготовленные предложения, неизвестные тестируемой группе.

Далее осуществлялся сбор введенных тестируемой группой текущих данных на стороне клиента из web-формы. Когда пользователь печатал 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.

Код обработчика событий Keyup и Keydown выглядит следующим образом:

```
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.pre-
ventDefault();
}
```

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, имя, время задержки клавиши и временное значение отпускания клавиши на сервер, используя при этом AJAX-запрос для дальнейшей обработки этих данных.

На следующем этапе осуществлялся сбор данных на стороне сервера. Web-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 пользователя на стороне сервера скрипт запрашивал БД для проверки, реализованной на MySQL, с целью определения, не были ли какие-либо предыдущие данные (нажатые клавиши) уже сохранены для текущего пользователя. Если данные уже были в базе, обработчик выбирал последние три и

использовал их временные метки для расчета временных интервалов диад, триад и тетрад путем разности значений временных меток.

В процессе эксперимента статистические данные были разделены на данные для обработки и эталонные.

Был выбран параметр «время удержания клавиш пользователями», так как он наиболее информативный в рассматриваемом случае, хотя это утверждение может быть неверно при работе с большим объемом данных.

По результатам тестов для каждого пользователя были рассчитаны и выстроены стандартные девиации текущих и эталонных данных. Результаты для всех шести тестируемых пользователей представлены в таблице 3. Кроме этого, был рассчитан критерий χ^2 для каждого тестируемого пользователя. Из таблицы видно, что пользователь под номером 2 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.

Таблица 3

Значение девиации и критерий хи-квадрат

Table 3

Deviation value and Chi-square criterion

Пользователь	Данные		
	Значения девиации (эталонные данные)	Значения девиации (измеренные)	Критерий Хи-квадрат
1	42	12	4.82
2	39	38	13.86
3	63	9	5.95
4	62	48	51.41
5	162	35	11.84
6	169	41	31.42

Пользователи № 1, 3 и 5 были аутентифицированы с использованием критерия χ^2, поскольку их значения критерия меньше, чем критическое значение 12.71, тогда с уверенностью в 97,5 % можно утверждать, что пользователи идентифицировались верно. Результаты расчетов показывают, что 50 % пользователей прошли проверку подлинности в процессе сравнения статистических значений параметров динамики нажатия клавиш. Но нужно 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 было убедиться в способности предложенной модели разграничить доступ для легитимных пользователей и нелегитимных, приспосабливаясь к незначительным изменениям параметров динамики нажатия клавиш, тем самым увеличивая точность аутентификации. В результате один из тестируемых пользователей был подвержен сравнению девиаций, в то время как другие три аутентифицированы с помощью критерия χ^2. Остальные пользователи не прошли процедуру 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

В работе представлен обзор решений в области аутентификации пользователей, предложена модель повышения точности аутентификации пользователей по динамике нажатия клавиш, приведены результаты экспериментального исследования.

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

Проблемы с эффективностью можно 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 исследованиях, где три тестируемых пользователя прошли процесс аутентификации после оценки критерия χ^2. В дальнейшем планируется создать метод, работающий на уровне логического сравнения с целью повышения точности. Он будет предназначен для создания усовершенствованного эталона, в котором учитывались бы особенности каждого пользователя, а также некоторые девиации, связанные с эмоциональными состояниями пользователей, что могло бы способствовать разработке систем непрерывной аутентификации пользователей с целью мониторинга их эмоционального состояния при работе на критических объектах.

Литература

1. Тесленко П.А., Барская И.С., Чумаченко Е.А. Проект создания автоматизированной системы управления контейнерным терминалом // Современные информационные и электронные технологии. 2013. Т. 1. № 14. С. 74–77.
2. Алферов В.П., Дровникова И.Г., Обухова Л.А., Рогозин Е.А. Вербальная модель управления слабо уязвимым процессом разграничения доступа пользователей к программным средствам системы электронного документооборота // Вестн. ДГТУ. 2019. Т. 46. № 2. С. 37–49.
3. Костюченко Е.Ю., Мещеряков Р.В. Распознавания пользователя по клавиатурному почерку на фиксированной парольной фразе в компьютерных системах // Изв. ЮФУ. 2003. № 4. С. 177–178.
4. Тумбинская М.В., Баянов Б.И., Рахимов Р.Ж., Кормильцев Н.В., Уваров А.Д. Анализ и прогнозирование вредоносного сетевого трафика в облачных сервисах // Бизнес-информатика. 2019. № 1. С. 71–81. DOI: 10.17323/1998-0663.2019.1.71.81.
5. Sharipov R.R., Tumbinskaya M.V., Abzalov A.R. Analysis of users keyboard handwriting based on Gaussian reference signals. Proc. Intern. RusAutoConf., 2019. DOI: 10.1109/RUSAUTOCON.2019.8867753.
6. Шарипов Р.Р., Сафиуллин Н.З. Аппаратурный анализ клавиатурного почерка с использованием эталонных гауссовских сигналов // Вестн. КГТУ им. А.Н. Туполева. 2006. № 2. С. 21–23.
7. D'Lima N., Mittal J. Password authentication using keystroke biometrics. Proc. IEEE Intern. Conf. ICCICT, Mumbai, 2015, pp. 1–6.
8. Промышленные компании: векторы атак. URL: https://www.ptsecurity.com/ru-ru/research/analytics/ics-attacks-2018/ (дата обращения: 15.11.2019).
9. Анализ уязвимостей компонентов АСУ ТП. URL: https://www.ptsecurity.com/ru-ru/research/analytics/ics-vulnerabilities-2019/ (дата обращения: 15.11.2019).
10. Анализ уязвимостей компонентов АСУ ТП. URL: https://www.ptsecurity.com/ru-ru/research/analytics/ics-vulnerabilities-2019/#3 (дата обращения: 15.11.2019).
11. Басыня Е.А. Распределенная система сбора, обработки и анализа событий информационной безопасности сетевой инфраструктуры предприятия // Безопасность информационных технологий. 2018. Т. 25. № 4. С. 42–51.
12. Абзалов А.Р., Самигуллина Р.Р., Жиганов А.В. Аутентификация пользователей по динамике нажатий клавиш при использовании систем автоматического прокторинга // Прикладная информатика. 2019. Т. 14. № 6. С. 25–35.
13. Flior E., Kowalski K. Continuous biometric user authentication in online examinations. Proc. IEEE 7th Intern. Conf. Inform. Tech., 2010, pp. 488–492.
14. Peralta D., Triguero I., Garcia S., Herrera F., Benitez J.M. DPD-DFF: A dual phase distributed scheme with double fingerprint fusion for fast and accurate identification in large databases. Information Fusion, 2016, vol. 32, pp. 40–51. DOI: 10.1016/j.inffus.2016.03.002.
15. Nelasa A.V., Krischuk V.M. Using of the user identification methods on keyboard handwriting at digital signature shaping. Proc. 6th Intern. Conf., CADSM, 2001, pp. 239–240. DOI: 10.1109/CADSM.2001.975824.
16. 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.
17. Самойлова Е.М. Построение экспертной системы поддержки принятия решения как интеллектуальной составляющей системы мониторинга технологического процесса // Вестн. ПНИПУ. 2016. Т. 18. № 2. С. 128–142.
18. Берман Г.Н. Сборник задач по курсу математического анализа. М.: Наука, 1977. 416 с.
19. Delattre M., Genon-Catalot V., Samson A. Mixtures of stochastic differential equations with random effects: application to data clustering. J. Stat. Plan. Inference, 2016, vol. 173, pp. 109–124.

Software & Systems Received 03.12.19
DOI: 10.15827/0236-235X.130.266-275 2020, vol. 33, no. 2, pp. 266–275

User authentication based on the keystroke dynamics in the process of using industrial control systems

M.V. Tumbinskaya [1], *Ph.D. (Engineering), Associate Professor, Department of Information Security Systems, tumbinskaya@inbox.ru*
N.F. Asadullin [1], *Student, nail.asadullin.1997@mail.ru*
R.R. Murtazin [1], *Student, ramichhh@gmail.com*

[1] *National Research Technical University, Kazan, 420111, Russian Federation*

Abstract. There are modern industrial control systems ubiquitously. The complex architecture of such systems, the requirements of process continuity and access to the Internet make them easily vulnerable to cybercriminals and cyber-attacks. Currently, there is no full protection for the components of industrial automated control systems, so there is a need for adequate protection and increasing the information security level.

As the domain knowledge analysis shows, 82 % of industrial enterprises cannot resist the internal intruder, who seeks to penetrate the technological network from the corporate network, after gaining access to the network technological segment, the attacker has wide opportunities for malicious influence on the components of industrial control systems. In 2018, the number of new vulnerabilities in the industrial control system components increased by 30 % compared to 2017. A significant proportion of vulnerabilities in industrial control systems have an association with incorrect authentication or excessive user access rights. Moreover, more than half of the vulnerabilities can be exploited remotely.

The paper presents an overview of information security vulnerabilities of industrial automated systems. Based on the analysis of the subject area, the authors propose a three-level model for improving the accuracy of user authentication based on the keystroke dynamics, which will increase the security level of automated control systems. The authors conducted an experimental study, the results of which showed a high ability of the proposed model for access differentiation for legal users and hackers, taking into account minor changes in the keystroke dynamic parameters, to increase the user authentication accuracy. The user authentication authenticity in practice was 97.5 %.

Keywords: authentication accuracy, typing biometrics, keystroke dynamics, biometrics, authentication, clustering.

References

1. Teslenko P.A., Barskaya I.S., Chumachenko E.A. The project of creating an automated container terminal management system. *Modern Information and Electronic Technologies*, 2013, vol. 1, no. 14, pp. 74–77 (in Russ.).

2. Alferov V.P., Drovnikova I.G., Obukhova L.A., Rogozin E.A. The verbal model of managing a weakly vulnerable process of delimiting user access to software tools of an electronic document management system. *Bull. DSTU*, 2019, vol. 46, no. 2, pp. 37–49 (in Russ.).

3. Kostyuchenko E.Yu., Meshcheryakov R.V. User recognition by keyboard handwriting on a fixed passphrase in computer systems. *News SFU*, 2003, no. 4, pp. 177–178 (in Russ.).

4. Tumbinskaya M.V., Bayanov B.I., Rakhimov R.Zh., Kormiltcev N.V., Uvarov A.D. Analysis and forecast of undesirable cloud services traffic. *Business Informatics*, 2019, no. 1, pp. 71–81 (in Russ.). DOI: 10.17323/1998-0663.2019.1.71.81.

5. Sharipov R.R., Tumbinskaya M.V., Abzalov A.R. Analysis of users keyboard handwriting based on Gaussian reference signals. *Proc. Intern. RusAutoConf.*, 2019. DOI: 10.1109/RUSAUTOCON.2019.8867753.

6. Sharipov R.R., Safiullin N.Z. Hardware analysis of keyboard script by using the reference gaussian signals. *Bull. KSTU*, 2006, no. 2, pp. 21–23 (in Russ.).

7. D'Lima N., Mittal J. Password authentication using keystroke biometrics. *Proc. IEEE Intern. Conf. ICCICT*, Mumbai, 2015, pp. 1–6.

8. *Industrial Companies: Attack Vectors*. Available at: https://www.ptsecurity.com/ru-ru/research/analytics/ics-attacks-2018/ (accessed November 15, 2019).

9. *Vulnerability Analysis of ICS Components*. Available at: https://www.ptsecurity.com/ru-ru/research/analytics/ics-vulnerabilities-2019/ (accessed November 30, 2019).

10. *Vulnerability Analysis of ICS Components*. Available at: https://www.ptsecurity.com/ru-ru/research/analytics/ics-vulnerabilities-2019/#3 (accessed November 15, 2019).

11. Basynya E.A. Distributed system for the collection, processing and analysis of information security events network infrastructure of the enterprise. *Information Technology Security*, 2018, vol. 25, no. 4, pp. 42–51 (in Russ.).

12. Abzalov A.R., Samigullina R.R., Zhiganov A.V. Authentication of users by the dynamics of keystrokes when using automatic proctoring systems. *Applied Informatics*, 2019, vol. 14, no. 6, pp. 25–35 (in Russ.).

13. Flior E., Kowalski K. Continuous biometric user authentication in online examinations. *Proc. IEEE 7th Intern. Conf. Inform. Tech*., 2010, pp. 488–492.

14. Peralta D., Triguero I., Garcia S., Herrera F., Benitez J.M. DPD-DFF: A dual phase distributed scheme with double fingerprint fusion for fast and accurate identification in large databases. *Information Fusion*, 2016, vol. 32, pp. 40–51. DOI: 10.1016/j.inffus.2016.03.002.

15. Nelasa A.V., Krischuk V.M. Using of the user identification methods on keyboard handwriting at digital signature shaping. *Proc. 6th Intern. Conf., CADSM*, 2001, pp. 239–240. DOI: 10.1109/CADSM.2001.975824.

16. Mukhamatkhanov R.M., Mikhaylov A.A., Bayanov B.I., Tumbinskaya M.V. Classification of DDos attacks based on a neural network model. *Applied Informatics*, 2019, no. 1, pp. 96–103 (in Russ.).

17. Samoylova E.M. Construction of an expert decision support system as an intellectual component of a process monitoring system. *Bull. PNRPU*, 2016, vol. 18, no. 2, pp. 128–142 (in Russ.).

18. Berman G.N. *Collection of Tasks on the Course of Mathematical Analysis*. Moscow, 1977, 416 p. (in Russ.).

19. Delattre M., Genon-Catalot V., Samson A. Mixtures of stochastic differential equations with random effects: application to data clustering. *J. Stat. Plan. Inference*, 2016, vol. 173, pp. 109–124.

Для цитирования

Тумбинская М.В., Асадуллин Н.Ф., Муртазин Р.Р. Моделирование аутентификации пользователей по динамике нажатий клавиш в промышленных автоматизированных системах // Программные продукты и системы. 2020. Т. 33. № 2. С. 266–275. DOI: 10.15827/0236-235X. 130.266-275.

For citation

Tumbinskaya M.V., Asadullin N.F., Murtazin R.R. User authentication based on the keystroke dynamics in the process of using industrial control systems. *Software & Systems*, 2020, vol. 33, no. 2, pp. 266–275 (in Russ.). DOI: 10.15827/0236-235X.130.266-275.