Результат вычисления S_2 используется для вынесения вердикта об отнесении программы соответствующей процессу *p* к вредоносным. Решение принимается по превышению некоторого заранее заданного порога принятия решения T, простым сравнением с заданной величиной. Величина T, при которой программа может быть признана вредоносной, подбирается экспериментально и сильно зависит от исходной таблицы экспертных оценок.

Выводы. Предложенный метод проактивной защиты опробован экспериментально. Модуль принятия решения, использованный в ходе экспериментов, показал достаточно хорошие результаты. Как и ожидалось, удалось значительно снизить количество ложных срабатываний. В тоже время, зависимость величины порога принятия решения от таблиц экспертных оценок можно отнести к недостаткам, для устранения которых требуется проведение дальнейших исследований. Вместе с тем, удалось показать, что система проактивной защиты, построенная на описанных в статье принципах, в подавляющем большинстве случаев способна самостоятельно без обращения к пользователю, принимать верные решения. Результаты работы уже на настоящем этапе исследований могут быть использованы для реализации методов проактивной защиты компьютерных систем от вредоносных программ.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. РД ФСТЭК «Базовая модель угроз безопасности персональных данных их обработке в информационных системах персональных данных», 15 февраля 2008 г.
2. *Касперский К.* Записки исследователя компьютерных вирусов. – СПб.: Питер, 2002. – 316 с.
3. *Aycock J.* Computer Viruses and Malware. Advances in information security. – Calgary: Springer, 2006. – 227 p.
4. *Алиев А.Т., Морозов А.П.* Защита информационных систем от вредоносного программного обеспечения // XIX науч. конф. “Современные информационные технологии: тенденции и перспективы развития”. – Ростов-на-Дону, 2012. – С. 26-27.
5. *Шрайбер С.* Недокументированные возможности Windows 2000. Библиотека программиста. – СПб.: Питер, 2002. – 544 с.
6. *Алиев А.Т.* Построение проактивной системы защиты от вредоносных программ // XII науч.-практ. конф. «Информационная безопасность – 2012». – Таганрог: ТТИ ЮФУ, 2012. Ч. 2. – С. 45-50.

Статью рекомендовал к опубликованию д.т.н., профессор О.Б. Макаревич.

Алиев Александр Тофикович – ООО НПО «Редут»; e-mail: aliev@nporedut.ru; 344002, г. Ростов-на-Дону, ул. Темерницкая, 44; генеральный директор; к.т.н.

Aliev Alexander Tofikovich – RPA “Redut”, Llc; e-mail: aliev@nporedut.ru; 44, Temernitskaya street, Rostov-on-Don, 344002, Russia; general director; cand. of eng. sc.

УДК 681.324

Ю.А. Брюхомицкий

ИММУНОЛОГИЧЕСКИЙ ПОДХОД К ОРГАНИЗАЦИИ КЛАВИАТУРНОГО МОНИТОРИНГА*

Клавиатурный мониторинг (КМ) позволяет вести непрерывную скрытную аутентификацию пользователей компьютерных систем и решать ряд других задач компьютерной безопасности. Известные системы КМ обладают недостаточной точностью и скоростью верификации работающего пользователя. С целью повышения точности и скорости

* Работа выполнена при поддержке гранта РФФИ 12-07-00081-а.

верификации предлагается анализировать параметры не одиночных событий клавиатуры, а цепочек лингвистически взаимосвязанных событий. Показано, что такой принцип анализа клавиатурных параметров по своей сути соответствует принципу работы искусственной иммунной системы при обнаружении аномалий данных. Предлагаемый иммунологический подход к организации КМ, основан на принципе отрицательного отбора. Суть подхода состоит в том, что при обучении системы КМ создается строковый шаблон «своего» пользователя и на его основе формируются детекторы для обнаружения «чужих». Причем формирование детекторов осуществляется на априори определенном множестве допустимых сочетаний клавиатурных параметров. Приводится алгоритм формирования детекторов. В рабочем режиме система КМ обнаруживает «чужих» по частоте срабатывания детекторов.

Клавиатурный мониторинг; скрытная непрерывная аутентификация; события клавиатуры; клавиатурный шаблон; иммунологический подход; искусственная иммунная система; принцип отрицательного отбора; детекторы для обнаружения «чужих».

Yu.A. Bryukhomitsky

THE IMMUNOLOGIC APPROACH TO KEYBOARD MONITORING ORGANIZATION

The Keyboard Monitoring (KM) makes it possible to carry on the nonstop secretive authentication of computer systems users and to solve other tasks of computer security. Known KM systems have scarce accuracy and verification rate of operating user. For the purpose of accuracy and verification rate increasing it's suggested to analyze the parameters not of keyboard single event, but of chains of linguistic interrelated events. It demonstrated that such principle of keyboard parameters analyzing inherently correspond to the principle of work of artificial immune system while the data anomalies detection. Suggested here immunological approach to a KM organization is based on a principal of negative selection. The main point of suggested approach is to create a string template of "well-known user" (while KM-system training), and then, based on it, the detectors of "strangers" recognition are created. Besides, the detectors creation is based on a priori definite set of acceptable combinations of keyboard parameters. The algorithm of detectors creation offered. While operating, the KM-system detects "strangers" by the response frequency of detectors.

Keyboard Monitoring; nonstop secretive authentication; keyboard events; keyboard template; immunological approach; artificial immune system; principal of negative selection; detectors of "strangers" recognition.

Клавиатурный мониторинг (КМ) позволяет вести непрерывный текст независимый анализ клавиатурного почерка (КП) работающих пользователей компьютерных систем. При необходимости КМ проводится прозрачно для пользователей. В таких случаях говорят о скрытном клавиатурном мониторинге (СКМ).

СКМ позволяет, в частности, решать следующие задачи [1, 2]:

- вести непрерывную аутентификацию пользователей компьютерных систем;
- выявлять пользователей компьютерных систем, совершающих злоупотребления и действия, выходящие за рамки их полномочий;
- выявлять психофизические и психофизиологические отклонения пользователей от их нормального состояния;
- проводить оценку достоверности сообщений («детектор лжи»).

Не умаляя степени общности предлагаемого здесь подхода к построению систем КМ для решения всех перечисленных задач, далее рассматривается решение первой из них – задачи ведения непрерывной аутентификации пользователей компьютерных систем – с целью их классификации по принципу «свой – чужой». Это позволяет целенаправленно использовать понятия, особенности и терминологию, специфичные для такого рода систем.

Задача классификации пользователей по принципу «свой – чужой» в биометрических системах может решаться двояко: как верификация или как идентификация. В первом случае система использует единственный шаблон «своего» пользователя, и сравнение параметров неизвестного пользователя осуществляется только с ним по принципу 1:1. Во втором случае решается общая задача сопоставления параметров неизвестного пользователя со всеми шаблонами зарегистрированных в системе пользователей по принципу 1:*M*. Далее рассматривается решение первой задачи, наиболее актуальной для персональных компьютеров.

Качество систем КМ определяется, прежде всего, точностью и скоростью верификации работающего пользователя. Эти характеристики, в свою очередь, тесно связаны со способами представления и классификации клавиатурных параметров.

Задача представления клавиатурных параметров, заключается в приведении событий клавиатуры к некоторому структурированному виду, который позволяет выявить характерные признаки КП данного пользователя, отличающие его от других пользователей. В режиме обучения системы КМ по этим признакам строятся клавиатурные эталоны зарегистрированных в системе пользователей. В рабочем режиме, путем сравнения характеристик КП фактически работающего пользователя с клавиатурным эталоном зарегистрированного пользователя, устанавливается степень их соответствия.

Наиболее простым и распространенным способом представления клавиатурных биометрических параметров пользователя является прямое измерение временных характеристик клавиатурного ввода [3]. При этом, как правило, контролируются три типа временных параметров событий клавиатуры: длительности удержаний клавиш, длительности пауз между удержаниями очередных клавиш, длительности одновременного удержания очередных клавиш. Последний параметр удобно интерпретировать как отрицательное значение длительности паузы между удержаниями очередных клавиш. В этом случае достаточно оперировать только двумя типами временных параметров событий клавиатуры при анализе КП пользователя.

В простейших методах анализа КП [3] постулируется, что для конкретного пользователя события клавиатуры являются независимыми, а их распределение носит нормальный характер с единственным центром. При таком подходе в качестве индивидуальных характеристик КП пользователя используются усредненные значения временных параметров одиночных событий клавиатуры. Использование таких методов в задачах КМ приводит к большой погрешности аутентификации, обусловленной, в первую очередь, недостаточной информативностью принятого представления клавиатурных параметров. Причем уровень ошибок существенно не снижается даже при больших объемах статистики. Это свидетельствует о наличии методической ошибки, которая не устраняется в рамках этих методов.

Последующие исследования показали, что для определенного пользователя статистические оценки временных параметров одних и тех же событий клавиатуры, наступающих в различных сочетаниях, заметно отличаются. Это свидетельствует о наличии в КП каждого пользователя своих устойчивых корреляционных зависимостей между временными параметрами последовательно наступающих событий клавиатуры. На основании этих результатов были предложены методы многосвязного [4] и цепочного [5] представления клавиатурных параметров, направленные на повышение точности систем КМ.

В процессе КМ в динамике регистрируются, по существу, два вида последовательно наступающих событий клавиатуры: удержание одной из *n* клавиш и пауза между удержаниями очередных клавиш. Используя терминологию формальных грамматик всю совокупность событий клавиатуры можно представить как алфавит

$A = A_{\text{у}} \cup A_{\bar{\text{у}}}$, объединяющий $A_{\text{у}}$ – подмножество событий, состоящих в удержании клавиш и $A_{\bar{\text{у}}}$ – подмножество событий, состоящих в наличии пауз между удержаниями клавиш.

Ограниченные последовательности событий, ориентированные слева направо, начинающиеся и оканчивающиеся событиями из подмножества $A_{\text{у}}$, представляют собой *цепочки событий* алфавита A. Длина r цепочки равна общему числу r событий алфавита A, входящих в эту цепочку. При клавиатурном наборе события подмножеств $A_{\text{у}}$ и $A_{\bar{\text{у}}}$ строго чередуются, поэтому каждая цепочка длины r будет содержать $(r + 1) / 2$ событий подмножества $A_{\text{у}}$ и $(r + 1) / 2 - 1$ событий подмножества $A_{\bar{\text{у}}}$. Частный случай такого представления при $r = 1$ соответствует простейшему методу [3], при $r \geq 3$ имеет место цепочный метод [5].

Экспериментальные исследования [6] подтвердили увеличение точности клавиатурной аутентификации пользователя с переходом от методов анализа усредненных одиночных событий клавиатуры к многосвязному и цепочному методам анализа КП. В то же время, непосредственная практическая реализация указанных методов достаточно сложна, что стимулирует поиск других подходов к построению хороших систем КМ.

Цепочный метод [5] основан, по существу, на учете особенностей вхождения символов в контекст, когда идентифицирующие пользователя клавиатурные параметры определяются не одиночными событиями клавиатуры, а цепочками лингвистически связанных событий. Это обстоятельство хорошо согласуется с иммунологическим принципом анализа аномалий в информационных потоках. В искусственной иммунной системе (ИИС) информационный поток анализируется на наличие аномалий путем сопоставления последовательно идущих событий с детекторами. Причем сопоставление производится не в одиночных, а сразу в нескольких смежных позициях. Ширина зоны сопоставления задается параметром r, который определяет, по существу, степень связности последовательно идущих событий и называется в иммунологии степенью аффинности. При сопоставлении цепочного метода КМ и иммунологического методов анализа становится очевидным, что длина цепочки лингвистически связанных событий в КМ содержательно адекватна параметру аффинности r в ИИС. Это обстоятельство оправдывает попытку использовать иммунологический подход к построению систем КМ, с целью получения удовлетворительного сочетания точности и простоты реализации.

Для решения задач КМ более подходящей, представляется иммунологическая модель, основанная на *отрицательном отборе.* Как и в любых системах распознавания типа «свой-чужой», в ней используется понятие шаблона «своего». Принципиальным отличием модели отрицательного отбора от обычных схем распознавания, является то, что шаблон «своего» формируется и используется лишь на этапе обучения ИИС для создания альтернативных ему детекторов образов «чужих». На этапе распознавания неизвестные образы сравниваются не с шаблоном «своего», а с детекторами «чужих». Такая схема распознавания имеет ряд существенных преимуществ [7].

Рассмотрим этап обучения системы КМ с использованием иммунологического подхода.

Пусть клавиатурная работа «своего» пользователя представлена произвольной последовательностью $P(t_i) = p_1, p_2, \ldots$ событий клавиатуры. Исключим из этой последовательности длительные паузы, не обусловленные КП пользователя, и условно ограничим ее числом событий N. Результатом будет цельная, лингвистически обусловленная конечная последовательность N событий клавиатуры $\hat{P}(t_i) = p_1, p_2, \ldots, p_N$ алфавита $A = A_{\text{у}} \cup A_{\bar{\text{у}}}$, отражающая КП «своего» пользователя [8].

Каждое событие алфавита A характеризуется временными параметрами: события подмножества $A_{\text{у}}$ – параметром τ_i – длительности удержания клавиши i; события подмножества $A_{\bar{\text{у}}}$ – параметром τ_{ij} – алгебраического значения длительности паузы между удержаниями клавиш i и j.

Статистическую оценку диапазонов вариации временных параметров τ_i, τ_{ij} событий клавиатуры удобно представлять минимаксными значениями:

$$\tau_i \to \left(\min_l \tau_i, \max_l \tau_i\right);\ \tau_{ij} \to \left(\min_l \tau_{ij}, \max_l \tau_{ij}\right), i, j = 1, 2, \dots, n,$$

где l – число измерений соответствующего параметра.

Если минимаксные диапазоны вариации временных параметров τ_i и τ_{ij} сильно отличаются, то их следует учитывать с разными масштабами. Для соразмерных минимаксных диапазонов τ_i и τ_{ij}целесообразно принять

$$\min_l \tau_i = \min_l \tau_{ij} = \min_l \tau;\ \ \max_l \tau_i = \max_l \tau_{ij} = \max_l \tau$$
$$i, j = 1, 2, \dots, n.$$

Границы вариации всех временных параметров при малых l могут устанавливаться путем прямой фиксации минимаксных значений. При больших l границы вариации параметров целесообразно вычислять с использованием числовых характеристик распределения: математического ожидания $\text{m}(\tau)$ и дисперсии $\sigma(\tau)$. Учитывая, что для оценки τ_i и τ_{ij} используются распределения выборочных статистик, границы вариации лучше задавать на основе *t*-распределения Стьюдента, учитывающего ошибку первого рода P_1:

$$\min_l \tau = \text{m}(\tau) - t[l, (1 - P_1)] \cdot \sigma(\tau);$$
$$\max_l \tau = \text{m}(\tau) + t[l, (1 - P_1)] \cdot \sigma(\tau),$$

где t – коэффициенты Стьюдента.

Для образования клавиатурного шаблона $\mathbf{T}^{\text{c}}$ «своего» пользователя последовательность событий клавиатуры $\hat{P}(t_i) = p_1, p_2, \dots, p_N$ с помощью скользящего временного окна с шагом сдвига h и длиной r разбивается на строки $S_{ij} = s_1, s_2, \dots, s_r,\ \ j = 1, 2, \dots\ n_S$ по $i = 1, 2, \dots, r$ событий в каждой строке. В результате последовательность $\hat{P}(t_i) = p_1, p_2, \dots, p_N$ будет представлена целым числом строк n_S, которое определится выражением

$$n_S = \left\lfloor \frac{N - r}{h} + 1 \right\rfloor,$$

где $\lfloor x \rfloor$ – пол функция числа x, (целая часть числа, образуемая путем округления x до ближайшего целого в меньшую сторону).

Для сохранения существа и преимуществ цепочного метода шаг сдвига скользящего временного окна следует принять $h = r$. В этом случае $n_S = \lfloor N / r \rfloor$.

В ИИС сопоставление строк образов «своего» и «чужого» осуществляется по принципу их частичного соответствия в r смежных позициях. Для реализации этого принципа для временных параметров КП, представленных действительными числами, их необходимо предварительно проквантовать по уровням. В [7] для этого предложено использовать двоично-кодированные целые числа. В общем случае, возможно кодирование с любым основанием d. Далее для простоты используется кодирование с основанием 10.

Установленный выше минимаксный диапазон вариации временных параметров τ КП пользователя нормируем к фиксированному диапазону целых десятичных чисел 0-(10^m–1). В результате каждое значение временного параметра τ в строке шаблона, исходно заданное в виде действительного числа, после квантования будет представлено целым десятичным числом, изменяющимся в диапазоне 0-(10^m–1). Разрядность m задает точность такого представления. Весь диапазон

вариации ($\min_l \tau$, $\max_l \tau$) будет содержать 10^m–1 интервалов с размером интервала d=1. Величина τ, реально изменяющаяся в диапазоне ($\min_l \tau$, $\max_l \tau$), будет отнесена к одному из интервалов диапазона 0-(10^m–1) с абсолютной ошибкой d и представлена десятичным кодом номера интервала. Принцип кодирования временных параметров событий клавиатуры поясняет рис. 1.

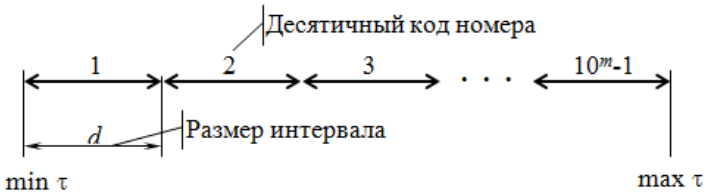

Рис. 1. Принцип кодирования событий информационных процессов

После применения указанной схемы кодирования клавиатурный шаблон пользователя $\mathbf{T}^c$ будет представлен массивом из n_S строк по r целых десятичных чисел в каждой строке. Общее число событий клавиатуры, необходимое для представления параметров КП пользователя приближенно определится как $N \cong n_S \cdot r \cdot l$.

Предлагаемую схему распознавания «чужого» в системе КМ удобно отобразить в виде следующей геометрической интерпретации. Строки массива $S_{ij} = s_1, s_2, \dots, s_r,\ i = 1, 2, \dots, r,\ j = 1, 2, \dots, n_S$ представим r-мерными векторами $\mathbf{S}_j,\ j = 1, 2, \dots, n_S$ с координатами τ_1, τ_{12}, τ_2, …, τ_r. В процессе формирования клавиатурного шаблона пользователя концы векторов $\mathbf{S}_j$, будут группироваться в ограниченных областях пространства, образуя кластеры клавиатурных параметров данного пользователя

$$C_k^c = \bigcup_{j=1}^{n_S} \mathbf{S}_j, k = 1, 2, \dots, n^2.$$

Постулируя нормальный характер распределения КП пользователя, кластеры будут иметь форму r-мерных эллипсоидов, а весь шаблон $\mathbf{T}^c$ «своего» пользователя будет представлен совокупностью n^2 кластеров по количеству неповторяющихся цепочек длины r:

$$\mathbf{T}^c = \bigcup_{k=1}^{n^2} C_k^c.$$

По аналогии строки событий клавиатуры $X_{ij} = x_1, x_2, \dots, x_r, i = 1, 2, \dots, r$, $j = 1, 2, \dots$ порожденных работой неизвестного пользователя, будучи представленные r-мерными векторами клавиатурных параметров $\mathbf{X}_j = x_1, x_2, \dots, x_r$, $j = 1, 2, \dots$ с координатами τ_1, τ_{12}, τ_2, …, τ_r также будут образовывать кластеры $C_k^x, k = 1, 2, \dots, n^2$ в форме r-мерных эллипсоидов, несущие информацию о клавиатурном шаблоне $\mathbf{T}^x$ неизвестного пользователя.

Следующим этапом реализации иммунологического подхода к организации КМ является использование шаблона $\mathbf{T}^c$ для создания детекторов, способных обнаруживать незарегистрированных в системе КМ пользователей.

Самый простой способ создания детекторов, применяемый в АОО, – их случайная генерация с равномерным законом распределения в областях пространства, свободных от кластеров «своего» [9]. Однако, случайная генерация детекторов, приводит к экспоненциальному росту вычислительных затрат при увеличении размера шаблона «своего», что ограничивает применение этого способа для многих приложений.

Известны другие способы создания детекторов, которые характеризуются меньшей трудоемкостью [10, 11]. Здесь предлагается один из подобных способов, ориентированный на задачу КМ, представляющийся оптимальным, как по вычислительным затратам, так и по эффективности обнаружения «чужого».

После применения десятичной схемы целочисленного кодирования временных параметров событий клавиатуры, каждая строка $S_{ij} = s_1, s_2, \dots, s_r,\ i = 1, 2, \dots, r,$ $j = 1, 2, \dots, n_S$ клавиатурного шаблона будет содержать $i = 1, 2, \dots, r$ параметров, представленных целыми m-разрядными десятичными числами. Совокупность n_S строк образует клавиатурный шаблон пользователя $\mathbf{T}^c$, который может быть представлен таблицей рис. 2.

S_j	$i = 1$	$i = 2$	…	$i = r$
S_1	c_{11}	c_{21}	…	c_{r1}
S_2	c_{12}	c_{22}	…	c_{r2}
…	…	…	…	…
S_{n_S-1}	$c_{1(n_S-1)}$	$c_{2(n_S-1)}$	…	$c_{r(n_S-1)}$
S_{n_S}	c_{1n_S}	c_{2n_S}	…	c_{rn_S}

Рис. 2. Клавиатурный шаблон пользователя

Значения временных параметров $i = 1, 2, \dots, r$ в столбцах таблицы выделяются из домена целых десятичных чисел, изменяющихся в диапазоне 0-(10^m–1). При этом каждая строка таблицы представляет собой кортеж, состоящий из r элементов, – по одному из каждого домена.

Клавиатурный шаблон пользователя $\mathbf{T}^c$ является подмножеством $\mathbf{T}^c \subset \mathbf{S}$ множества $\mathbf{S}$ – всех возможных комбинаций элементов кортежей $S_k = s_1, s_2, \dots, s_r,$ $i = 1, 2, \dots, r,\ k = 1, 2, \dots, d.$

Множество $\mathbf{S}$ можно получить, как декартово произведение доменов целых десятичных чисел, изменяющихся в диапазоне 0-(10^m–1):

$$\mathbf{S} = \mathbf{D}_1 \times \mathbf{D}_2 \times \dots \times \mathbf{D}_r,$$

где

$$\mathbf{D}_1 = \{c_{11}, c_{12}, \dots, c_{1d}\};$$
$$\mathbf{D}_2 = \{c_{21}, c_{22}, \dots, c_{2d}\};$$
$$\dots\dots\dots\dots\dots$$
$$\mathbf{D}_r = \{c_{r1}, c_{r2}, \dots, c_{rd}\}.$$

Из проведенных рассуждений вытекает предлагаемый способ создания детекторов, который реализуется по следующему алгоритму:

1. Определяется множество $\mathbf{S}$ всех возможных комбинаций элементов кортежей $S_{ik} = s_1, s_2, \dots, s_r,\ i = 1, 2, \dots, r,\ k = 1, 2, \dots, d$ как декартово произведение доменов: $\mathbf{S} = \mathbf{D}_1 \times \mathbf{D}_2 \times \dots \times \mathbf{D}_r$.

2. Формируется клавиатурный шаблон $\mathbf{T}^c$ зарегистрированного в системе пользователя. Для этого последовательность событий клавиатуры $\hat{P}(t_i) = p_1, p_2, \dots, p_N$ с помощью скользящего временного окна с шагом сдвига h и длиной r разбивается на строки $S_j = s_1, s_2, \dots, s_r, j = 1, 2, \dots n_S$ по $i = 1, 2, \dots, r$ событий в каждой строке.

3. Находится разностное множество $\Delta = \mathbf{S} \setminus \mathbf{T}^c$ путем удаления из множества $\mathbf{S}$ кортежей $S_{ij} = s_1, s_2, \dots, s_r,\ i = 1, 2, \dots, r,\ j = 1, 2, \dots, n_S,$ составляющих клавиатурный шаблон пользователя $\mathbf{T}^c$.

4. Подмножество Δ, представляющее собой массив строк $S_{i\delta} = s_1, s_2, \dots, s_r$, $i = 1, 2, \dots, r,\ \delta = 1, 2, \dots, d - n_S$, используется для создания детекторов.

Детекторы формируются в формате целых десятичных чисел, изменяющихся в диапазоне 0-(10^m–1), двумя возможными способами:

- путем прямого перебора строк $S_{i\delta} = s_1, s_2, \dots, s_r,\ i = 1, 2, \dots, r,$ $\delta = 1, 2, \dots, d - n_S$ из подмножества Δ;
- путем случайной генерации строк $S_{i\delta} = s_1, s_2, \dots, s_r,\ i = 1, 2, \dots, r,$ $\delta = 1, 2, \dots, d - n_S$ из подмножества Δ.

Числовые характеристики рассматриваемого приложения позволяет использовать оба упомянутых способа формирования детекторов. Причем способ создания детекторов исключает их сопоставление со строками шаблона на предмет частичного совпадения, как это делается в традиционном алгоритме отрицательного отбора [9].

После создания детекторов процесс обучения системы КМ заканчивается, и ее можно использовать в рабочем режиме – обнаружения незарегистрированных в системе пользователей – «чужих». При этом реализуется режим «верификации», при котором набор детекторов соответствует единственному клавиатурному шаблону «своего» пользователя. Как и на этапе обучения, из последовательности $P^x(t_i) = p_1, p_2, \dots$ событий клавиатуры, образованной работой неизвестного *x*-пользователя, исключаются длительные паузы, не обусловленные его КП, и формируется цельная, лингвистически обусловленная последовательность $\hat{P}^x(t_i) = p_1, p_2, \dots$. С помощью скользящего временного окна с шагом сдвига *h* и длиной *r* она разбивается на строки $S_{ij}^x = s_1, s_2, \dots, s_r, j = 1, 2, \dots$ по $i = 1, 2, \dots, r$ событий в каждой строке. Образованные строки последовательно сопоставляется с детекторами. Совпадение строки последовательности S_{ij}^x с любым детектором свидетельствует о появлении сочетания клавиатурных параметров, отсутствующего в эталоне $\mathbf{T}^{\text{с}}$ «своего» пользователя. Статистическая вероятность клавиатурного присутствия в системе «чужого» $\hat{P}^{\text{ч}}$ определяется частотой срабатывания детекторов f:

$$\hat{P}^{\text{ч}} \approx f = \frac{n_{\text{с}}^+}{n_{\text{с}}},$$

где $n_{\text{с}}^+ -$ число положительных исходов при сравнении строк; $n_{\text{с}}$ – общее число проведенных сравнений строк.

Принятие решения о наличии подмены «своего» пользователя «чужим» считается правомерным при превышении частоты f некоторого порогового значения $f^{\text{п}}$:

$$S_{ij}^x \equiv \begin{cases} \hat{S}_{ij}^{\text{с}}, \text{если } f < f^{\text{п}}; \\ \hat{S}_{ij}^{\text{ч}}, \text{если } f \geq f^{\text{п}}. \end{cases}$$

Причем, в соответствии с процедурой верификации, «чужим» будет признан любой пользователь, КП которого не соответствует шаблону $\mathbf{T}^{\text{с}}$ «своего».

Выводы. Предлагаемый иммунологический подход позволяет получить системы, с принципиально новыми свойствами:

1. В известных системах КМ сопоставление образцов осуществляется в пространстве Хэмминга, в то время как АОО производит анализ на соответствие в пространстве *r*-смежности.

2. В известных системах КМ эталон каждого «своего», формируется на этапе обучения и хранится в памяти системы. Если в сеансе одновременно работают несколько «своих», то это приводит к пропорциональному увеличению числа эталонов. Для АОО увеличение числа пользователей приводит к противоположному эффекту. Количество детекторов независимо от свойств интегрального набора «своего» при фиксированной вероятности успеха остается практически постоянным.

3. Известные системы КМ весьма чувствительны к вариации биометрических параметров «своих» пользователей. АОО, за счет высоко распределенной массовой обработки данных, проявляет бóльшую устойчивость к шумам.

4. В известных системах КМ распознавание «чужих» производится путем итоговых вычислений по всей совокупности сопоставляемых данных. В АОО обнаружение «чужих» производится в темпе поступления входных данных и отображается частотой активации детекторов.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Брюхомицкий Ю.А., Казарин М.Н.* Система скрытного клавиатурного мониторинга // Известия ТРТУ. – 2006. – № 9 (64). – С. 153-154.
2. *Брюхомицкий Ю.А.* Клавиатурная идентификация личности. Lambert Academic Publishing, Saarbrűcken. – Germany, 2012. – 140 с.
3. *Широчин В.П., Кулик А.В., Марченко В.В.* Динамическая аутентификация на основе анализа клавиатурного почерка. // http: // www.masters.donntu.edu.ua / 2002 / fvti / aslamov / files / bio_autentification.htm.
4. *Брюхомицкий Ю.А., Казарин М.Н.* Методы многосвязного представления клавиатурного почерка // Материалы III Международной конференции «Нелокальные краевые задачи и родственные проблемы математической биологии, информатики и физики. – Нальчик, 5-8 декабря 2006. – С. 68-69.
5. *Брюхомицкий Ю.А.* Цепочный метод клавиатурного мониторинга // Известия ЮФУ. Технические науки. – 2009. – № 11 (100). – С. 128-138.
6. *Казарин М.Н.* Разработка и исследование методов скрытного клавиатурного мониторинга: дис…. канд. техн. наук. – Таганрог, 2006. – С. 114-115.
7. Искусственные иммунные системы и их применение / Под ред. Д. Дасгупты: Пер. с англ. / Под ред. А.А. Романюхи. – М.: Физматлит, 2006. – 344 с.
8. *Брюхомицкий Ю.А.* Иммунологические принципы организации клавиатурного мониторинга пользователей компьютерных систем // Материалы XII Международной научно-практической конференции «Информационная безопасность». Ч. I. – Таганрог: Изд-во ТТИ ЮФУ, 2012. – С. 10-19.
9. *Forrest S., Perelson A.S., Allen L., Cherukuri R.* Self-nonself discrimination in a computer // In: Proc. of Ieee symposium on research in security, Oakland, CA, 16-18 May 1994. – P. 202-212.
10. *D'haeseleer P., Forrest S., Helman P.* An immunological approach to change detection: algorithms, analysis, and implications // In: Proc. of Ieee symposium on research in security, Oakland, CA, May 1996.
11. *Брюхомицкий Ю.А., Гончаров С.Б.* Модификации иммунологического алгоритма отрицательного отбора для систем компьютерной безопасности // Материалы Всероссийской научной конференции «Теоретические и методические проблемы эффективного функционирования радиотехнических систем» («Системотехника 2012»). http://rts.tti.sfedu.ru/. – Таганрог, 2012. – С. 126-136.

Статью рекомендовал к опубликованию к.т.н. М.Ю. Руденко.

Брюхомицкий Юрий Анатольевич – Южный федеральный университет; e-mail: bya@tgn.sfedu.ru; 347928, г. Таганрог, ул. Чехова, 2; тел.: 88634371905; кафедра безопасности информационных технологий; доцент.

Bryukhomitsky Yuriy Anatol'evich – Southern Federal University; e-mail: bya@tgn.sfedu.ru; 2, Chekhov street, Taganrog, 347928, Russia; phone: +78634371905; the department of security in data processing technologies; associate professor.