УДК 004.93'1
doi:10.31799/1684-8853-2020-4-61-77

Научные статьи
Articles

Высоконадежная аутентификация по рукописным паролям на основе гибридных нейронных сетей с обеспечением защиты биометрических эталонов от компрометации

А. Е. Сулавко[a], *канд. техн. наук, доцент, orcid.org/0000-0002-9029-8028, sulavich@mail.ru*
[a]*Омский государственный технический университет, Мира пр., 11, Омск, 644050, РФ*

Введение: *нейросетевые преобразователи «биометрия-код» являются идеологической основой для серии стандартов ГОСТ Р 52633 (не имеющих к настоящему моменту мировых аналогов), которые могут быть востребованы при разработке средств высоконадежной биометрической аутентификации и электронной подписи с биометрической активацией.* ***Цель:*** *разработать модель преобразователя «биометрия-код» для высоконадежной биометрической аутентификации по рукописным паролям с высокой устойчивостью к атакам на извлечение знаний.* ***Результаты:*** *продемонстрирована уязвимость нейросетевых преобразователей «биометрия-код», позволяющая совершать быстрый направленный перебор конкурирующих примеров для компрометации биометрического образа и личного ключа его владельца. Описан метод эффективной защиты от данной атаки. Предложена гибридная модель нейросетевого преобразователя «биометрия-код» (основанная на новом типе гибридных нейронных сетей), не компрометирующего биометрический эталон и ключ (пароль) пользователя и устойчивого к подобным атакам. Экспериментально подтверждена высокая надежность и эффективность предложенной модели в задачах верификации рукописных паролей. Показатели надежности генерации ключа из рукописного пароля составили: FRR = 11,5 %, FAR = 0,0009 % при длине ключа 1024 бит (с учетом предъявления подделок рукописного образа).* ***Практическая значимость:*** *результаты будут востребованы в приложениях информационной безопасности и при реализации электронного документооборота.*

Ключевые слова *— распознавание образов, разностные функционалы Байеса, обработка коррелированных биометрических параметров, защита информации, автоматическая настройка нейронных сетей, плотности вероятности, «широкие» нейронные сети, преобразователи «биометрия-код», рукописный почерк.*

Для цитирования: Сулавко А. Е. Высоконадежная аутентификация по рукописным паролям на основе гибридных нейронных сетей с обеспечением защиты биометрических эталонов от компрометации. *Информационно-управляющие системы*, 2020, № 4, с. 61–77. doi:10.31799/1684-8853-2020-4-61-77

For citation: Sulavko A. E. Highly reliable authentication based on handwritten passwords using hybrid neural networks with protection of biometric templates from being compromised. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 4, pp. 61–77 (In Russian). doi:10.31799/1684-8853-2020-4-61-77

Введение

Наступила эпоха Big Data. Растет количество интеллектуальных технологий и онлайн-сервисов, которые несложно освоить массовому потребителю. При этом обеспечить безопасность виртуального образа пользователя становится все сложней — вместе с количеством личных кабинетов возрастает число паролей, которые нужно хранить. Сегодня пользователь нуждается не только в надежной аутентификации, но и в защите аутентификационных данных от компрометации.

Пароли и криптографические ключи являются отчуждаемыми от владельца и поэтому подвержены «человеческому фактору». Длинный ключ (пароль) является надежным, только если были соблюдены все правила при его генерации — информационная энтропия ключа (пароля) должна быть сопоставима с его длиной. Но случайный длинный пароль почти невозможно запомнить. Выходом из ситуации является «привязка» всех ключей и паролей субъекта к его биометрическим параметрам с помощью преобразователя «биометрия-код» (ПБК) [1]. ПБК можно сравнить с интеллектуальным «черным ящиком», который «знает» своего владельца и безопасно хранит его пароль или криптографический ключ. ПБК обучается формировать и отдавать пользователю его пароль (ключ) при предъявлении биометрического образа. При предъявлении образа любого другого субъекта ПБК должен формировать случайный бинарный код, близкий по информационной энтропии к «белому шуму». Предполагается, что сами пароли и ключи генерируются перед обучением ПБК в соответствии с принятыми нормами. Данные обученного ПБК (ключ и биометрический эталон) должны быть защищены от компрометации при хранении и передаче по каналам связи без применения сторонних средств шифрования [2]. Хакеры не должны иметь возможность извлечь знания из обученного ПБК. Концепция ПБК может использоваться как основа для средств высоконадежной биометрической

аутентификации, а также электронной подписи с биометрической активацией.

Большинство методов биометрической аутентификации базируется на статических биометрических образах (отпечатках пальцев, радужке и т. п.). Статические образы наиболее уязвимы перед атаками представления, так как их невозможно держать в секрете. Открытый образ может быть изучен злоумышленником и фальсифицирован. Поэтому для аутентификации желательно использовать тайный образ, характеризующий особенности воспроизведения пароля его владельцем.

Настоящее исследование посвящено разработке модели ПБК для высоконадежной аутентификации на основе рукописных паролей.

Описание проблемы

Биометрический образ (пример образа) — это биометрические данные человека, подвергающиеся в дальнейшем масштабированию, удалению шумов и другой обработке в целях вычисления вектора биометрических параметров (признаков). При обучении ПБК создается биометрический эталон пользователя, который связывается с криптографическим ключом или паролем. Высвобождение ключа (пароля) происходит на этапе аутентификации, шифрования/дешифрования контента или создания электронной подписи. Предполагается, что обучение должно выполняться в доверенной среде, но обученный ПБК размещается в потенциально враждебной среде. При этом во многих приложениях, где требуется обеспечить анонимность пользователей (например, в медицине), ПБК каждого пользователя должен быть обезличен (не связан с его персональными данными).

Далее под *ключом* будет подразумеваться как непосредственно ключ шифрования или электронной подписи, так и пароль пользователя. Принципиального отличия в реализации ПБК при связывании биометрии с паролем или ключом нет. В зависимости от применения ПБК на практике могут предъявляться требования к длине и информационной энтропии ключа, а также соответствующим свойствам ПБК.

Наконец, следует дополнительно пояснить, что понимается под высокой надежностью работы ПБК. Надежность определяется показателями *FRR* и *FAR* — вероятностями ошибок «ложного отказа» (ответ ПБК не совпадает с ключом субъекта при предъявлении образа легитимным пользователем («Своим»)) и «ложного допуска» (ответ ПБК совпадает с ключом пользователя при предъявлении образа нелегитимным субъектом («Чужим»)).

Чтобы балансировать FRR и FAR, требуется изменять *порог принятия*. На практике для установки ненулевого порога требуется дополнительно корректировать ответ ПБК, например, с помощью кодов, исправляющих ошибки, и хранить синдромы ошибок, что снижает защищенность биометрического эталона и ключа от компрометации. *Нулевой порог принятия* означает, что система принимает ответ как правильный, только если он строго равен ключу пользователя (расстояние между ответом ПБК и ключом равно нулю).

В настоящей работе рассматривается атака *«извлечения знаний» из ПБК*, направленная на фальсификацию биометрического образа и получение злоумышленником личного ключа субъекта быстрее, чем при реализации атаки полного перебора ключей. Как и в случае атак на парольные системы, злоумышленник пытается сократить количество вариантов для перебора биометрических образов.

Сценарии атак на биометрические системы классифицируют по уровню знаний нарушителя об атакуемой стороне [3]. Будем исходить из того, что алгоритмы извлечения признаков из рукописного образа и функционирования ПБК не являются секретными. Кроме того, злоумышленник обладает параметрами обученного ПБК и выборкой образов «Чужие» неограниченного объема. Он может построить обученный ПБК, чтобы реализовать перебор образов «Чужие» с целью получить на выходах ПБК личный ключ пользователя, правильность которого он может проверить (рис. 1), не имея представления о том, какие именно выходные биты ПБК не совпали с соответствующими битами ключа пользователя (какие разряды оказались ошибочными). Битовую последовательность на выходе из ПБК будем называть *ответом*.

■ ***Рис. 1.*** Попытка подбора рукописного пароля для дешифрования контента

■ ***Fig. 1.*** Trying to matching handwriting password for decrypting the content

Тем не менее злоумышленник не обладает цифровой копией образа пользователя-жертвы (т. е. полными данными о написании пароля, включая динамику изменения скорости и давления пера), считается, что в этом случае ПБК будет скомпрометирован. Однако разглашение текстового содержания и даже компрометация внешнего вида рукописного пароля не должна приводить к компрометации ПБК. В этом случае пользователю требуется время, чтобы повторно обучить ПБК с использованием другого тайного биометрического образа. Именно такой сценарий атаки рассматривается в работе: злоумышленнику известны все алгоритмы, у него имеется база данных с параметрами обученных ПБК, имеется изображение рукописного образа пользователя-жертвы.

Стандарты ПБК и достигнутые ранее результаты

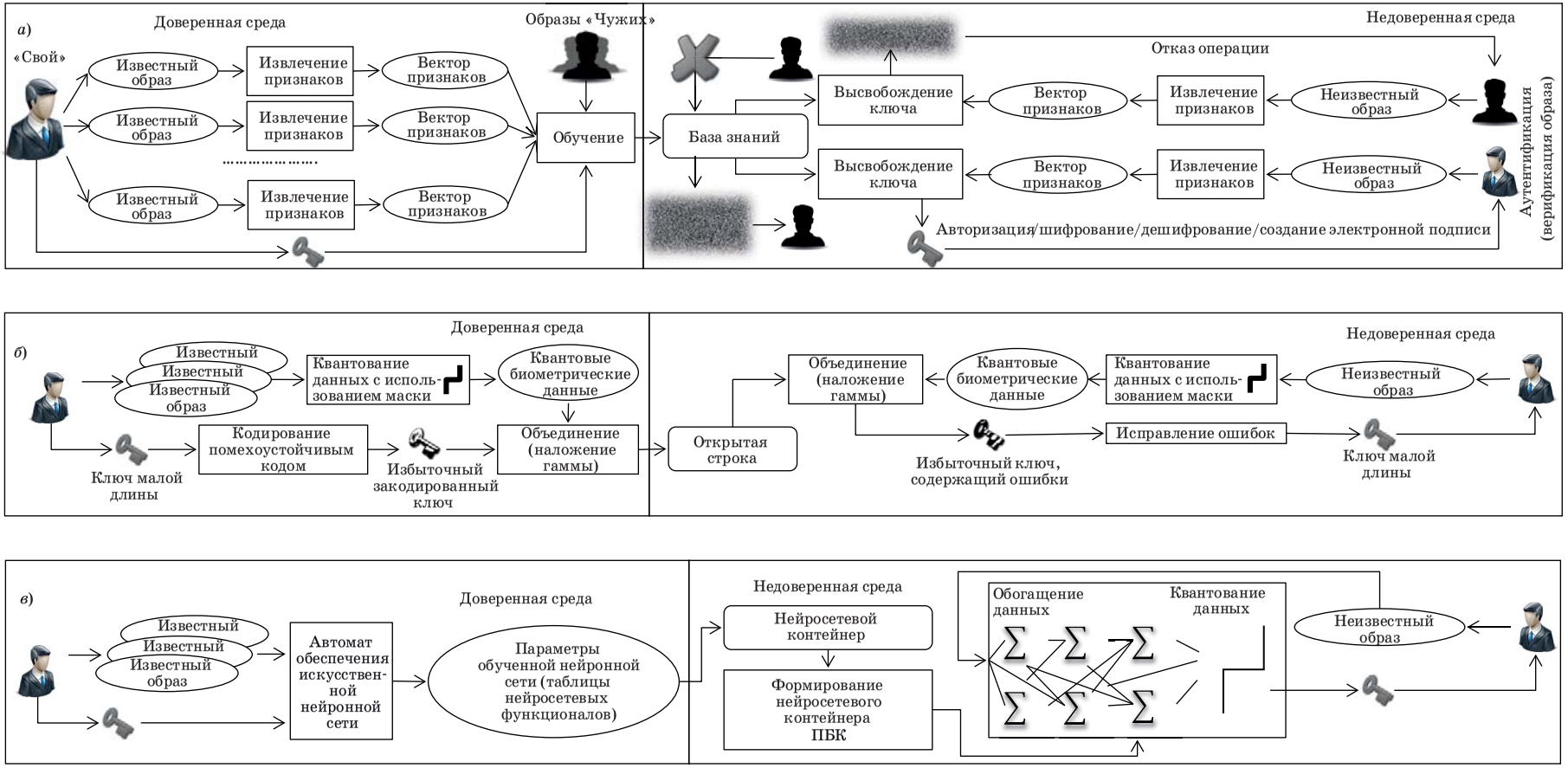
Общая схема работы ПБК представлена на рис. 2, *а*. На текущий момент сложилось два основных подхода к построению ПБК [4]: на основе нечеткого экстрактора (рис. 2, *б*), в рамках которого для дополнительной защиты применяется криптография (стандарты ISO/IEC 19792:2009 [5], 24761:2009 [6] и 24745:2011 [7]), и нейросетевой подход, поддерживаемый серией ГОСТ Р 52633 [8–14] (рис. 2, *в*), не имеющих к настоящему времени мировых аналогов.

В основе классического нечеткого экстрактора [15] применяются методы помехоустойчивого кодирования для исправления ошибок, возникающих из-за невозможности точного повторного воспроизведения биометрического образа. Этот подход имеет принципиальные недостатки [4]. В работе [16] приводятся уязвимости нечетких экстракторов, позволяющие выполнять направленный перебор биометрических данных для получения ключа. Об утечке конфиденциальности в некоторых схемах нечетких экстракторов идет речь и в других работах [17]. Также нечеткие экстракторы не способны к полноценному обучению. Они квантуют «сырые» биометрические данные, при этом подавляются шумы оборудования, но не учитывается характер распределения значений признаков пользователей. По этой причине их можно применять, только если образы высокоинформативные (как отпечаток пальца или радужка). Для рукописных образов число ошибок оказывается значительным [18].

Искусственные нейронные сети (ИНС) кодируют данные об особенностях признаков пользователей весовыми коэффициентами, что не дает прямого наблюдения за биометрическими эталонами [19]. Нейросетевой ПБК строится персонально для каждого субъекта, при этом формируется ИНС, количество входов которой равно числу признаков, а количество выходов — длине его личного ключа (см. рис. 2, *в*). Каждый нейрон последнего слоя генерирует один бит. Нейронная сеть обучается на биометрических образах пользователя и образах «Чужих», чтобы вырабатывать ключ субъекта при поступлении на вход его биометрического образа. Хорошо обученная нейронная сеть не нуждается в дополнительной корректировке выходов (ответа). Обучение нейросетевого ПБК должно быть абсолютно устойчивым, при этом объем обучающей выборки «Чужие» может быть сколь угодно большим. Разработчик биометрической системы может заготовить репрезентативную выборку «Чужие» заранее и использовать ее для обучения каждого ПБК. Однако число примеров образа «Свой» должно быть малым (по ГОСТ Р 52633.5-2011 достаточно 11 примеров [13]), нельзя заставлять пользователя сотни раз вводить биометрический образ. Это обстоятельство накладывает существенные ограничения на архитектуру ИНС, используемую в основе ПБК.

На сегодня построить ПБК на основе многослойных нейронных сетей затруднительно [4]. Исследования показывают, что на практике для обучения «глубокой» сети требуется более сотни примеров рукописного образа «Свой», чтобы надежность решений была приемлемой [1]. Практически все итерационные алгоритмы обучения теряют устойчивость при изменении параметров ИНС или объема обучающей выборки. Чем ниже качество биометрического образа, тем больший объем выборки требуется. Поэтому построение ПБК с большим количеством бинарных выходов на базе многослойных ИНС видится затруднительным (уже при длине ключа/пароля 32 бита количество классов нейронной сети должно составлять $2^{32} = 4\ 294\ 967\ 296$, что больше половины населения планеты). Кроме того, сверточные сети имеют ряд уязвимостей, обусловленных их изначальной ориентированностью на обработку графических образов [3] (наложение различного вида шума на изображение подписи существенно увеличивает FAR).

Активные исследования ведутся в области неглубоких сетей (shallow networks), способных к универсальной аппроксимации (для этого требуется потенциально неограниченное число скрытых нейронов, которое играет роль сложности модели ИНС). На данный момент проведена оценка ограничений малых ИНС, сформулирован ряд теорем [20], получены нижние оценки сложности ИНС в зависимости от соотношения между областью значений аппроксимируемой функции и размерностью входа [21].



■ ***Рис. 2.*** Иллюстрация принципов работы ПБК (слева — обучение ПБК, справа — высвобождение ключа): *а* — общая схема; *б* — нечеткий экстрактор; *в* — нейросетевой ПБК

■ ***Fig. 2.*** Illustration of the principles of operation of converters “biometrics to code” (on the left — training, on the right — the release of the key): *а* — general scheme; *б* — fuzzy extractor; *в* — neural network converter

■ ***Таблица 1***. Достигнутый уровень надежности при аутентификации по рукописным образам

■ ***Table 1***. Achieved reliability level for handwritten authentication

Метод/подход, особенности	FRR, %	FAR, %	С учетом подделок
Многослойный персептрон (MLP) + метод главных компонент PCA [23]	6,4	7,4	+
Авторский метод реализации ПБК [24]. Ключ 256 бит. Защита повышает число ошибок	38,75	13,45	–
CNN + метод опорных векторов [25]. Обучающая выборка «Свой», 30 примеров	2,17	13	+
CNN [26]. Обучающая выборка «Свой», 30 примеров	1,48 2,63	1,48 2,63	+ +
CNN [27]	2,42	2,42	+
Рекуррентные ИНС [27]	2,37	2,37	+
Нечеткий экстрактор [28]	9	9	–
«Широкие» ИНС [29]	10	10^{-7}	–

Большие нейронные сети из малого числа слоев (одного или двух) легли в основу стандартов ГОСТ Р 52633. Эти сети принято называть «широкими» [4]. Важным отличием «широких» сетей является процедура автоматического и абсолютно устойчивого послойного обучения (без использования алгоритма градиентного спуска). Идеологом и основателем научного направления, связанного с «широкими» ИНС и нейросетевыми ПБК на их основе, является А. И. Иванов [22]. За последние 20 лет под его авторством вышло множество работ, посвященных данному направлению. Сегодня развитием нейросетевых ПБК занимаются преимущественно ученые из России и Казахстана [1, 19, 22].

Приведем сопоставительные данные о надежности биометрической аутентификации на основе рукописных образов (подписей и паролей) при помощи методов, которые позволяют защитить эталоны и личные ключи субъектов от компрометации (табл. 1).

Преимущество «широкой» сети — в устойчивом обучении, увеличение числа нейронов влияет на время обучения линейно, в отличие от «глубокой» сети.

База рукописных образов

Для проведения исследований собрана база естественных рукописных образов. В соответствии с ISO/IEC 19795-3 [30] при формировании базы были учтены следующие факторы:

— «старение эталона»: база включает рукописные образы испытуемых, полученные в разные дни (с интервалом до нескольких недель);

— пол и возраст распределены равномерно на интервале от 18 до 35 лет;

— усилия злоумышленника: образ пароля каждого испытуемого пытались повторить еще пять субъектов по 10 раз, предварительно изучив его внешний вид на экране монитора и имея представление о темпе почерка его владельца. Согласно ISO/IEC 19795-3 [30], такой вид подделки рукописного образа соответствует 4-й степени фальсификации.

Для ввода рукописных образов использовались планшеты фирмы Wacom с частотой опроса 200 точек в секунду и 1024 уровнями давления пера на планшет. Ввод осуществлялся при помощи программного модуля, разработанного на языке C++ для семейства ОС Windows, все примеры сохранены как тестовые файлы. На момент эксперимента база насчитывала более 27 000 примеров рукописных паролей 260 испытуемых, включая примеры их подделок.

Предлагаемая процедура извлечения признаков

В компьютерном представлении рукописный образ состоит из функций положения пера $x(t)$, $y(t)$ и давления пера на планшет $\boldsymbol{p(t)}$ (аналогом является сила нажатия, которую способны регистрировать некоторые модели), где t — это время в дискретной форме. Образ преобразуется в вектор признаков $\bar{\mathbf{a}} = \{a_1, \ldots, a_N\}$ фиксированной длины ($N = 782$). «Широкие» ИНС позволяют использовать как можно больше признаков и не рассматривать разные методы их извлечения как альтернативные. Поэтому различные подходы к извлечению признаков могут и должны применяться совместно. В настоящей работе комбинировались следующие способы извлечения признаков [4, 18]:

— образ делится на 16 равных отрезков, строится матрица расстояний между их краями в двух- и трехмерном пространстве ($p(t)$ — третье измерение);

вычисляются:

— коэффициенты корреляции между $x(t)$, $y(t)$, $p(t)$, $x'(t)$, $y'(t)$, $p'(t)$ и функцией скорости пера $v_{xy}(t)$, производной от $x(t)$, $y(t)$;

— параметры внешнего вида образа: угол наклона, отношение длины к ширине, центр в трехмерном пространстве с координатами x, y, p;

— средние значения фрагментов функций $p(t)$, $x'(t)$, $y'(t)$, $v_{xy}(t)$ (образ делится на пять равных по числу точек отрезков);

— детализирующие коэффициенты быстрого вейвлет-преобразования Хаара (алгоритм Малла), полученные на четырех уровнях разложения (низкие частоты) для $x(t)$, $y(t)$, $p(t)$, $v_{xy}(t)$;

— усредненный амплитудный спектр, полученный с помощью Short-Time Fourier Transform (размер окна — 128 отчетов, шаг — 16 отчетов) для $x(t)$, $y(t)$, $p(t)$, $v_{xy}(t)$.

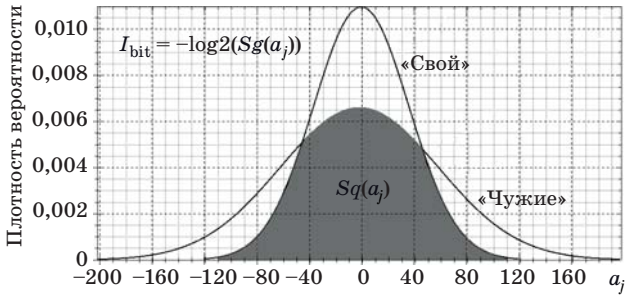
Сочетание нескольких методов получения признаков не снижает вероятность случайного совпадения и даже затрудняет попытки намеренной подделки.

В работе [31] предложена универсальная шкала для оценки информационной емкости биометрического образа. Информативность j-го признака оценивается через построение функций плотности вероятности для классов образов «Свой» и «Чужой» (рис. 3). Закон распределения большинства представленных признаков близок к нормальному [18] (реже встречается логнормальное и двойное экспоненциальное распределения).

После построения функций плотности вероятности вычисляется площадь их пересечения $Sq(a_j)$, которая равна вероятности ошибки при классификации образа по признаку a_j (интеграл от функции плотности вероятности равен вероятности). Вероятность переводится в собственную информацию по формуле

$$I_{\text{bit}} \approx -\log_2 Sq(a_j).$$

Для каждого подписанта следует выбирать наиболее информативные признаки индивидуально — при обучении нейросетевого ПБК. Информативность отдельно взятого признака I_{bit} — это показатель его уникальности. Но на общее количество информации (о субъекте) в биометрическом образе влияют также корреляционные связи между признаками. Часть информации всегда «переходит» в матрицу парных коэффициентов корреляции признаков, которая почти уникальна у каждого рукописного образа. Эту дополнительную информацию (как будет показано далее) можно использовать при распознавании.



■ ***Рис. 3.*** Информативность j-го признака для одного из испытуемых

■ ***Fig. 3.*** Information capacity of the j-th feature for some subject

Нейросетевые ПБК на базе классических нейронов

Первый слой классической «широкой» нейронной сети, обучаемой по алгоритму ГОСТ 52633.5, обогащает входные данные, второй — играет роль кодов, исправляющих ошибки [4]. Однако по сравнению с нечеткими экстракторами нейросетевая коррекция ошибочных разрядов ключа обладает гораздо меньшей избыточностью [19].

Рассмотрим однослойные «широкие» сети. Классический нейрон базируется на функционале

$$y = \sum_{j=1}^{n} \mu_j a_j \qquad (1)$$

и пороговой функции активации

$$f(y) = \begin{cases} 0, \text{ если } y < \mu_0 \\ 1, \text{ если } y > \mu_0 \end{cases}; \qquad (2)$$

модули весов нейронов первого слоя вычисляются по формуле [4]

$$\mu_j = |m_s(a_j) - m_o(a_j)| \,/\, \sigma_s(a_j) \cdot \sigma_o(a_j), \qquad (3)$$

где y — отклик нейрона на образ «Свой» или «Чужой»; n — количество входов нейрона; a_j — значение j-го признака (входа нейрона); $f(y)$ — ответ нейрона; μ_0 — порог активации нейрона; $m_o(a_j)$ и $\sigma_o(a_j)$ — математическое ожидание и среднеквадратичное отклонение значений j-го признака для образа «Свой»; $m_s(a_j)$ и $\sigma_s(a_j)$ — аналогичные показатели образов для «Чужих». Таким образом, сеть из нейронов (1) с функцией активации (2) после обучения представляет собой нейросетевой ПБК. Ответ нейросетевого ПБК складывается из битовых значений на выходах нейронов (путем их конкатенации).

Если нейрон настроен на выход «1» при поступлении образа «Свой», то знак весового коэффициента выбирается исходя из правила: «+» при $m_s(a_j) < m_o(a_j)$, иначе «−». Если нейрон настраивается на «нулевой» бит, то знаки инвертируются. Параметры $m_o(a_j)$, $\sigma_o(a_j)$, $m_s(a_j)$ и $\sigma_s(a_j)$ после обучения удаляются, чтобы не компрометировать эталон. Остаются таблицы связей и весов μ, из которых нельзя непосредственно вычислить $m_o(a_j)$. Параметры обученных нейронов (связи и веса) называют *нейросетевым контейнером*.

Порог μ_0 нейрона обычно настраивается исходя из откликов нейрона на обучающие примеры «Чужой» по правилу

$$\mu_0 = m_s(y)\alpha, \quad (4)$$

где α — единый для всех нейронов эмпирически подбираемый коэффициент, влияющий на баланс между FRR и FAR. В теории такой способ должен дать вероятность ошибки «ложного принятия» нейроном образа «Чужой», приближенно равную 0,5. Если допустить, что выходы нейронов независимы, каждый добавляемый нейрон, который настраивается по формуле (4), должен снижать FAR примерно в 2 раза. Информационная энтропия (далее просто энтропия) ответов нейросетевого ПБК на образы «Чужих» в этом случае должна быть близка к длине ключа. Но в действительности чем больше входов у нейронов, тем выше корреляция между их выходами и FAR, но чем больше информации поступает на вход каждому нейрону, тем ниже FRR. Действует и обратная логика: чем больше нейронов, тем ниже FAR, но выше FRR.

Выполненные в настоящей работе оценки показали высокий уровень FRR при настройке порогов нейронов по формуле (4), поэтому предложен альтернативный способ настройки порогов нейронов исходя из откликов y на примеры «Свой», не использовавшиеся при вычислении весов:

$$\mu_0 = m_o(y) - \sigma_o(y)\alpha.$$

Число входов каждого нейрона предлагается определять так, чтобы сумма информативности связанных с ними признаков была $\Sigma I_{\text{bit}} > 1$. Номера связанных с нейроном признаков определяются случайно, но среди тех, коэффициент корреляции r между которыми принимает высокие значения ($r > 0{,}5$).

Предлагаемая модель гибридного ПБК на базе разностных нейронов Байеса и классических нейронов

Функционал (1) теряет мощность при усилении корреляционных связей между признаками. По этой причине количество информации на входе классического нейрона всегда меньше, чем сумма собственной информации (см. рис. 3) всех признаков. Абсолютно иначе дело обстоит, если вместо функционала (1) использовать разностный байесовский функционал

$$d_t = \sum_{j=1}^{n}\left|\frac{|m_o(a_t)-a_t|}{\sigma_o(a_t)} - \frac{|m_o(a_j)-a_j|}{\sigma_o(a_j)}\right|, \; j \neq t. \quad (5)$$

Многомерный разностный функционал Байеса (5) дает тем меньшее значение, чем выше коэффициент корреляции признака под номером t с признаками под номерами j [4]. На его основе возможна нейросетевая обработка коррелированных сочетаний признаков. Недостатком является то, что он полностью компрометирует биометрический эталон (параметры $m_o(a_j)$ не должны использоваться при расчете близости после обучения ПБК).

Предлагается функционал, обладающий аналогичными свойствами, но компрометирующий биометрический эталон лишь частично:

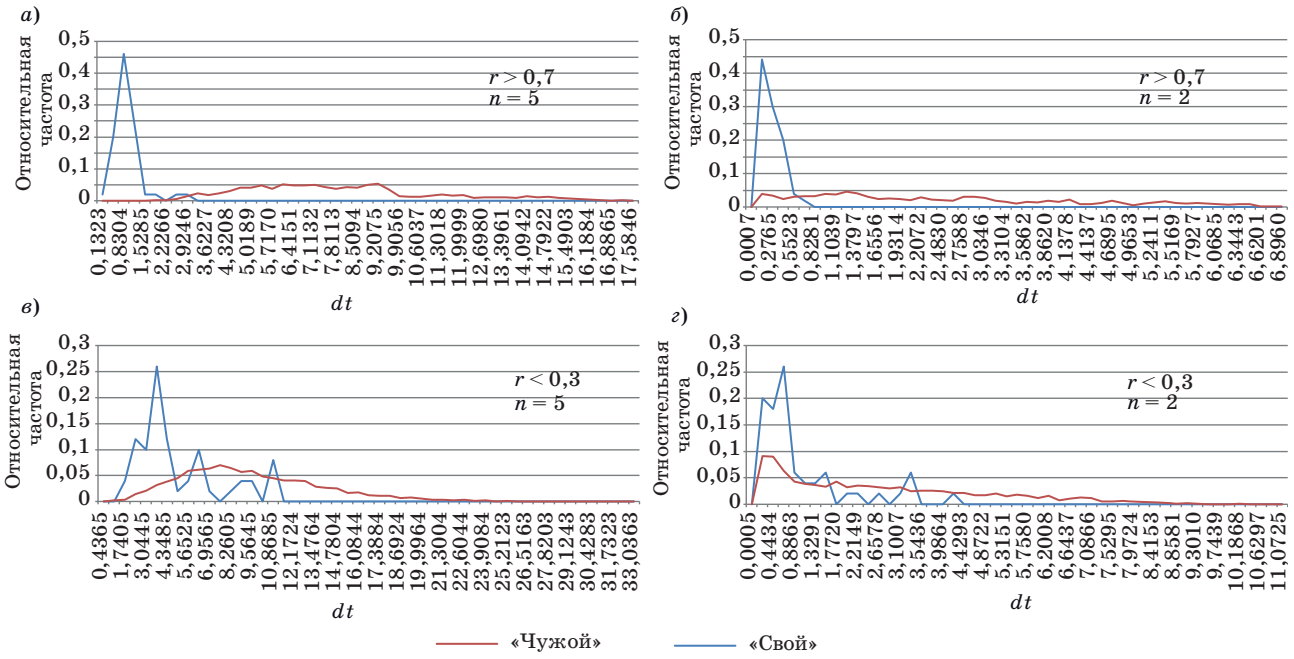
$$d_t = \sum_{j=1}^{n}\left|\left|\frac{a_t}{\sigma_o(a_t)} - \frac{a_j}{\sigma_o(a_j)}\right| - \Delta m_{tj}\right|,$$
$$\Delta m_{tj} = \left|\frac{m_o(a_t)}{\sigma_o(a_t)} - \frac{m_o(a_j)}{\sigma_o(a_j)}\right|, \; j \neq t. \quad (6)$$

Свойства функционала (6) демонстрируются на рис. 4, *а–г*: чем больше n и выше корреляция между признаками r, тем меньше вероятность ошибок распознавания образа. Это справедливо для любых признаков, независимо от их физического смысла. Убедиться в этом несложно, достаточно сгенерировать описания абстрактных классов образов в двух пространствах признаков с нормальным законом распределения: независимых и зависимых. Сгенерировать значения независимых признаков под заданные параметры распределения можно методом Монте-Карло (классы должны отличаться параметрами $m_o(a_j)$ и $\sigma_o(a)$). Воссоздать положительную зависимость между признаками можно, отдельно сортируя значения каждого признака по возрастанию [31]. Далее следует построить нейрон Байеса на основе функционала (6), сформировать образы из сгенерированных данных (в двух вариантах — на основе независимых признаков и зависимых) и обработать эти образы при помощи нейрона Байеса, после чего построить эмпирические плотности вероятности откликов (6) этого нейрона (см. рис. 4).

Разностные нейроны Байеса на базе функционала (6) с функцией активации (2) обучаются путем вычисления параметров Δm и $\sigma_o(a)$. Это позволяет не хранить параметры $m_o(a)$. Введем также функционал

$$d_t = \sum_{j=1}^{n}\left|\frac{a_t}{\sigma(a_t)} - \frac{a_j}{\sigma(a_j)}\right|, \; j \neq t \quad (7)$$

обладающий идентичными свойствами в плане обработки сильно коррелированных признаков. Преимущество метрики (7) заключается в том,



■ ***Рис. 4.*** Эмпирические плотности вероятности откликов разностных нейронов Байеса (6) при распознавании абстрактных образов при $I_{\text{bit}} \approx 0{,}5$ в случае зависимых ($r > 0{,}7$) и независимых ($r < 0{,}3$) признаков: *а*, *в* — $n = 5$; *б*, *г* — $n = 2$

■ ***Fig. 4.*** Empirical probability densities of responses of Bayes difference neurons (6) when recognizing abstract images at $I_{\text{bit}} \approx 0.5$: in the case of dependent ($r > 0.7$) and independent ($r < 0.3$) features: *а*, *в* — $n = 5$; *б*, *г* — $n = 2$

что для ее корректной работы требуется хранить только параметры $\sigma(a)$. Это могут быть среднеквадратичные отклонения признаков как для класса образов «Свой» $\sigma_o(a)$, так и для класса «Чужие» $\sigma_s(a)$, что также допустимо. При этом $\sigma_s(a)$ вообще не компрометирует биометрический эталон пользователя.

Порог срабатывания функции активации (2) для разностного нейрона Байеса вычисляется по правилу

$$\mu_0 = m_o(d_t) + \sigma_o(d_t)\beta,$$

где β — коэффициент баланса FRR и FAR для нейронов Байеса. Несмотря на идентичные свойства метрик (5)–(7), стопроцентной корреляции между их откликами d_t не наблюдается, что позволяет строить разностные нейроны Байеса на основе всех указанных метрик и объединять их в единую сеть. Каждый нейрон должен быть связан с признаками, коэффициент корреляции между которыми принимает достаточно высокие значения ($r_{t,j} > 0{,}5$). Количество входов нейрона не лимитируется: чем больше признаков с примерно равным (и высоким) уровнем взаимной корреляции входит в нейрон, тем стабильнее он работает.

Классические нейроны показывают хороший результат, если признаки достаточно информативны и слабо коррелированы ($r_{t,j} < 0{,}5$). Разностные нейроны Байеса обладают почти противоположными свойствами: они ориентированы на обработку сильно зависимых признаков ($r_{t,j} > 0{,}5$), так как неявно извлекают дополнительную информацию из корреляционной матрицы признаков, что было продемонстрировано на рис. 4. Таким образом, если объединить классические нейроны и разностные нейроны Байеса в гибридную сеть, можно снизить показатели FRR и FAR, а также повысить энтропию ответов ПБК. При формировании гибридной сети следует отдельно настраивать сегмент из классических нейронов и сегмент из байесовских нейронов, объединяя их в один слой. Формирование связей и настройка классических нейронов может производиться в соответствии с ГОСТ Р 52633.5 [13], с единственным обязательным отличием — при создании связей для каждого нейрона выбираются признаки с уровнем взаимной корреляции $r_{t,j} < 0{,}5$. При создании связей для нейрона Байеса признаки выбираются случайно [4], но среди тех, что имеют взаимную корреляционную зависимость $r_{t,j} > 0{,}5$.

При построении гибридного ПБК необходимо учитывать тот факт, что разностные нейроны Байеса компрометируют биты ключа пользователя. Каждый нейрон должен давать на выходе один бит ключа (верный или неверный, зависит от преодоления порога μ_0 нейроном), который необходимо хранить (решение этой проблемы будет показано далее).

Об оценке стойкости ПБК к атакам

Энтропия ответов ПБК на образы «Чужие» является важным показателем, так как она связана с FAR: чем ниже FAR, тем выше энтропия [19]. Приблизительную оценку энтропии можно получить, вычислив собственную информацию события «ложного допуска»:

$$E(\mathrm{FAR}) \approx -\log_2 \mathrm{FAR}.$$

Точный расчет многомерной энтропии длинных бинарных последовательностей прямым численным экспериментом является технически нерешаемой задачей, как и расчет сверхнизких показателей FAR. Чтобы получить $E(\mathrm{FAR}) = 256$ бит, требуется, чтобы $\mathrm{FAR} < 10^{-77}$. Для проверки такой экстремально низкой вероятности прямым численным экспериментом не хватит населения планеты (даже если каждый человек придумает тысячи независимых рукописных паролей). По причине сложности сбора больших выборок к биометрическим системам предъявляются не столь жесткие требования, как к паролям.

Рассмотрим подходы к оценке FAR в биометрических системах.

Первый подход заключается в проведении прямого численного эксперимента и вычислении FAR как отношения числа ошибок к числу опытов по распознаванию «Чужих» с определением доверительных вероятности и интервала. Такая идеология лежит в основе ISO/IEC 19795-1 [5]. Если следовать только этому подходу (например, правилам «трех» или «тридцати»), то оценить надежность ПБК, обладающего действительно высокой энтропией, невозможно.

Второй подход основан на построении двух функций плотности вероятности для расстояний Хэмминга между ключом (паролем) пользователя и ответом ПБК (нейронной сети или нечеткого экстрактора) на образы «Свой» и образы «Чужие» соответственно. Далее вычисляется площадь их пересечения (по аналогии с процедурой оценки информативности признака, см. рис. 3). Если задать *порог принятия*, то можно вычислить FRR и FAR [19]. Однако данный способ дает приблизительную оценку. Для большей точности вводятся поправки, учитывающие показатели стабильности, информативности и коррелированности признаков. Плотности вероятности для ПБК можно описать нормальным законом распределения лишь условно, так как для иных архитектур ИНС (например, гибридных, состоящих не только из классических нейронов) закон распределения расстояний Хэмминга может быть другой.

Третий подход изложен в ГОСТ Р 52633.3-2011 [11]. Согласно стандарту, в эксперименте необходимо использовать не только естественные образы «Чужой», но и синтетические, генерируемые на основе скрещивания естественных (по методике ГОСТ Р 52633.2-2010 [10]). Если для естественных образов «Чужих» не наблюдается ошибок 2-го рода («ложного совпадения ключа»), то естественные образы скрещиваются. При скрещивании следует выбирать пары из «Чужих», которые дают ответы ПБК, наиболее близкие в метрике Хэмминга к ключу пользователя. Далее по аналогичному принципу могут скрещиваться синтетические образы. Каждая новая популяция синтетических образов «Чужих» все ближе к образу «Свой». Процесс тестирования ПБК прекращается, когда для очередной популяции будут зафиксированы ошибки 2-го рода или синтезировано определенное число популяций «Чужих». В процессе тестирования сокращается количество попыток предъявления конкурирующих примеров, а точность оценки FAR увеличивается на порядки при сохранении статистической значимости.

Результаты по оценке вероятности ошибочных решений нейросетевых ПБК

При разработке алгоритма тестирования ПБК решено опираться на третий подход (ГОСТ 52633.3-2011 [11]) и метод перекрестного сравнения (кроссвалидации) [4, 18].

По требованиям ГОСТ 52633.3 для тестирования средств высоконадежной биометрической аутентификации при условии высокого уровня взаимного доверия между донором биометрии и владельцем средства биометрической аутентификации (тестирования) необходимо не менее 128 примеров естественных образов «Чужих», которые ранее не были использованы при обучении тестируемого ПБК. Используемые при тестировании биометрические образы должны быть независимыми и формироваться по ГОСТ Р 52633.1 [9]. В соответствии с ГОСТ 52633.5 для обучения нейросетевого ПБК требуется не менее 64 независимых образов «Чужих» [13]. Под независимыми образами подразумеваются примеры различных рукописных паролей, воспроизведенных разными субъектами.

Для каждого испытуемого генерировался случайный ключ и формировался ПБК. Для обучения ПБК пользователя использовалось 15 примеров его образа и 64 примера образов других случайно выбранных испытуемых («Чужих»).

Далее проводились испытания надежности работы ПБК. Для оценки FRR использовалось по 30 образов от каждого испытуемого, не вошедших в обучающую выборку. FRR определялась как отношение числа зарегистрированных ошибок «ложного отказа» к общему количеству опытов ($30 \times 260 = 7800$).

Чтобы оценить вероятность ошибки «ложного допуска», относительно каждого испытуемого формировалась тестовая выборка образов «Чужих», которая состояла из примеров рукописного пароля оставшихся 195 подписантов, не вошедших в выборку обучения (бралось по 10 образов на подписанта). Кроме того, в тестовую выборку испытуемого вошли 50 подделок его образа. Так для каждого испытуемого сформирована тестовая выборка из 2000 естественных образов «Чужих» (всего $2000 \times 260 = 520\,000$ примеров).

Тестирование FAR выполнялось независимо для каждого испытуемого на основании отобранных 2000 примеров «Чужих». Если ошибки не фиксируются, производится скрещивание 10 % «Чужих» (20 образов из 2000, воспроизведенных различными подписантами), которые дают ответы ПБК, наиболее близкие в метрике Хэмминга к ключу пользователя, по следующей формуле:

$$a_{j,k} = \frac{c+1-k}{c+1} a_{j,A} + \frac{k}{c+1} a_{j,B},$$

где c — количество синтетических примеров, порождаемых парой «сильных Чужих» A и B предыдущего поколения; k — номер синтетического примера; j — номер признака. Этот способ синтеза «Чужих» сохраняет естественные корреляционные связи образов A и B.

Формируется новая популяция из 2000 синтетических образов, для которых также вычисляются ответы ПБК. Далее по аналогичному принципу скрещиваются синтетические образы. Тестирование повторяется, пока не будут зафиксированы ошибки или расстояние Хэмминга между ответами ПБК и ключом испытуемого не перестанет уменьшаться. FAR определялась по формуле

$$FAR = \frac{\sum_{i=1}^{260} \sum_{p=0}^{p\_\max_i} \frac{er_{i,p}}{2 \cdot 10^{3+p}}}{260},$$

где $p\_\max_i$ — количество популяций «Чужих» для i-го испытуемого; $er_{i,p}$ — количество ошибок «ложного допуска» для i-го испытуемого из блока при тестировании на примерах из p-й популяции.

Описанная методика дает точную и достоверную ненулевую оценку FAR. Выборочные результаты тестирования нескольких вариаций ПБК на базе однослойных «широких» сетей представлены в табл. 2.

При обучении MLP на больших объемах данных «глубину» персептрона обычно стараются повысить, а размерность пространства признаков снизить (например, с помощью PCA). Для нейросетевых автоматов биометрической аутентификации увеличение количества слоев не актуально. Гораздо эффективнее соизмеримо повышать число входов и выходов сети. Это приводит к медленному росту качества решений — снижению вероятностей ошибок и повышению энтропии ответов ПБК при предъявлении образов «Чужих» (см. табл. 2). Однако рост постепенно останавливается. После этого наращивать число нейронов не имеет смысла. По этой причине энтропия ответов нейросетевых ПБК на образы «Чужих» не соответствует их длине. Такой подход к повышению стойкости ПБК имеет общие черты с повышением стойкости парольной защиты путем увеличения длины их контрольных сумм (хотя реальная стойкость также зависит от степени случайности самих паролей).

■ ***Таблица 2***. Результаты тестирования ПБК на базе классических нейронов

■ ***Table 2***. Results of testing classical neural net "biometrics to code" converters

Число нейронов L и входов n	FRR, % (порог = 0)	FAR, % (порог = 0)	EER, % (порог > 0)	α
$L = 1024$, $n \geq 5(\Sigma I > 1)$	14	0,0016	1,9	4,75
$L = 512$, $n \geq 5(\Sigma I > 1{,}5)$	17,5	0,0072	2,2	4
$L = 512$, $n \geq 5(\Sigma I > 1)$	14,5	0,0019	1,9	4,5
$L = 512$, $n = 10$	16	0,0058	2	4
$L = 512$, $n = 20$	23	0,0207	2,7	3
$L = 512$, $n = 150$	50	0,233	8,3	1,5
$L = 256$, $n = 10$	25	0,0094	2,3	3
$L = 128$, $n = 10$	19	0,0333	2,8	3
$L = 64$, $n \geq 5(\Sigma I > 1)$	32	0,0348	3,9	3
$L = 64$, $n = 10$	26,5	0,0462	4,1	2,5
$L = 64$, $n = 10$	53,5	0,0125	4,1	2
$L = 64$, $n = 50$	12	0,4671	6	2

Установлено, что свыше 50 % «наиболее близких "Чужих"» попали в намеренные подделки. Обычно это происходит, когда пароль испытуемого является достаточно простым (из четырехпяти символов). Самые сильные синтетические образы часто получались из двух подделок, воспроизведенных разными людьми. Таким образом, динамика рукописного пароля не является абсолютно устойчивой к подделкам, выполняемым путем копирования внешнего вида.

Результаты по тестированию нейросетевых ПБК на предмет возможности извлечения знаний

Процедура обучения «широкой» ИНС является однонаправленной и не подразумевает обратной разработки. Однако восстановление биометрического образа и личного ключа пользователя из таблиц нейросетевых функционалов все же возможно.

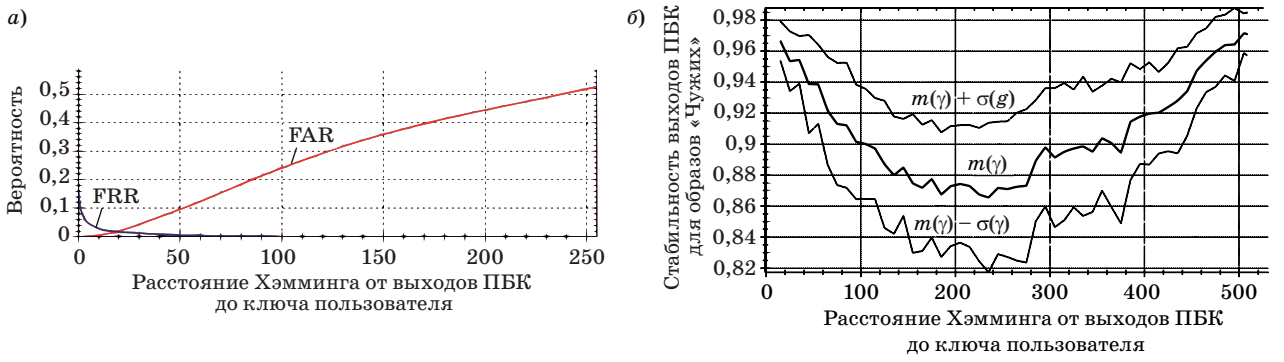
Контролируя допустимое число ошибочных бит в ответе «широкой» сети, можно балансировать FRR и FAR (например, применяя корректирующие коды или второй слой нейронов для исправления нескольких неверных бит) (рис. 5, *а*). Однако эта возможность одновременно является уязвимостью. Хакер может собрать большую базу примеров произвольных паролей, воспроизведенных различными подписантами («Чужими»), и оценить среднюю стабильность ответов ПБК для каждого подписанта («Чужого») по формуле

$$\gamma_k = \sum_{l=1}^{L} 2\left|P_l(1) - 0{,}5\right|, \qquad (8)$$

где k — номер подписанта («Чужого»); L — количество нейронов; l — номер нейрона; $P_l(1)$ — вероятность (или относительная частота) появления «единицы» (можно заменить на $P_l(0)$) в l-м разряде ответа ПБК (в выходе l-го нейрона) на примеры образа k-го «Чужого». Оценка относительных частот $P_l(1)$ и $P_l(0)$ может проводиться на основании нескольких образов подписанта (в настоящей работе использовалось по 10 примеров от каждого «Чужого»).

Вычислены показатели стабильности ответов ПБК (8) при предъявлении примеров каждого «Чужого». По результатам эксперимента показатель стабильности (8) для каждого «Чужого» оказался зависим от среднего количества ошибочных бит соответствующих ответов ПБК (рис. 5, *б*). Таким образом, даже в отсутствие явной индикации близости ответов ПБК и ключа пользователя хакер может осуществить направленный перебор синтетических образов, скрещивая примеры рукописных паролей разных «Чужих», которые дают наиболее стабильный ответ. Через несколько поколений скрещивания, «двигаясь» в направлении повышения стабильности ответов ПБК, удается подобрать «Чужого», почти или полностью идентичного образу «Свой» (в зависимости от объема исходной базы «Чужих», наличия в базе злоумышленника подделок образа «Свой» и того, насколько качественно они выполнены). Данная атака снижает количество вариантов перебора на несколько порядков. Даже если злоумышленник не обладает примерами подделок и какой-либо информацией о пароле пользователя-жертвы, данная атака вполне осуществима (в этом случае нарушителю потребуется гораздо больше времени).

Также видно (см. рис. 5, *б*), что образы «Чужие» обладают так называемой *«симметрией» стабильности ответов относительно образа «Свой»* (свойством «симметрии»). Это означает, что стабильность ответов ПБК при предъявлении образов «Чужих» возрастает, но не только если ответы близки (в метрике Хэмминга) к ключу пользователя, но и если они близки к инверсии ключа (инверсный код возникает, если все биты ответа ПБК являются ошибочными). Инверсный код можно обратить и получить ключ пользователя. Из этого следует, что у каждого образа «Свой» в нейросетевом логическом базисе существует его инверсия. Если на вход классического нейросетевого ПБК (обученного по ГОСТ 52633.5 [13]) по-



■ ***Рис. 5.*** Результаты тестирования классических нейросетевых ПБК с параметрами $L = 1024$, $n \geq 5$: *а* — вероятности ошибок в зависимости от порога принятия; *б* — стабильность ответов ПБК на образы «Чужих» в зависимости от количества ошибочных бит

■ ***Fig. 5.*** Test results of classical neural net “biometrics to code” converter with parameters $L = 1024$, $n \geq 5$: *а* — probability of errors depending on the threshold of acceptance; *б* — stability of converter responses to “Strangers” depending on the number of error bits

дать инверсию образа «Свой», то на выходе ПБК появится инверсный ключ, который можно обратить. Данное свойство позволяет ускорить процедуру направленного перебора биометрических образов в 2 раза (осуществляя одновременно поиск наиболее близкого и наиболее дальнего образа «Чужого» относительно образа «Свой»).

Предлагаемый способ защиты гибридных нейросетевых контейнеров

В работе [19] предложено защищать нейросетевые контейнеры путем применения обратимых и необратимых преобразований. Усовершенствуем данный подход, чтобы защитить гибридный ПБК.

Каждый нейрон имеет таблицы связей и весов. Для защиты таблиц нейросетевых функционалов нужно применять механизм защищенного нейросетевого контейнера (ЗНК) (рис. 6, *а* и *б*). Нейроны можно выстроить в цепочку (см. рис. 6, *а*). После обучения ПБК таблицы каждого нейрона шифруются наложением гаммы, представляющей собой контрольную сумму выходов всех предыдущих нейронов в цепочке:

$$tables'_l = \\ = XOR(tables_l, hash(pass, bit_1, \ldots, bit_{l-1})), \quad (9)$$

где $tables_l$ — таблицы параметров соответствующего нейрона; $hash()$ — криптографическая хеш-функция (например, md5); $pass$ — пароль, который является опциональным и служит для дополнительной (двухфакторной) защиты; bit_l — выход, на который настраивается l-й нейрон в цепочке. В настоящем исследовании пароль не использовался.

При обработке биометрического образа нейросетевым ПБК в режиме ЗНК происходит процесс «распаковки» нейронов — параметры каждого следующего нейрона в цепочке дешифруются по той же формуле (9). Для получения на выходе ПБК верного ключа пользователя требуется, чтобы все нейроны «проголосовали» правильно. Если хотя бы один нейрон в цепочке выдаст ошибочный бит, это повлечет неверную дешифровку параметров всех последующих нейронов. В свою очередь последующие нейроны будут давать случайные выходы, и возникнет эффект хеширования биометрического образа «Чужого». В итоге ответы нейросетевого ПБК становятся случайными, их энтропия возрастает. При этом важен тот факт, что FRR и FAR в режиме ЗНК не меняются при пороге принятия, равном нулю (рис. 7, *а*). Стабильность ответов ПБК при поступлении на вход образов «Чужих» становится низкой и перестает возрастать, если образ «Чужого» близок к образу «Своего» (график на рис. 5, *б* в режиме ЗНК становится почти прямым, рис. 7, *б*). Однако режим ЗНК все же накладывает ограничения: балансировать FRR и FAR, корректируя несколько ошибочных ответов ПБК, становится невозможно.

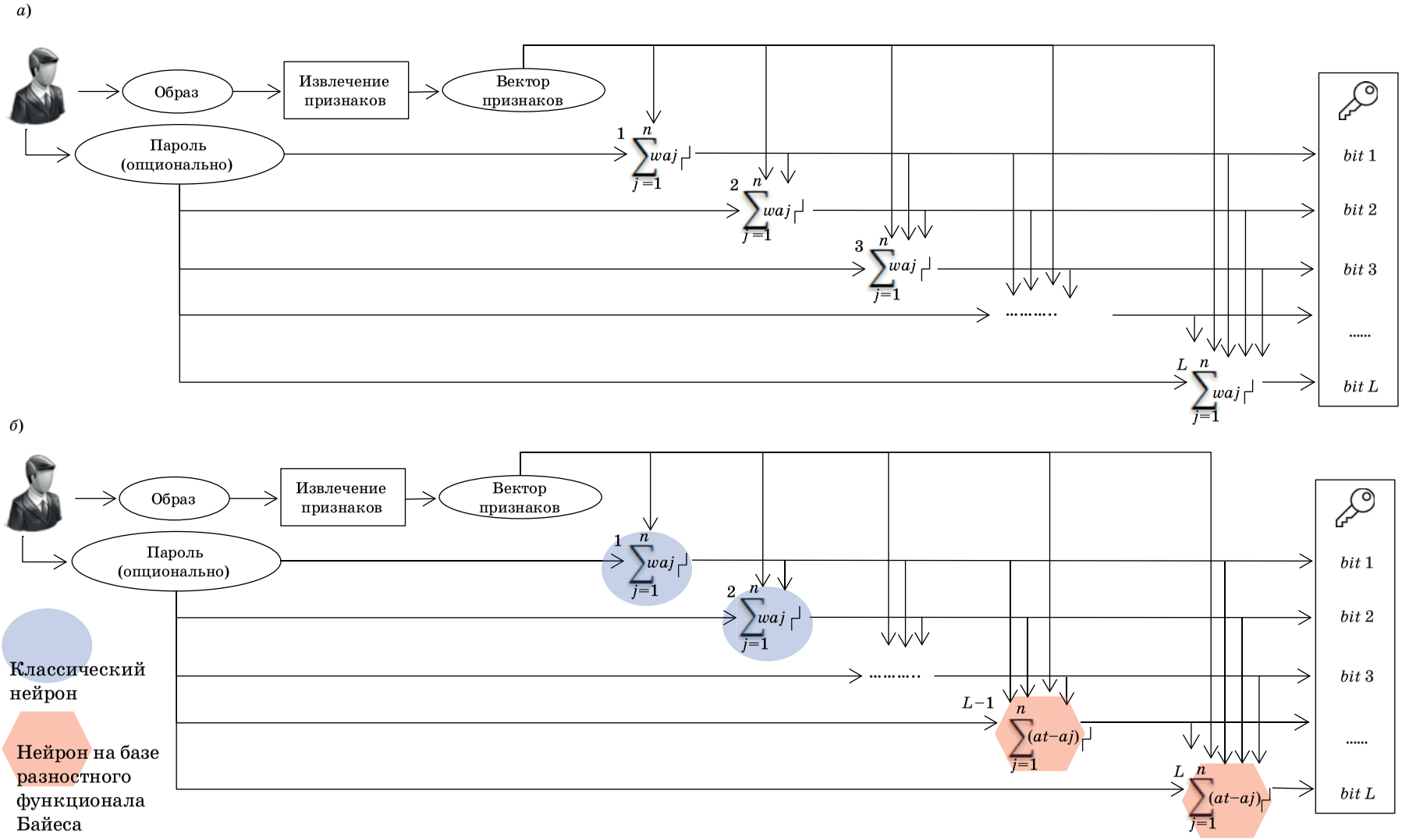
Результаты по оценке вероятностей ошибочных решений гибридных ПБК с применением предложенной схемы защиты

Разностные нейроны Байеса нужно использовать совместно с классическими нейронами, применяя механизм ЗНК (см. рис. 6, *б*). Предлагается размещать классические нейроны в начале цепочки, а байесовские нейроны — в конце (в силу того, что последние в незащищенном виде компрометируют часть ключа). Нейроны Байеса будут надежно защищены, если классических нейронов будет много (достаточно 256).

В режиме ЗНК порядок нейронов в цепочке (см. рис. 6, *б*) играет важную роль. Предлагается расположить классические нейроны в порядке уменьшения суммарной информативности их входов (ΣI_{bit}), а разностные нейроны Байеса расположить в порядке снижения размерности и корреляционной зависимости входов. В этом случае удается повысить энтропию ответов ПБК при идентичных показателях FRR и FAR. Этот эффект имеет простое объяснение. Если сначала располагаются нейроны с более стабильной статистикой выходов, то при поступлении образа «Чужой» эти нейроны среагируют первыми, и процесс «хеширования» запустится раньше (больше нейронов будет дешифровано неверно), при поступлении образа «Свой» ничего не изменится.

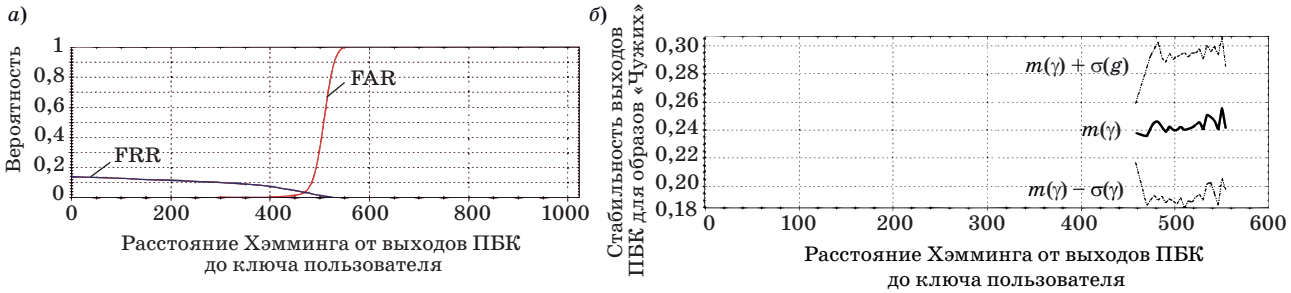
По результатам эксперимента (рис. 8, *а–в*) установлено, что средние показатели стабильности ответов гибридного ПБК при поступлении образов «Чужих» в режиме ЗНК гораздо ниже ($m(\gamma) = 0{,}24$), чем без защиты ($m(\gamma) = 0{,}92$), и еще ниже ($m(\gamma) = 0{,}23$) при ранжировании нейронов в соответствии с информативностью и коррелированностью входов. Также можно видеть, что у гибридного ПБК отсутствует уязвимость, связанная со свойством «симметрии» (см. рис. 8, *а*).

Установлено, что комплексирование двух видов нейронов существенно снижает вероятности ошибок: FRR = 11,5 %, FAR = 0,0009 %, EER ≈ ≈ 1,6 %, $L = 1024$ (512 классических и 512 нейронов Байеса), $\alpha = 4{,}25$, $\beta = 25{,}5$ (рис. 9, *а–в*). Результат [29] превосходит полученный в настоящей работе потому, что при тестировании в работе [29] не учитывались подделки и для вычисления FAR применялся менее точный метод



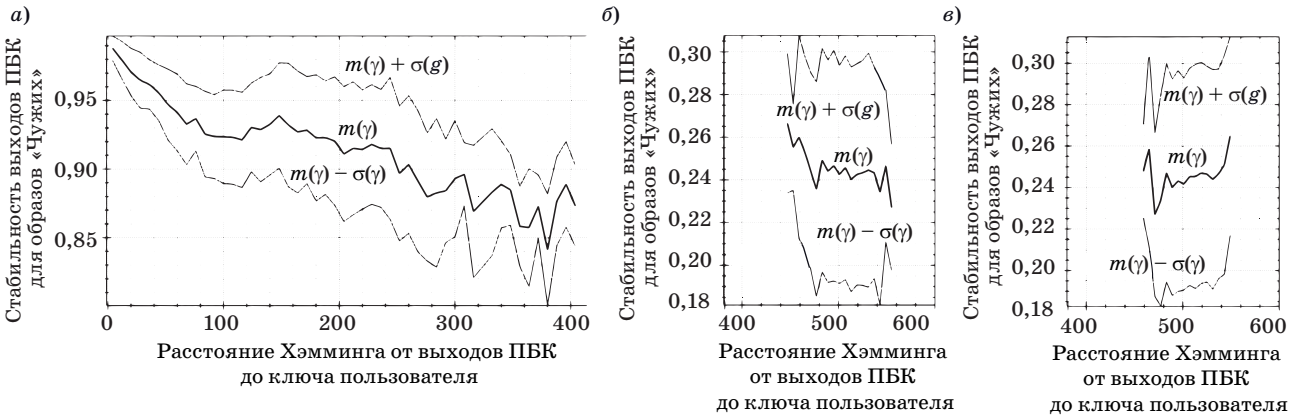
■ ***Рис. 6***. Механизм ЗНК применительно к нейросетевому ПБК на базе классических нейронов (*а*) и гибридному ПБК (*б*)

■ ***Fig. 6***. The mechanism of a protected neural network container in relation to converter “biometrics to code” based on classical neurons (*a*) and hybrid converter (*б*)



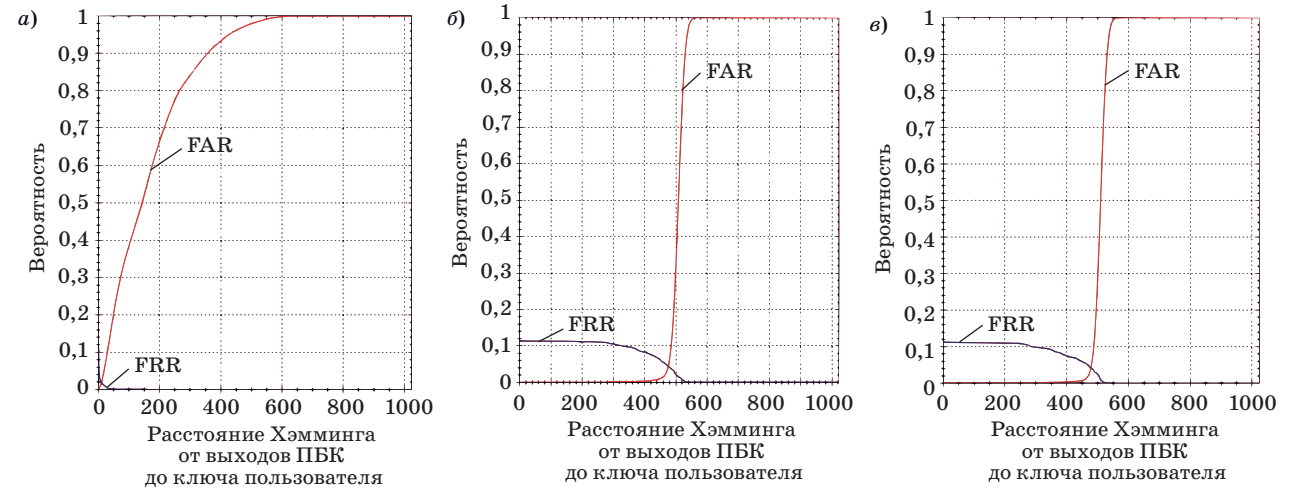
■ ***Рис. 7.*** Результаты тестирования классических нейросетевых ПБК в режиме ЗНК с параметрами $L = 1024$, $n \geq 5$: *а* — вероятности ошибок в зависимости от порога принятия; *б* — стабильность ответов ПБК на образы «Чужих» в зависимости от количества ошибочных бит

■ ***Fig. 7.*** Testing results of classical neural net "biometrics to code" converters in protection mode with parameters $L = 1024$, $n \geq 5$: *а* — probability of errors depending on the threshold; *б* — stability of converters responses to "Strangers" images depending on the number of erroneous bit



■ ***Рис. 8.*** Стабильность ответов гибридных ПБК при поступлении образов «Чужих» в зависимости от числа ошибочных бит: *а* — обычный режим; *б* — режим ЗНК; *в* — режим ЗНК с ранжированием нейронов

■ ***Fig. 8.*** Stability of hybrid "biometrics to code" responses to "Strangers" images depending on the number of erroneous bits: *а* — normal mode; *б* — protection mode; *в* — protection mode with ranking of neurons



■ ***Рис. 9.*** Вероятности ошибок в зависимости от порога принятия: *а* — обычный режим; *б* — режим ЗНК; *в* — режим ЗНК с ранжированием нейронов

■ ***Fig. 9.*** Probabilities of errors depending on the threshold: *а* — normal mode; *б* — protection mode; *в* — protection mode with ranking of neurons

(на базе второго подхода — оценки пересечения функций плотности вероятности для расстояний Хэмминга между ответами ПБК и ключами).

Снижение вероятностей ошибок указывает на то, что ошибочные решения нейронов Байеса слабо коррелированы с ошибками классических нейронов.

Заключение

Установлено, что механизм защищенного нейросетевого контейнера можно успешно применять в гибридных нейронных сетях, состоящих из классических нейронов и разностных нейронов Байеса. Предположены новые варианты построения разностных нейронов Байеса, не компрометирующих и частично компрометирующих биометрический эталон пользователя (даже без применения метода защиты нейросетевых контейнеров). Продемонстрирована их эффективность при распознавании образов в пространстве сильно коррелированных признаков.

Экспериментально подтверждена высокая надежность верификации рукописных образов на базе предложенной модели гибридной нейробайесовской сети (с учетом предъявления подделок рукописных паролей испытуемых). Показатели ошибок аутентификации (высвобождения ключа пользователя) составили: FRR = 11,5 %, FAR = 0,0009 % (EER ≈ 1,6 %) при длине ключа 1024 бита. Достигнутые показатели не являются предельными.

Направления будущих исследований могут быть связаны с применением механизмов защиты нейросетевых контейнеров в отношении других архитектур гибридных нейронных сетей, способных к быстрому обучению.

Финансовая поддержка

Работа выполнена при поддержке Российского научного фонда по гранту № 17-71-10094.

Financial support

This work was supported by the Russian Science Foundation, No. 17-71-10094.

Литература

1. **Akhmetov B. S., Ivanov A. I., Alimseitova Z. K.** Training of neural network biometry-code converters. *2018 News of the National Academy of Sciences of the Republic of Kazakhstan. Series of Geology and Technical Sciences*, 2018, vol. 1, no. 427, pp. 61–68.
2. **Jain A. K., Nandakumar K., Nagar A.** Biometric template security. *EURASIP J. Adv. Signal Process*, 2008, pp. 113:1–113:17.
3. **Hafemann L. G., Sabourin R., Oliveira L. S.** Characterizing and evaluating adversarial examples for offline handwritten signature verification. *IEEE Transactions on Information Forensics and Security*, 2019, vol. 14, iss. 8, pp. 2153–2166. doi:10.1109/TIFS.2019.2894031
4. **Ivanov A. I., Lozhnikov P. S., Sulavko A. E.** Evaluation of signature verification reliability based on artificial neural networks, Bayesian multivariate functional and quadratic forms. *Computer Optics*, 2017, no. 5, pp. 765–774. doi:10.18287/2412-6179-2017-41-5-765-774
5. ISO/IEC 19792:2009. Information technology — Security techniques — Security evaluation of biometrics. International Organization for Standardization, 2009. 37 p.
6. ISO/IEC 24761:2009. Information technology — Security techniques — Authentication context for biometrics. International Organization for Standardization, 2011. 50 p.
7. ISO/IEC 24745:2011. Information technology — Security techniques — Biometric information protection. International Organization for Standardization, 2011. 50 p.
8. ГОСТ Р 52633.0-2006. Защита информации. Техника защиты информации. Требования к средствам высоконадежной биометрической аутентификации. М., Стандартинформ, 2007. 25 с.
9. ГОСТ Р 52633.1-2009. Защита информации. Техника защиты информации. Требования к формированию баз естественных биометрических образов, предназначенных для тестирования средств высоконадежной биометрической аутентификации. М., Стандартинформ, 2010. 24 с.
10. ГОСТ Р 52633.2-2010. Защита информации. Техника защиты информации. Требования к формированию синтетических биометрических образов, предназначенных для тестирования средств высоконадежной биометрической аутентификации. М., Стандартинформ, 2011. 22 с.
11. ГОСТ Р 52633.3-2011. Защита информации. Техника защиты информации. Тестирование стойкости средств высоконадежной биометрической защиты к атакам подбора. М., Стандартинформ, 2012. 16 с.
12. ГОСТ Р 52633.4-2011. Защита информации. Техника защиты информации. Интерфейсы взаимодействия с нейросетевыми преобразователями биометрия-код. М., Стандартинформ, 2012. 46 с.
13. ГОСТ Р 52633.5-2011. Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа. М., Стандартинформ, 2012. 20 с.

14. ГОСТ Р 52633.6-2012. Защита информации. Техника защиты информации. Требования к индикации близости предъявленных биометрических данных образу «Свой». М., Стандартинформ, 2012. 24 с.
15. **Dodis Y., Reyzin L., Smith A.** Fuzzy extractors: How to generate strong keys from biometrics and other noisy. *EuroCrypt*, 2004, pp. 523–540.
16. **Иванов А. И., Сомкин С. А., Андреев Д. Ю., Малыгина Е. А.** О многообразии метрик, позволяющих наблюдать реальные статистики распределения биометрических данных «нечетких экстракторов» при их защите наложением гаммы. *Вестник УрФО. Безопасность в информационной сфере*, 2014, № 2(12), с. 16–23.
17. **Ignatenko T., Frans M. J.** Willems. Information leakage in fuzzy commitment schemes. *IEEE Transactions on Information Forensics and Security*, 2010, vol. 5, no. 2, pp. 337–348. doi:10.1109/TIFS.2010.2046984
18. **Ложников П. С., Сулавко А. Е., Еременко А. В., Волков Д. А.** Экспериментальная оценка надежности верификации подписи сетями квадратичных форм, нечеткими экстракторами и персептронами. *Информационно-управляющие системы*, 2016, № 5, с. 73–85. doi:10.15217/issn1684-8853.2016.5.73
19. **Ахметов Б. С., Иванов А. И., Фунтиков В. А., Безяев А. В., Малыгина Е. А.** *Технология использования больших нейронных сетей для преобразования нечетких биометрических данных в код ключа доступа*. Алматы, LEM, 2014. 144 с.
20. **Kůrková V., Sanguineti M.** Probabilistic lower bounds for approximation by shallow perceptron networks. *Neural Networks*, 2017, vol. 91, pp. 34–41.
21. **Kůrková V., Sanguineti M.** Model complexities of shallow networks representing highly varying functions. *Neurocomputing*, 2016, vol. 171, pp. 598–604.
22. **Иванов А. И.** Нейросетевая защита конфиденциальных биометрических образов гражданина и его личных криптографических ключей. Пенза, ПНИЭИ, 2014. 57 с.
23. **Iranmanesh V., Ahmad S. M. S., Adnan W. A. W., Yussof S., Arigbabu O. A., Malallah F. L.** Online handwritten signature verification using neural network classifier based on principal component analysis. *Scientific World Journal*, 2014, vol. 2014, pp. 1–8.
24. **Iranmanesh V.** Online signature template protection by shuffling and one time pad schemes with neural network verification. *Proceedings of the International Conference on Computer Science and Computational Mathematics (ICCSCM '13)*, 2013, pp. 53–59.
25. **Hafemann L. G., Sabourin R., Oliveira L. S.** Writer-independent feature learning for offline signature verification using deep convolutional neural networks. *2016 International Joint Conference on Neural Networks (IJCNN)*, 2016. doi:10.1109/IJCNN.2016. 7727521
26. **Souza V. L. F., Oliveira A. L. I., Sabourin R.** A writer-independent approach for offline signature verification using deep convolutional neural networks features. *2018 7th Brazilian Conference on Intelligent Systems (BRACIS)*, 2018. doi:10.1109/BRACIS.2018. 00044
27. **Díaz M., Fischer A., Ferrer M. A., Plamondon R.** A perspective analysis of handwritten signature technology. *ACM Computing Surveys*, 2019, vol. 51, iss. 6, article 117, pp. 1–37.
28. **Maiorana E., Campisi P.** Fuzzy commitment for function based signature template protection. *IEEE Signal Processing Letters*, 2010, vol. 17, pp. 249–252.
29. **Malygin A., Seilova N., Boskebeev K., Alimseitova Zh.** Application of artificial neural networks forhandwritten biometric images recognition. *Computer Modelling and New Technologies*, 2017, vol. 21(1), pp. 31–38.
30. ГОСТ Р ИСО/МЭК ТО 19795-3-2009. Автоматическая идентификация. Идентификация биометрическая. Эксплуатационные испытания и протоколы испытаний в биометрии. Часть 3. Особенности проведения испытаний при различных биометрических модальностях. М., Стандартинформ, 2010. 28 с.
31. **Sulavko A. E., Zhumazhanova S. S., Fofanov G. A.** Perspective neural network algorithms for dynamic biometric pattern recognition in the space of interdependent features. *Proceedings of 2018th Conference "Dynamics of Systems, Mechanisms and Machines"*, Omsk, 2018, pp. 1–12.

UDC 004.93'1
doi:10.31799/1684-8853-2020-4-61-77

Highly reliable authentication based on handwritten passwords using hybrid neural networks with protection of biometric templates from being compromised

A. E. Sulavko[a], PhD, Tech., Associate Professor, orcid.org/0000-0002-9029-8028, sulavich@mail.ru
[a]Omsk State Technical University, 11, Mira Pr., 644050, Omsk, Russian Federation

Introduction: Biometrics-to-code converters based on neural networks are the ideological basis for a series of GOST R 52633 standards (unparalleled anywhere in the world) and can be used in the development of highly reliable biometric authentication and electronic signature with biometric activation. **Purpose:** Developing a model of a biometrics-to-code converter for highly reliable biometric authentication by handwritten passwords with high resistance to attacks on knowledge extraction. **Results:** We demonstrated the vulnerability of neural networks which makes it possible to perform quick directed enumeration of competing examples in order to compromise a biometric pattern and the personal key of its owner. We described a method of effective protection against this attack, and proposed a hybrid model for a biometrics-to-code converter based on a new type of hybrid neural networks, which does not compromise

the biometric standard and the user's key (password), being resistant to such attacks. The high reliability and effectiveness of the proposed model has been experimentally confirmed in handwritten password verification. The reliability indicators for generating a key from a handwritten password were: FRR = 11.5%, FAR = 0.0009% with a key length of 1024 bits (taking into account the presented fakes of a handwritten pattern). **Practical relevance:** The results can be used in information security applications or electronic document management.

Keywords — pattern recognition, Bayesian differential measures, correlated biometric features, information protection, quick tuning of neural networks, probability density, "wide" neural networks, biometrics-to-code converters, handwritten patterns.

For citation: Sulavko A. E. Highly reliable authentication based on handwritten passwords using hybrid neural networks with protection of biometric templates from being compromised. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 4, pp. 61–77 (In Russian). doi:10.31799/1684-8853-2020-4-61-77

References

1. Akhmetov B. S., Ivanov A. I., Alimseitova Z. K. Training of neural network biometry-code converters. *2018 News of the National Academy of Sciences of the Republic of Kazakhstan. Series of Geology and Technical Sciences*, 2018, vol. 1, no. 427, pp. 61–68.
2. Jain A. K., Nandakumar K., Nagar A. Biometric template security. *EURASIP J. Adv. Signal Process*, 2008, pp. 113:1–113:17.
3. Hafemann L. G., Sabourin R., Oliveira L. S. Characterizing and evaluating adversarial examples for offline handwritten signature verification. *IEEE Transactions on Information Forensics and Security*, 2019, vol. 14, iss. 8, pp. 2153–2166. doi:10.1109/TIFS.2019.2894031
4. Ivanov A. I., Lozhnikov P. S., Sulavko A. E. Evaluation of signature verification reliability based on artificial neural networks, Bayesian multivariate functional and quadratic forms. *Computer Optics*, 2017, no. 5, pp. 765–774. doi:10.18287/2412-6179-2017-41-5-765-774
5. ISO/IEC 19792:2009. Information technology — Security techniques — Security evaluation of biometrics. International Organization for Standardization, 2009. 37 p.
6. ISO/IEC 24761:2009. Information technology — Security techniques — Authentication context for biometrics. International Organization for Standardization, 2011. 50 p.
7. ISO/IEC 24745:2011. Information technology — Security techniques — Biometric information protection. International Organization for Standardization, 2011. 50 p.
8. State Standard 52633.0-2006. Data protection. Information security technique. High Reliability Biometric Authentication Requirements. Moscow, Standardinform Publ., 2007. 25 p. (In Russian).
9. State Standard 52633.1-2009. Data protection. Information security technique. Requirements for the formation of databases of natural biometric images intended for testing highly reliable biometric authentication. Moscow, Standardinform Publ., 2010. 24 p. (In Russian).
10. State Standard 52633.2-2010. Data protection. Information security technique. Requirements for the formation of synthetic biometric images intended for testing highly reliable biometric authentication tools. Moscow, Standardinform Publ., 2011. 22 p. (In Russian).
11. State Standard 52633.3-2011. Data protection. Information security technique. Testing the resistance of highly reliable biometric protection to selection attacks. Moscow, Standardinform Publ., 2012. 16 p. (In Russian).
12. State Standard 52633.4-2011. Data protection. Information security technique. Interfaces for interaction with neural network biometrics to code converters. Moscow, Standardinform Publ., 2012. 46 p. (In Russian).
13. State Standard 52633.5-2011. Data protection. Information security technique. Automatic training of neural network biometrics to code converters. Moscow, Standardinform Publ., 2012. 20 p. (In Russian).
14. State Standard 52633.6-2012. Data protection. Information security technique. Requirements for indicating the proximity of biometric data presented to the image of "Own". Moscow, Standardinform Publ., 2012. 24 p. (In Russian).
15. Dodis Y., Reyzin L., Smith A. Fuzzy extractors: How to generate strong keys from biometrics and other noisy. *EuroCrypt*, 2004, pp. 523–540.
16. Ivanov A., Somkin S., Andreev D., Malygina E. Diversity metrics to watch actual biometric data distribution statistics "fuzzy extractors" in their protection of a range. *UrFR Newsletter. Information Security*, 2014, no. 2(12), pp. 16–23 (In Russian).
17. Ignatenko T., Frans M. J. Willems. Information leakage in fuzzy commitment schemes. *IEEE Transactions on Information Forensics and Security*, 2010, vol. 5, no. 2, pp. 337–348. doi:10.1109/TIFS.2010.2046984
18. Lozhnikov P. S., Sulavko A. E., Eremenko A. V., Volkov D. A. Experimental evaluation of reliability of signature verification by quadratic form networks, fuzzy extractors and perceptrons. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2016, no. 5, pp. 73–85 (In Russian). doi:10.15217/issn1684-8853.2016.5.73
19. Ahmetov B. S., Ivanov A. I., Funtikov V. A., Bezjaev A. V., Malygina E. A. *Tekhnologiia ispol'zovaniia bol'shikh neironnykh setei dlia preobrazovaniia nechetkikh biometricheskikh dannykh v kod kliucha dostupa* [Technology of using large neural networks for fuzzy transformation of biometric data in the access code key]. Almaty, LEM Publ., 2014. 144 p. (In Russian).
20. Kůrková V., Sanguineti M. Probabilistic lower bounds for approximation by shallow perceptron networks. *Neural Networks*, 2017, vol. 91, pp. 34–41.
21. Kůrková V., Sanguineti M. Model complexities of shallow networks representing highly varying functions. *Neurocomputing*, 2016, vol. 171, pp. 598–604.
22. Ivanov A. I. *Neirosetevaia zashchita konfidentsial'nykh biometricheskikh obrazov grazhdanina i ego lichnykh kriptograficheskikh kliuchei* [Neural protection of sensitive biometric images of the citizen and his personal cryptographic keys]. Penza, PNIEI Publ., 2014. 57 p. (In Russian).
23. Iranmanesh V., Ahmad S. M. S., Adnan W. A. W., Yussof S., Arigbabu O. A., Malallah F. L. Online handwritten signature verification using neural network classifier based on principal component analysis. *Scientific World Journal*, 2014, vol. 2014, pp. 1–8.
24. Iranmanesh V. Online signature template protection by shuffling and one time pad schemes with neural network verification. *Proceedings of the International Conference on Computer Science and Computational Mathematics (ICCSCM '13)*, 2013, pp. 53–59.
25. Hafemann L. G., Sabourin R., Oliveira L. S. Writer-independent feature learning for offline signature verification using deep convolutional neural networks. *2016 International Joint Conference on Neural Networks (IJCNN)*, 2016. doi:10.1109/IJCNN.2016.7727521
26. Souza V. L. F., Oliveira A. L. I., Sabourin R. A writer-independent approach for offline signature verification using deep convolutional neural networks features. *2018 7th Brazilian Conference on Intelligent Systems (BRACIS)*, 2018. doi:10.1109/BRACIS.2018.00044
27. Díaz M., Fischer A., Ferrer M. A., Plamondon R. A perspective analysis of handwritten signature technology. *ACM Computing Surveys*, 2019, vol. 51, iss. 6, article 117, pp. 1–37.
28. Maiorana E., Campisi P. Fuzzy commitment for function based signature template protection. *IEEE Signal Processing Letters*, 2010, vol. 17, pp. 249–252.
29. Malygin A., Seilova N., Boskebeev K., Alimseitova Zh. Application of artificial neural networks forhandwritten biometric images recognition. *Computer Modelling and New Technologies*, 2017, vol. 21(1), pp. 31–38.
30. ISO/IEC TR 19795-3:2007. Information technology — Biometric performance testing and reporting — Part 3: Modality-specific testing. International Organization for Standardization Publ., 2007. 19 p. (In Russian).
31. Sulavko A. E., Zhumazhanova S. S., Fofanov G. A. Perspective neural network algorithms for dynamic biometric pattern recognition in the space of interdependent features. *Proceedings of 2018th Conference "Dynamics of Systems, Mechanisms and Machines"*, Omsk, 2018, pp. 1–12.