# Высоконадёжная двухфакторная биометрическая аутентификация по рукописным и голосовым паролям на основе гибких нейронных сетей

*А.Е. Сулавко* [1]
[1] *ФГБОУ ВО «Омский государственный технический университет» (ОмГТУ), 644050, г. Омск, проспект Мира, д. 11*

***Аннотация***

В работе рассматривается проблема высоконадежной биометрической аутентификации на основе преобразователей тайных биометрических образов в длинный ключ или пароль, а также их тестирования на сравнительно малых выборках (тысячи образов). Статические образы являются открытыми, поэтому при удаленной аутентификации доверие к ним ограничено. Описан процесс вычисления биометрических параметров голосового и рукописного паролей, предложен метод автоматического формирования гибкой гибридной сети, состоящей из нейронов различного типа, и абсолютно устойчивый алгоритм ее обучения на малых выборках «Свой» (7–15 примеров). Предложен метод обученного гибридного преобразователя «биометрия-код» от извлечения знаний. Достигнуты низкие показатели FAR.

*Ключевые слова*: гибридные сети, квадратичные формы, функционалы Байеса, особенности воспроизведения рукописных образов, параметры голоса, широкие нейронные сети, преобразователи «биометрия-код», защищенные нейросетевые контейнеры.

*Цитирование*: **Сулавко, А.Е.** Высоконадежная двухфакторная биометрическая аутентификация по рукописным и голосовым паролям на основе гибких нейронных сетей // Компьютерная оптика. – 2020. – Т. 44, № 1. – С. 82-91. – DOI: 10.18287/2412-6179-CO-567.

*Citation*: Sulavko AE. Highly reliable two-factor biometric authentication based on handwritten and voice passwords using flexible neural networks. Computer Optics 2020; 44(1): 82-91. DOI: 10.18287/2412-6179-CO-567.

## *Введение*

Мы живём в эру информационных технологий, когда обостряются проблемы информационной безопасности. Пароли и криптографические ключи являются отчуждаемыми от владельца и поэтому подвержены «человеческому фактору». Длинный ключ (пароль) является надёжным, только если были соблюдены все правила при его генерации – энтропия ключа должна быть сопоставима с его длиной. Но случайный длинный пароль невозможно запомнить. Выход из ситуации – «привязка» ключей и паролей субъекта к его биометрическим параметрам с помощью преобразователя «биометрия–код» (ПБК) [1], который настраивается на выдачу ключа пользователя при предъявлении его биометрического образа («Свой»). При предъявлении неизвестного образа («Чужой») ПБК должен формировать случайный бинарный код, близкий по энтропии к «белому шуму» (рис. 1). Сами пароли и ключи генерируются до обучения ПБК в соответствии с принятыми нормами.

Статические биометрические образы (отпечаток пальца, радужка и другие) являются достаточно уникальными для высокоточной идентификации человека. Однако они уязвимы перед атаками представления (спуфинг), так как их невозможно держать в секрете. Открытый образ может быть изучен злоумышленником и подделан (можно снять данные со стакана, ручки двери, фотографии и т.д.). Процесс ввода биометрического образа можно непосредственно контролировать только по месту прохождения аутентификации (эти вопросы рассматриваются в серии из 3 стандартов ISO/IEC 30107). Для удалённой аутентификации аналогичные методы контроля не действуют – биометрические данные могут быть синтезированы искусственно. Таким образом, при удалённой аутентификации открытую биометрическую характеристику стоит рассматривать не в качестве подтверждающей информации, а только как идентификатор.

К открытым образам также относится рукописная подпись (автограф). Вероятность ложного принятия подделки подписи в десятки раз выше, чем случайного совпадения [2, 3]. Текстонезависимое распознавание диктора тоже строится на открытых образах. Последние достижения в области синтеза речи показали, что можно сгенерировать речевой сигнал конкретного человека произвольного содержания, почти идентичный натуральному голосу, предварительно обучив нейросетевой синтезатор на нескольких часах голосовой информации (со стенограммой) за ночь на обычном компьютере [4]. Созданы архитектуры синтезаторов речи (на свёрточных или рекуррентных сетях) – WaveNet, Char2Wav, DeepVoice, Tacotron.

Поэтому для аутентификации нужно использовать тайный образ, отражающий особенности воспроизведения пароля его владельцем. У злоумышленника не должно быть информации о том, какой биометрический пароль синтезировать. Настоящее исследование посвящено разработке ПБК, ориентированного на об-

работку относительно слабых биометрических данных рукописного и голосового паролей.

### *Об оценке надёжности ПБК*

К уникальности паролей и криптографических ключей предъявляются серьёзные требования. В соответствии с ГОСТ Р 34.10-2012 длина ключа электронной подписи должна быть не менее 256 бит, вероятность его случайного совпадения менее $10^{-77}$. К биометрическим системам предъявляются заниженные требования – $FAR \leq 10^{-9}$ считается хорошим показателем. Но для защиты от атак грубого и направленного перебора, реконструкции шаблона [5], генерации синтетических [4] или конкурирующих примеров [6] (когда за основу берётся естественный образ и видоизменяется путём наложения шума и т.д.) этот показатель недостаточен. Сгенерировать и перебрать $10^9$ независимых биометрических образов можно даже на штатном компьютере [7]. Это дольше перебора $10^9$ паролей, но тем не менее возможно. По этой причине вероятность «ложного допуска» должна быть сопоставима с вероятностью совпадения случайного пароля средней длины (56–64 бит). Чтобы достигнуть или хотя бы приблизиться к такой вероятности, требуется комплексирование нескольких видов тайных биометрических образов (голоса, клавиатурного и рукописного почерка), один фактор даёт слишком высокие показатели FRR и FAR.

С повышением надёжности биометрической аутентификации меняются требования к тестированию. Серия стандартов ISO/IEC 19795 не предназначена для проверки экстремально низких вероятностей, так как основана на традиционном подходе к определению FAR через отношение числа ошибок к количеству проведённых опытов (с последующим определением доверительных вероятности и интервала). Но для оценки очень низких вероятностей прямым численным экспериментом не хватит ресурсов. Например, для доказательства $FAR \approx 1{,}38 \cdot 10^{-17}$ (совпадения 56-битного пароля) сбор биометрических данных займёт более 3 лет, если привлечь все население планеты (при времени воспроизведения одного образа 10 секунд, без отдыха и сна, при этом пароли не должны повторяться). Это невозможно даже в теории. Поэтому для оценки стойкости ПБК к атакам подбора в ГОСТ 52633.3 предлагаются специальные методы. Согласно стандарту в эксперименте могут использоваться не только естественные образы «Чужой», но и синтетические, которые генерируются на основе скрещивания естественных. Скрещиваются пары из 1 % «Чужих», которые по результатам тестирования оказались наиболее близкими к образу «Свой». Далее по аналогичному принципу скрещиваются уже синтетические образы. Каждая новая популяция всё ближе к образу «Свой». Процесс повторяется, пока образ «Свой» не будет скомпрометирован либо «сближение» образов не прекратится. Тем самым сокращается количество попыток предъявления конкурирующих примеров и увеличивается точность оценки FAR при сохранении статистической значимости. FAR может быть вычислен как соотношение числа ошибок к числу опытов (принимая во внимание, что если в 1 % близких «Чужих» ошибок не произошло, то в 99 % более далёких тоже). Для получения итоговой оценки FAR весь процесс тестирования нужно «прогнать» через множество ПБК, обученных на данных разных испытуемых [8]. Для начала испытаний необходимо иметь исходную тестовую базу (*нулевую популяцию*) независимых образов «Чужих» (не вошедших в обучающую выборку), численностью в несколько тысяч.

Для получения средней оценки FRR с точностью до десятой процента вполне достаточно $10^4$ опытов (суммарно) [8]. Однако для каждого испытуемого число опытов должно быть идентичным.

Распространённой практикой является оценка равной вероятности ошибок (EER). Однако EER мало говорит о стойкости защиты к попыткам взлома. Реальная стойкость определяется показателем FAR. Нужно одновременно обеспечить *высокую конфиденциальность* ($10^{-20} \leq FAR \leq 10^{-17}$) при *приемлемой доступности* ($FRR \leq 0{,}25$) [8].

### *Подходы к построению ПБК*

Важной задачей является защита биометрических образов, ключей и паролей от компрометации при хранении. Формат данных обученного ПБК должен быть таким, чтобы даже без применения сторонних средств шифрования биометрический эталон пользователя и его личный ключ (пароль) были скрыты от непосредственного наблюдения и надёжно защищены от восстановления. Хакеры не должны иметь возможность извлечения знаний из обученного ПБК (рис. 1). Таковы требования современных стандартов по информационной безопасности.

На текущий момент сложилось два основных подхода к построению ПБК [1]: на основе «нечёткого экстрактора» (этот подход лёг в основу стандартов ISO/IEC 19792:2009, 24761:2009 и 24745:2011) и нейросетевой подход (поддерживается серией из шести стандартов ГОСТ Р 52633). Нечёткие экстракторы имеют принципиальные недостатки, этот подход может быть востребован при обработке статических образов, для динамических образов уровень ошибок оказывается неприемлемо высоким [9].

Нейросетевые ПБК (НПБК) позволяют достичь более низких показателей FRR и FAR, не налагая ограничения на длину ключа [9]. НПБК строится персонально для каждого субъекта, при этом формируется искусственная нейронная сеть (ИНС), количество входов которой равно числу признаков (биометрических параметров), а количество выходов – длине личного ключа. Каждый нейрон последнего слоя генерирует один бит. Нейронная сеть обучается на примерах

образа пользователя и примерах, не принадлежащих пользователю, чтобы синтезировать ключ при поступлении образа «Свой» (рис. 1).

Обучение НПБК должно быть абсолютно устойчивым. Объём обучающей выборки «Чужие» может быть сколь угодно большим. Разработчик биометрической системы может подготовить репрезентативную выборку «Чужие» и использовать её для обучения каждого НПБК. Но число примеров образа «Свой» должно быть малым (10–20), нельзя заставлять пользователя сотни раз вводить образ. Это обстоятельство накладывает ограничения на возможную архитектуру ИНС, лежащую в основе НПБК.

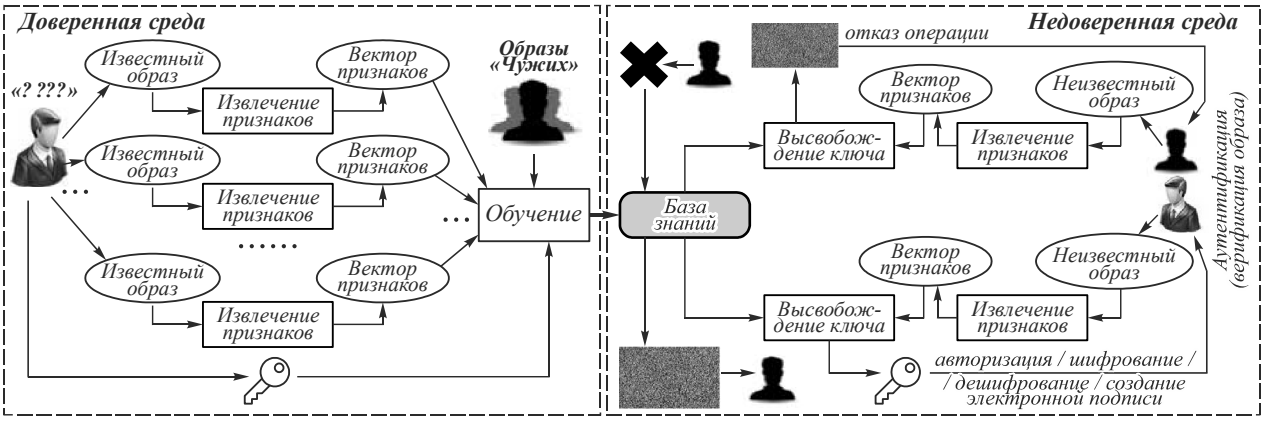



*Рис. 1. Общая схема преобразователя биометрия-код*

В частности, для построения НПБК затруднительно использовать многослойные сетевые архитектуры, особенно персептроны, обучаемые по принципу обратного распространения ошибки и имеющие обычно один выход (при классификации двух образов – «Свой» и «Чужой») [1]. Свёрточные нейронные сети в целом дают более показательные результаты, но при анализе слабых биометрических образов вероятности ошибок оказываются значительными (особенно для голосовых образов) [10]. Дело в том, что итерационные алгоритмы настройки весов нейронов неустойчивы на малых обучающих выборках [1] (методы оптимизации сети не делают алгоритм обучения абсолютно устойчивым, что требуется в биометрии). Попытки настроить ПБК на примерах подписи итерационным алгоритмом не увенчались успехом [11].

Чем выше информативность биометрического образа (т.е. уникальность и стабильность признаков), тем меньший объём выборки нужен и тем более стабильной оказывается итерационная процедура обучения. Например, для настройки свёрточных нейронных сетей в задаче биометрической идентификации личности по лицу [5] достаточно пяти примеров изображения лица от каждого идентифицируемого субъекта (в различных ракурсах). Однако образ лица гораздо информативней, чем голосовой или рукописный образ [9]. Поэтому при распознавании диктора или подписи процедуры итерационного обучения менее стабильны. Результат верификации подписи сильно варьируются для разных подписантов при обучающих выборках в 15–30 примеров на человека [2]. При верификации диктора увеличение числа обучающих примеров может вызвать повышение EER [10] (что нелогично и как раз говорит о неустойчивости обучения). Таких примеров можно привести множество.

Недостатки «глубоких» нейронных сетей [1], наряду с невозможностью применения для высоконадежной (при $\mathrm{FAR} \le 10^{-12}$) биометрической аутентификации «нечётких экстракторов», привели к появлению серии стандартов ГОСТ Р 52633, в которых рассматриваются только однослойные и двухслойные ИНС, обучаемые без «обратного распространения ошибки». Все нейроны настраиваются независимо друг от друга, исходя из параметров распределения признаков, и имеют пороговые функции активации, что позволяет генерировать код ключа [1]. Данную категорию ИНС принято называть «*широкими*» сетями [1] (они имеют много нейронов, но мало слоёв). Алгоритм обучения ГОСТ Р 52633.5 всегда остаётся устойчивым. Добавление нейронов, увеличение числа входов и выходов сети не приводит к усложнению процесса обучения, а ведёт к снижению количества ошибок (до определённого предела) [12], при этом не требуется изменять гиперпараметры ИНС или алгоритма обучения.

### *Пространство признаков*

После оцифровки биометрический образ преобразуется в вектор признаков $\bar{a} = \{a_1, ..., a_N\}$ фиксированной длины. Оцифрованный рукописный образ состоит из функций координат $x(t)$, $y(t)$ и давления (силы нажатия) пера на планшет $p(t)$, где $t$ – это время в дискретной форме. Для устройств начального уровня поддержки давления нет (только индикация касания). В ISO/IEC 19794-7 предусмотрена возможность учёта угла наклона и азимута пера. Эти данные регистрируются профессиональными устройствами, которые не имеют широкого распространения. Поэтому в работе рассмотрено 2 набора признаков: для устройств с поддержкой давления ($N_{HXYP} = 782$) и без ($N_{HXY} = 521$).

В настоящей работе не преследуется целей выбора «наилучших» характеристик. Широкие ИНС позволяют брать как можно больше признаков и не рассматривать разные методы их извлечения как альтернативные. Каждый признак может «отсеять» определённого «Чужого» [8]. Поэтому мы использовали различные подходы совместно (табл. 1).

Голосовой пароль представлен звуковым сигналом, записанным в файл формата wav (без сжатия). Мнения относительно параметров оцифровки сигнала (с сохранением информации о дикторе) в различных работах расходятся. Наиболее информативные признаки диктора можно получить путём оценки индивидуальной фундаментальной частоты его голоса (частоты основного тона, ЧОТ), которая обусловлена строением гортани и проявляется при произношении гласных фонем. У мужчин ЧОТ варьируется от 80 до 150 Гц, у женщин – от 120 до 400 Гц [13]. Однако информацию о дикторе несут и более высокочастотные составляющие (обертоны). Частоты, отвечающие за разборчивость речи, сконцентрированы в основном в диапазоне 300 и 3400 Гц. Поэтому считается, что речь человека при $F = 8$ кГц различима. В [13] сообщается, что при вейвлет-анализе речи удаётся получить хорошее частотно-временное разрешение при $F = 8$ кГц. Такая частота характерна для низко информативных каналов передачи данных, таких как телефонная сеть. Тем не менее обычно используется частота дискретизации $F = 16$ кГц [14] или выше [4, 8] (при увеличении $F$ уменьшаются ошибки квантования). Касательно уровней квантования существенных расхождений нет – размер семпла в 16 бит считается достаточным для «машинного восприятия» [4, 8, 13].

Не все устройства могут обеспечить качественную оцифровку звука. Встроенные в мобильные гаджеты средства (низкого ценового сегмента) не всегда способны охватить даже диапазон частот 80–3400 Гц. Для более качественных «голосовых» микрофонов среднего уровня реальный диапазон частот записи обычно составляет от 70 Гц до 12 кГц. По теореме Котельникова достаточно ограничиться $F = 24$ кГц. На практике важно, чтобы технология была доступна широкому кругу лиц. Поэтому решено апробировать 2 набора признаков (табл. 1): для частоты 24 кГц ($N_{f24} = 570$) и 8 кГц ($N_{f8} = 350$).

*Табл. 1. Признаки, используемые в настоящем и более ранних исследованиях [1, 9, 16, 17]*

| № | Краткое описание группы признаков (1 – рукописный, 2 – голосовой пароль) | Число признаков |
|---|---|---|
| 1.1 | Образ делится на 16 равных по числу точек отрезков, строится матрица расстояний между их краями в 2- и 3-мерном пространстве ($p(t)$ – третье измерение) | 240<br>120 – без $p(t)$ |
| 1.2 | Коэффициенты корреляции между $x(t)$, $y(t)$, $p(t)$, $x'(t)$, $y'(t)$, $p'(t)$ и функцией скорости пера $v_{xy}(t)$, производной от $x(t)$, $y(t)$ | 21<br>10 – без $p(t)$ |
| 1.3 | Параметры внешнего вида образа: угол наклона, отношение длины к ширине, центр в трех-(двух-)мерном пространстве, описываемый тремя (двумя) координатами | 5<br>4 – без $p(t)$ |
| 1.4 | Средние значения фрагментов функций $p(t)$, $x'(t)$, $y'(t)$, $v_{xy}(t)$ (образ делится на 5 равных по числу точек отрезков) | 20<br>15 – без $p(t)$ |
| 1.5 | Детализирующие коэффициенты быстрого вейвлет-преобразования Хаара (алгоритм Малла), полученные на 4 уровнях разложения (низкие частоты) для $x(t)$, $y(t)$, $p(t)$, $v_{xy}(t)$ (функции сначала приводились к 128 отчетам (интерполяция)) | 240<br>180 – без $p(t)$ |
| 1.6 | Усредненный амплитудный спектр, полученный с помощью STFT (размер окна – 128 отчетов, шаг – 16 отчетов) для $x(t)$, $y(t)$, $p(t)$, $v_{xy}(t)$ | 256<br>192 – без $p(t)$ |
| 2.1 | Часть усредненного по всем окнам амплитудного спектра речевого сигнала (коэффициенты нижних частот), вычисленного с помощью STFT (размер окна – 2048(512) при $F$=24 (8) кГц, шаг – 16). Предварительно речевой сигнал нормируется по энергии, удаляется тишина | 40 – $F$=24 кГц<br>20 – $F$=8 кГц |
| 2.2 | Часть кепстра (коэффициенты нижних частот), который берется от полного усредненного по всем окнам амплитудного спектра, получаемого в соответствии с 2.1 | 40 – $F$=24 кГц<br>20 – $F$=8 кГц |
| 2.3 | Часть кепстра (коэффициенты нижних частот), который берется от полного логарифмированного усредненного амплитудного спектра, получаемого в соответствии с 2.1 | 40 – $F$=24 кГц<br>20 – $F$=8 кГц |
| 2.4 | Кепстры второго порядка от полных кепстров 2.1 и 2.2 (кепстры 2.1 и 2.2 повторно подвергаются прямому преобразованию Фурье) | 256 – $F$=24 кГц<br>128 – $F$=8 кГц |
| 2.5 | Подсчет частоты переходов сигнала через нулевое деление окном (размер окна – 2048(512) при $F$=24 (8) кГц, шаг – 16), грубо характеризуют ЧОТ (нулевую форманту) | 64 – $F$=24 кГц<br>32 – $F$=8 кГц |
| 2.6 | Полный амплитудный спектр функции автокорреляции речевого сигнала | 128 |
| 2.7 | Частота переходов через «ноль» и частота экстремумов функции автокорреляции | 2 |

ЧОТ вариативна. Для её оценки применяются методы вейвлет-анализа [13], анализа периодичности автокорреляционных функций [15], частоты экстремумов сигнала (переходов через «ноль» [16]). Также для поиска признаков анализируют спектры и кепстры сигнала. Для этого принято использовать кратковременное быстрое преобразование Фурье (STFT). Распространен подход с расчётом мел-кепстральных

коэффициентов (MFCC) [14] или логарифмических энергий (MFEC) [10]. Вейвлет-анализ редко применяется в задачах биометрической аутентификации из-за низкой скорости работы. Быстрый вейвлет-анализ (разложение алгоритмом Малла по разным базисам) не даёт видимых преимуществ при поиске информативных признаков диктора [17]. В настоящей работе использована компиляция нескольких подходов к вычислению признаков голоса (табл. 1).

В стандартах ISO/IEC 29794-1 и ГОСТ Р 52633 представлены разные подходы к оценке качества (уникальности, стабильности) образов и признаков. В [18] предложен альтернативный подход (и шкала) оценки информативности признака $I_j$ через определение площади пересечения функций плотности вероятности его значений для образов «Свой» и «Чужой». Признаки из табл. 1 условно можно описать нормальным законом распределения, что проверялось критерием хи-квадрат (у некоторых испытуемых отдельные признаки имеют распределение Лапласа или логнормальное, но эти отклонения не оказали существенного влияния на конечный результат). Информативность отдельно взятого признака можно считать аналогом его уникальности. Но на общее количество информации образа влияют корреляционные связи признаков. Часть информации «переходит» в матрицу коэффициентов корреляции между признаками [1], которая почти уникальна у каждого тайного образа.

### *Гибкие гибридные сети в основе ПБК*

Оставаясь в рамках концепции широких сетей, не обязательно придерживаться традиционной схемы построения нейрона [1]. Исследовано несколько классов функционалов (мер близости), которые способны более эффективно обогащать слабые биометрические данные, чем функционал взвешенного суммирования (1) классического нейрона. Каждый из них можно рассматривать как основу нейрона с пороговой функцией активации (заменив сумматор (1) на соответствующий функционал). Однако входы такого нейрона стоит выбирать не просто случайно, а исходя из свойств функционала, лежащего в его основе. В работе использовались следующие меры близости для создания уникальных нейронов гибридной сети:

- многомерные функционалы Байеса (разностные [1], гиперболические (2) [19]) обладают способностью к высокоэффективной обработке сильно зависимых признаков (чем выше корреляционная зависимость, тем меньше ошибок распознавания) [1, 19]. Для нейрона Байеса важно выбирать признаки так, чтобы коэффициенты парной корреляции $r_{j,t}$ между ними были как можно выше ($r_{j,t}>0{,}5$) и близки по значению. Чем больше таких признаков, тем эффективнее работает нейрон;
- квадратичные функционалы (меры Пирсона (3), хи-модуль), обрабатывающие только слабо зависимые сочетания признаков ($r_{j,t}<0{,}3$), дают меньший процент ошибок при меньшей размерности, чем (1) [9];
- «гравитационные» метрики (в работе [20] предложено 6 мер) являются своего рода противоположностью квадратичным функционалам, так как работают с сильно зависимыми признаками ($r_{j,t}<0{,}5$);
- функционал наибольшего правдоподобия Байеса (4) [16] подходит для обработки независимых ($r_{j,t}<0{,}3$) и зависимых признаков ($r_{j,t}>0{,}3$). Для «наивного» Байесовского классификатора не столько важна сила взаимной корреляционной зависимости признаков, сколько равенство парных коэффициентов корреляции между ними. Это показали результаты исследований [9]. Нейроны наибольшего правдоподобия при $r_{j,t}>0{,}3$ настраиваются аналогично многомерным Байесовским функционалам, при $r_{j,t}<0{,}3$ – аналогично квадратичным формам;
- меры близости, основанные на критериях проверки гипотез о законе распределения случайной величины [18] (сравнения двух распределений). В данной работе применялось 7 критериев, показавших наилучшие результаты при обработке слабо зависимых биометрических признаков ($r_{j,t}<0{,}3$): Крамера–фон Мизеса (интегральный), Джини (интегральный и дифференциальный), максимума пересечения функций плотности вероятности (ПВ), среднего геометрического ПВ, Ватсона (интегральный), Колмогорова–Смирнова (дифференциальный);
- предложено преобразование [18], позволяющее адаптировать многие критерии к обработке зависимых признаков ($r_{j,t}>0{,}3$), в данной работе для создания нейронов использовалось 9 таких критериев: Крамера–фон Мизеса, Смирнова–Крамера–фон Мизеса, Джини (интегральные и дифференциальные), Купера, Колмогорова–Смирнова (дифференциальные), максимума пересечения функций ПВ.

Всего 28 типов нейронов, включая классический (1). Нет смысла приводить все формулы вычисления близости (их обоснования и описания весьма объёмны, с этой информацией можно ознакомиться в соответствующих работах, указанных выше). Ниже приведены только 4 из них в качестве примера:

$$y=\sum_{j=1}^{n}\mu_j\cdot a_j, \tag{1}$$

$$y=\sum_{j=1}^{n}\left(\frac{(m_{0,t}-a_{0,t})^2}{\sigma_{0,t}{}^2}-\frac{(m_{0,j}-a_{0,j})^2}{\sigma_{0,j}{}^2}\right), j\neq t, \tag{2}$$

$$y_h=\sum_{j=1}^{n}\left(\frac{(m_{h,j}-a_j)^2}{\sigma_{h,j}{}^2}\right), \tag{3}$$

$$y_h = \prod_{j=1}^{n} p_{h,j}(a_j), \quad (4)$$

где $a_j$ – значение $j$-го входа нейрона, связанного с определенным признаком; $n$ – размерность нейрона (число входов); $h$ – номер гипотезы («Свой», «Чужой»); $m_{h,j}$, $\sigma_{h,j}$ – математическое ожидание и среднеквадратичное отклонение значений признака, соответствующего $j$-му входу, характерные для гипотезы $h$; $\mu_j$ – весовой коэффициент $j$-го входа классического нейрона; $p_{h,j}(.)$ – функция вычисления плотности вероятности для нормального закона распределения с параметрами $m_{h,j}$ и $\sigma_{h,j}$. Порог $\mu_j$ вычисляется по ГОСТ Р 52633.5, исходя из $m_{h,j}$, $\sigma_{h,j}$ и $b$ – бита ключа, который должен генерировать нейрон [1, 9].

Совокупность параметров $m_{h,j}$, $\sigma_{h,j}$ можно назвать эталоном образа «Свой» ($h=0$) и «Чужой» ($h=1$). В режиме идентификации гипотез могло быть множество – для каждого субъекта создаётся эталон.

Гибридная сеть из разнотипных нейронов, каждый из которых обрабатывает сочетания признаков с определённым уровнем взаимной зависимости или суммарной информативности, названа *гибкой сетью* (хронология развития данного направления представлена в [21]). Такая сеть подстраивается под пространство признаков субъекта, определяя свою конфигурацию, исходя из обучающей выборки. Эти сети могут быть многослойными [17, 21]. Положительный эффект достигается, если на разных слоях используются различные типы нейронов [17]. Однако аналогичного эффекта можно добиться, объединяя разнородные нейроны в один большой слой. В этом случае можно сохранить структурную пластичность сети и менять конфигурацию в процессе функционирования (добавляя, убирая нейроны), что затруднительно, если слоёв несколько (мутации гибких сетей в настоящей работе не рассматриваются).

Если гибридный ПБК (ГПБК) стоится на базе однослойной гибкой сети, то классические нейроны, а также разностные и гиперболические нейроны Байеса принимают решение, исходя из порогового значения $\mu_0$ (5), для остальных нейронов предусмотрена иная функция активации (6).

$$A(y) = \begin{cases} 0, \text{ если } y < \mu_0 \cdot w_0, \\ 1, \text{ если } y > \mu_0 \cdot w_0, \end{cases} \quad (5)$$

$$A(y_0, y_1) = \begin{cases} b, \text{ если } y_0 < y_1 \cdot w_z, \\ \neg b, \text{ если } y_0 > y_1 \cdot w_z, \end{cases} \quad (6)$$

где $b$ – бит ключа, на который «настраивается» нейрон, $w_z$ – весовой коэффициент, влияющий на балансировку порогов всех нейронов $z$-типа.

Рассмотрим процесс создания и обучения гибридного ПБК (ГПБК) на базе однослойной гибкой сети:

– Вычисляются параметры $m_{h,j}$, $\sigma_{h,j}$, $I_j$ (информативность) для каждого $j$-го признака.
– Определяется 10 групп информативности признаков $G_1$–$G_{10}$, начиная с наиболее информативных $0 < I_j < 0{,}1$ ($G_1$), заканчивая наименее $0{,}9 < I_j < 1$ ($G_{10}$), $\Delta I = 0{,}1$. Признаки из $G_{10}$ не учитываются, они почти не влияют на результат распознавания [18].
– По обучающим данным «Свой» вычисляется матрица $R_C$ коэффициентов корреляции $r_{j,t}$ признаков.
– Создаются нейроны для обработки независимых признаков. Для обработки каждой группы $G_i$ создаются отдельные нейроны, причём при переходе от $G_i$ к $G_{i+1}$ количество входов каждого нейрона возрастает в 1,5 раза ($n_{i+1} = \Delta n \cdot n_i$, $\Delta n = 1{,}5$): для квадратичных форм $n_1 = 2$, для «критериев» $n_1 = 4$.
– Формируются нейроны, обрабатывающие зависимые признаки. Нейроны Байеса настраиваются по методике [1]. Для «критериев» и «гравитационных» метрик устанавливается фиксированная размерность $n_c = 20$ и $n_g = 2$ соответственно. Зависимые признаки не нужно группировать по информативности, они имеют схожие формы функций плотности вероятности и близки по информативности.

Указанные значения гиперпараметров ($\Delta I$, $n_i$, $\Delta n$, $n_c$, $n_g$) были получены эмпирически и близки к оптимальным. Каждый признак может входить в несколько нейронов, но в один нейрон не более одного раза.

Нейроны, относящиеся к одной категории (например, квадратичные), настраиваются идентично. Поэтому допустимо формировать единую конфигурацию для всех типов нейронов, относящихся к категории (например, создав «подсеть» разностных нейронов Байеса, она полностью копируется и для гиперболических). Это ускоряет настройку. Желательно, чтобы количество совпадающих синапсов для нейронов, базирующихся на одной мере близости, не превышало 25%. Иначе ИНС будет избыточной.

В состав ГПБК может входить НПБК (ГПБК+).

Настройка многослойной гибкой сети отличается тем, что нейроны последующих слоёв воспринимают выходы предыдущего как признаки, поэтому обучающая выборка «пропускается» через очередной слой для настройки следующего. Процесс повторяется циклически, пока не будут настроены все слои.

Описанный процесс обучения абсолютно устойчив. Время обучения гибкой сети сопоставимо со временем обучения широкой сети (при равном числе нейронов и объёме выборки), однако при этом количество ошибочных решений у гибкой сети гораздо меньше (табл. 2). Сравнимый с НПБК уровень ошибок для ГПБК получается на почти двукратно меньшем числе обучающих примеров «Свой» (8–9), что эквивалентно увеличению скорости обучения.

### *Обучающая и тестовая выборки*

Пока единой открытой базы с обучающей и тестовой выборкой, соответствующей всем требованиям репрезентативности [8], не существует. В данной работе использована проприентарная база рукописных и голосовых паролей [17]. Рукописные образы формировались с использованием планшета Wacom (частота опроса – 200 Гц, 1024 уровня давления), голосовые – с использованием микрофонов Pioneer, Sony (диапазон частот не менее 70–12000 Гц).

Открытые голосовые базы [22] в основном состоят из фрагментов произвольной речи (для тестирования FRR и обучения нужны примеры произношения пароля или фразы). Открытые базы рукописных образов [23] разнородны (разные устройства ввода, мало данных динамики и давления пера, недостаточно примеров «Свой»). Поэтому из открытых источников взяты примеры (с параметрами записей не ниже используемых в работе), но только для расширения тестовой выборки «Чужих». В качестве основных источников использованы: VoxForge [http://www.voxforge.org] (русский набор включает 630 дикторов), обезличенная база рукописных образов веб-сервиса SignToLogin [24]. Всего для эксперимента отобрано:

– выборка «Свои»: 160 подписантов и 160 дикторов (пол и возраст распределены равномерно от 18 до 35 лет), каждый воспроизвёл образ своего пароля 80 раз (15 обучающих и 65 тестовых примеров «Свой» собраны в два этапа с интервалом в несколько недель);
– выборка «Чужие»: по одному тестовому примеру других 650 рукописных и 650 голосовых паролей, воспроизведенных «Чужими» (другими субъектами).

Для построения каждого ПБК использовалось по 15 примеров «Свой» и по 159 примеров «Чужой» (каждый из 160 испытуемых является «Чужим» по отношению ко всем остальным).

### *Расширение исходной тестовой выборки естественных образов «Чужие»*

Метод высокоточной быстрой оценки стойкости НПБК из ГОСТ Р 52633.3 подходит и для ГПБК, однако для полностью корректного тестирования желательно иметь несколько десятков тысяч исходных примеров «Чужих». При недостатке естественных образов исходную тестовую выборку допустимо расширить путём генерации независимых синтетических аналогов, которые обладают соответствующими свойствами биометрического образа. При проведении испытаний двухфакторной системы допустимо объединять рукописные и голосовые образы, принадлежащие разным людям, в один общий, так как они независимы. При скрещивании 650 примеров «каждый с каждым» можно получить 422500, для эксперимента также отобрано 50000 тестовых рукописно-голосовых тайных образов. Однако при тестировании однофакторной системы 650 примеров явно недостаточно.

Для тестирования достаточно сгенерировать векторы признаков $\bar{a}$, принадлежащих несуществующим субъектам. Каждый признак $a_j$ имеет некоторое распределение всех возможных значений для всех субъектов. Приближённо эту информацию о глобальной совокупности можно получить, рассчитав $m_{1,j}$, $\sigma_{1,j}$ для всех $N$ признаков по данным выборки «Чужие» ($h=1$) и получив векторы $\bar{a}_{Чm}=\{m_{1,1},...,m_{1,N}\}$ и $\bar{a}_{Чσ}=\{\sigma_{1,1},...,\sigma_{1,N}\}$. Однако все признаки субъекта в той или иной степени зависимы. Эту информацию также необходимо учитывать при генерации реалистичного образа.

Ковариационная матрица $Q_C$ признаков для каждого субъекта почти уникальна (она также является симметричной неотрицательно определённой). Следует рассчитать матрицы $Q_{Ck}$ для как можно большего числа испытуемых. Эти расчёты можно выполнить, используя выборку «Свои» (каждый коэффициент ковариации $q_{j,t}$ из матрицы $Q_{Ck}$ рассчитывался на основании 80 примеров образа $k$-го испытуемого, $k=1..160$, размерность всех матриц $Q_{Ck}$ одинакова и равна $N{\cdot}N$). Выборку «Чужие» для этой цели использовать нельзя, так как в ней образ каждого субъекта представлен всего одним примером. Далее вычисляются аналогичные параметры $m(q_{j,t})$ и $\sigma(q_{j,t})$, из которых формируются 2 квадратные матрицы $Q_m$ и $Q_\sigma$ ($N{\cdot}N$). Таким образом, $\bar{a}_{Чm}$, $\bar{a}_{Чσ}$, $Q_m$ и $Q_\sigma$ содержат всю необходимую информацию о глобальной совокупности значений признаков и их взаимосвязи.

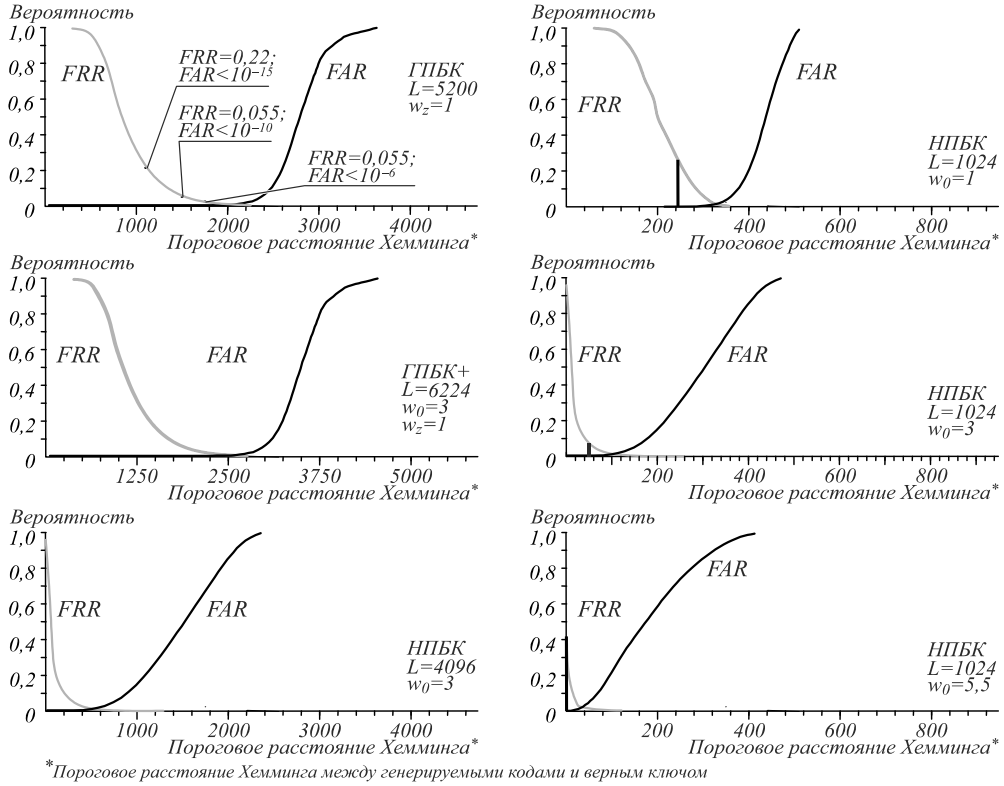
Для генерации образа нового «Чужого» можно выполнить следующие шаги:

1. Сгенерировать методом Монте–Карло вектор независимых признаков $\bar{a}'$ с поправкой под соответствующие параметры из $\bar{a}_{Чm}$, $\bar{a}_{Чσ}$. Выбор генератора нормальных величин не принципиален (можно использовать преобразование Бокса–Мюллера).
2. С учётом данных $Q_m$ и $Q_\sigma$ аналогичным образом сгенерировать элементы $q_{j,t}$ для матрицы $Q_Ч$. На данном шаге создаётся индивидуальная ковариационная матрица признаков субъекта.
3. С помощью разложения Холецкого вычислить нижнюю треугольную матрицу $L$ ($Q_Ч=L{\cdot}L^T$). Если $Q_Ч$ не оказывается положительно определённой, нужно повторить шаг 2.
4. Сгенерировать вектор случайных стандартных нормальных независимых величин $\bar{u}$.
5. Вычислить $\bar{a}=\bar{u}\cdot\mathrm{L}+\bar{a}'$.

Полученный вектор $\bar{a}$ можно использовать в качестве тестового примера, его корреляционные связи $Q_Ч$ и значения признаков будут подчиняться соответствующим распределениям глобальной совокупности, т.е. $\bar{a}$ близок к естественному образу. Но при значительных отклонениях закона распределения признаков от нормального процедура становится некорректной. Альтернативный подход может заключаться в построении нейросетевого генератора реалистичных «Чужих» на основе расширяющейся генеративной нейронной сети (так называемый кодировщик) [2, 3], но этот вопрос требует отдельной проработки.

### *Тестирование ГПБК*

Проведена высокоточная оценка FAR (со скрещиванием пар близких «Чужих» по методу ГОСТ Р 52633.3 через нахождение промежуточных $\vec{a}$). Основные результаты представлены в табл. 2. На рис. 2 видно, что увеличение $L$ (размерности) НПБК с 1024 до 2048 нейронов не даёт преимуществ, в то время как решения ГПБК становятся более надёжными (значит, корреляция между разрядами генерируемого кода ослабевает, что говорит об эффективности использования разнородных нейронов). При использовании ГПБК+ существенного различия с ГПБК по надёжности решений не наблюдается.




*Пороговое расстояние Хемминга между генерируемыми кодами и верным ключом

*Рис. 2. Характеристические кривые вероятностей ошибок двухфакторной аутентификации от допустимого числа неверных бит ключа. На графиках, расположенных СПРАВА, жирной линией показаны точки, в которых FAR($w_0$=5,5)≈FAR($w_0$=3)≈FAR($w_0$=1) при разных FRR*

Из табл. 2 видно, что надёжность для двумерного рукописного и низкокачественного (8 кГц) голосового паролей соизмерима (как и для трёхмерного рукописного и высококачественного (24 кГц) голосового).

В ГОСТ Р 52633.3 для НПБК предусмотрена процедура экспресс-проверки FAR, которая даёт грубую оценку на основании опытов с 128 тестовыми примерами «Чужих» (чтобы пользователь мог убедиться в корректности работы ПБК). Эта процедура основана на гипотезе о том, что расстояния Хемминга между выходами НПБК и верным ключом субъекта для тестовых образов «Чужие» имеют нормальное распределение. Для НПБК это справедливо (с некоторыми оговорками). Однако для ГПБК этого не наблюдается. Более того, распределение меняется в зависимости от соотношения количества нейронов разных типов, входящих в состав ГПБК. Поэтому для ГПБК создание аналогичного метода «ускоренного быстрого тестирования» проблематично. Тем не менее можно вывести эмпирические соотношения, которые позволяют грубо оценивать FAR для ГПБК, через аналогичную оценку для НПБК, при $FRR_{ГПБК} \approx FRR_{НПБК}$.

*Табл. 2. Показатели FFR/FAR по результатам проведённого эксперимента*

| Тип ПБК | Голосовой пароль | | Рукописный пароль | | Два фактора | |
|---|---|---|---|---|---|---|
| | 24 кГц | 8 кГц | XYP | XY | 24 кГц + XYP | 8 кГц + XY |
| НПБК | 0,234 / $<10^{-6}$ | 0,245 / $<10^{-4}$ | 0,24 / $<10^{-6}$ | 0,19 / $<10^{-4}$ | 0,181 / $<10^{-10}$ \| 0,071 / $<10^{-6}$ | 0,163 / $<10^{-9}$ |
| ГПБК | 0,142 / $<10^{-6}$ | 0,125 / $<10^{-4}$ | 0,137 / $<10^{-6}$ | 0,145 / $<10^{-4}$ | 0,055 / $<10^{-10}$ \| 0,22 / $<10^{-15}$ | 0,044 / $<10^{-9}$ |
| ГПБК+ | 0,144 / $<10^{-6}$ | 0,127 / $<10^{-4}$ | 0,141 / $<10^{-6}$ | 0,148 / $<10^{-4}$ | 0,055 / $<10^{-10}$ \| 0,22 / $<10^{-15}$ | 0,044 / $<10^{-9}$ |

### *Защита нейросетевых контейнеров*

К сожалению, все меры близости, кроме (1), компрометируют биометрический эталон пользователя ($m_{h,j}$, $\sigma_{h,j}$ нужно хранить), что противоречит требованиям ГОСТ Р 52633.0. В [7] предлагается защищать нейросетевые контейнеры путём размножения ошибок образа «Чужой» с применением обратимых и необратимых преобразований. Каждый классический нейрон имеет таблицы связей и весов. Для защиты таблиц нейросетевых функционалов нужно применять механизм защищённого нейросетевого контейнера (ЗНК). Нейроны выстраиваются в цепочке путём создания перекрёстных связей (рис. 3). После обучения таблицы каждого нейрона шифруются наложением гаммы, представляющей собой контрольную сумму выходов всех предыдущих нейронов в цепочке (7):

$$tables_l' = \\ = XOR(tables_l, hash(pass, b_1, ..., b_{l-1})), \quad (7)$$

где $l$ – номер нейрона в цепочке, $hash()$ – криптографическая хеш-функция, $pass$ – пароль, который является опциональным и служит как дополнительный фактор защиты. Первый нейрон остаётся незащищённым, однако все классические нейроны имеют «встроенную» защиту в виде весов $\mu_j$, из которых нельзя прямым численным методом восстановить данные обучающей выборки ($m_{h,j}, \sigma_{h,j}, b_l$). Поэтому параметры $b_l$ также можно считать неизвестными для злоумышленника, как и pass.

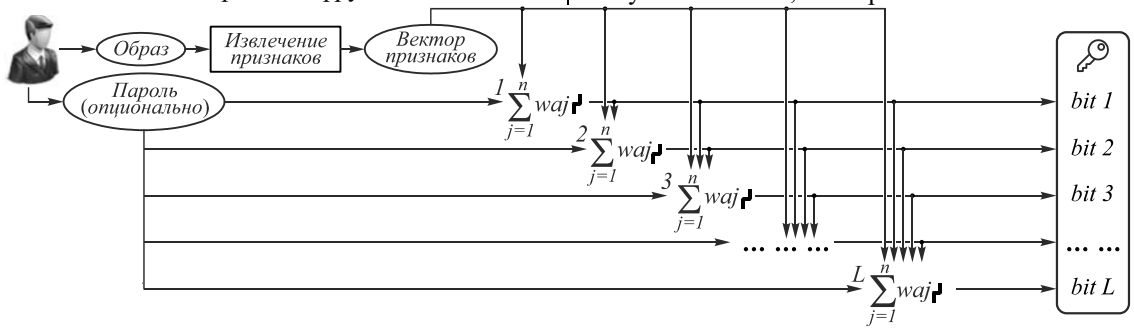



*Рис. 3. Механизм ЗНК. Каждый следующий нейрон связан со всеми предыдущими. Так нейроны последовательно расшифровывают связи и веса последующих нейронов*

ЗНК можно применить для защиты ГПБК+. Объединение НПБК с ГПБК необходимо, так как классические нейроны следует размещать в начале цепочки, а гибридные – в конце (как незащищённые). Гибридные нейроны будут надёжно защищены, если классических нейронов будет много (достаточно 256).

В режиме ЗНК энтропия выходов НПБК (как и ГПБК) при поступлении образа «Чужой» становится близкой к «белому шуму» [7]. Это объясняется тем, что если хоть один нейрон ошибается, таблицы параметров следующего нейрона расшифровываются неверно, что приводит к лавинообразному накоплению ошибок (рис. 3). ЗНК препятствует осуществлению направленного перебора образов «Чужой» для несанкционированного восстановления ключа (или его отдельных бит $b_l$).

Настройка НПБК при активации ЗНК сопряжена с изменением FRR и FAR. В режиме ЗНК при обработке образа «Свой» ни один нейрон не должен ошибиться (или это вызовет накопление ошибок). Для этого $w_0$ для НПБК необходимо снизить (см. рис. 2, справа). Настройку ЗНК для ГПБК+ выполнить сложнее. Необходимо настроить все гиперпараметры $w_z$, заставить разнородные нейроны работать «сообща», не допуская ошибок 1-го рода, при сохранении низкого уровня FAR. На данный момент удалось балансировать только 9 типов нейронов из 28 (классический, квадратичные, Байесовские и некоторые критерии). В результате в режиме ЗНК для двухфакторной аутентификации получены оценки: FRR = 0,07 при $FAR < 10^{-6}$ (FRR = 0,24 при $FAR < 10^{-12}$). Если длина ключа ниже, чем количество нейронов первого слоя, обеспечивающее необходимое качество решений, на выходе защищённого ГПБК можно расположить второй слой необучаемых ($\mu_j = 1$) классических нейронов (1) с функцией активации (5), где $w_0 = 0{,}5$. Количество нейронов второго слоя должно быть равно длине ключа. Если нейрон второго слоя настраивается на $b = 1$, то количество его входов с ожидаемым значением «1» должно быть на один больше, чем с ожидаемым значением «0». При $b = 0$ должно быть наоборот. Под «ожидаемым» подразумевается значение $b$, на которое настраивается соответствующий нейрон первого слоя сети.

### *Заключение*

На базе разработанного аппарата гибких нейронных сетей предложена новая модель гибридного преобразователя «биометрия–код» (ГПБК), который превосходит классический нейросетевой ПБК по надёжности биометрической аутентификации при равном объёме обучающей выборки (и сопоставим при почти двукратно меньшем объёме). ГПБК в незащищённом виде компрометирует эталон пользователя, но комплексируя ГПБК с НПБК можно успешно применить механизм защиты таблиц нейросетевых функционалов, создав защищённый гибридный нейросетевой контейнер, из которого будет очень сложно извлечь биометрический образ и ключ субъекта.

Разработан метод двухфакторной биометрической аутентификации по голосовому и рукописному паролям (FRR = 0,03 при FAR < $10^{-6}$, FRR = 0,055 при FAR<$10^{-10}$, FRR = 0,22 при FAR < $10^{-15}$). Для обработки образов предложено два варианта извлечения признаков – образов низкого и высокого качества. На базе метода ГОСТ Р 52633.3 проведён эксперимент по точной оценке FAR для ГПБК в режиме однофакторной и двухфакторной аутентификации. Для этого предложена методика расширения выборки «Чужих».

### *Благодарности*

Исследование выполнено за счёт гранта Российского научного фонда (проект №17-71-10094).

### *Литература*

1. **Иванов, А.И.** Оценка надёжности верификации автографа на основе искусственных нейронных сетей, сетей многомерных функционалов Байеса и сетей квадратичных форм / А.И. Иванов, П.С. Ложников, А.Е. Сулавко // Компьютерная оптика. – 2017. – Т. 41, № 5. – С. 765-774.
2. **Hafemann, L.G.** Writer-independent feature learning for offline signature verification using deep convolutional neural networks / L.G. Hafemann, R. Sabourin, L.S. Oliveira. // 2016 International Joint Conference on Neural Networks (IJCNN). – 2016. – P. 2576-2583.
3. **Souza, V.L.F.** A writer-independent approach for offline signature verification using deep convolutional neural networks features / V.L.F. Souza, A.L.I. Oliveira, R. Sabourin // 2018 7th Brazilian Conference on Intelligent Systems (BRACIS). – 2018. – P. 212-217.
4. **Tachibana, H.** Efficiently trainable text-to-speech system based on deep convolutional networks with guided attention / H. Tachibana, K. Uenoyama, Sh. Aihara // 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). – 2018. – P. 4784-4788.
5. **Mai, G.** On the reconstruction of face images from deep face templates / G. Mai, K. Cao, P.C. Yuen, A.K. Jain // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2019. – Vol. 41, Issue 5. – P. 1188-1202.
6. **Hafemann, L.G.** Characterizing and evaluating adversarial examples for offline handwritten signature verification / L.G Hafemann, R. Sabourin, L.S. Oliveira // IEEE Transactions on Information Forensics and Security. – 2019. – Vol. 14, Issue 8. – P. 2153-2166. – DOI: 10.1109/TIFS.2019.2894031.
7. **Гулов, В.П.** Перспектива нейросетевой защиты облачных сервисов через биометрическое обезличивание персональной информации на примере медицинских электронных историй болезни (краткий обзор литературы) / В.П. Гулов, А.И. Иванов, Ю.К. Язов, О.В. Корнеев // Вестник новых медицинских технологий – 2017. – Т. 24, № 2 – С. 220-225.
8. **Ахметов, Б.С.** Алгоритмы тестирования биометрико-нейросетевых механизмов защиты информации / Б.С. Ахметов, В.И. Волчихин, А.И. Иванов, А.Ю. Малыгин. – Алматы: КазНТУ имени К.И. Сатпаева, 2013. – 152 с.: ил.
9. **Ложников, П.С.** Биометрическая защита гибридного документооборота / П.С. Ложников. – Новосибирск: Изд-во СО РАН, 2017. – 130 с.
10. **Torfi, A.** Text-independent speaker verification using 3D convolutional neural networks / A. Torfi, J. Dawson, N.M. Nasrabadi // 2018 IEEE International Conference on Multimedia and Expo (ICME). – 2018. – P. 1-6.
11. **Akhmetov, B.S.** Training of neural network biometry-code converters / B.S. Akhmetov, A.I. Ivanov, Z.K. Alimseitova // News of the National Academy of Sciences of the Republic of Kazakhstan, Series of Geology and Technical Sciences. – 2018. – p. 61-68.
12. **Malygin, A.** Application of artificial neural networks forhandwritten biometric images recognition / A. Malygin, N. Seilova, K. Boskebeev, Zh. Alimseitova // Computer Modelling and New Technologies. – 2017. – Vol. 21, Issue 1. – P. 31-38.
13. **Горшков, Ю.Г.** Обработка речевых и акустических биомедицинских сигналов на основе вейвлетов / Ю.Г. Горшков. – М.: Радиотехника, 2017. – 240 с.
14. **Lukic, Y.** Speaker identification and clustering using convolutional neural networks / Y. Lukic, C. Vogt, O. Dürr, T. Stadelmann // IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP). – 2016. – P. 1-6.
15. **Жиляков, Е.Г.** Алгоритмы обнаружения основного тона речевых сигналов / Жиляков Е.Г., Фирсова А.А., Чеканов Н.А. // Научные ведомости БелГУ. Сер. Экономика. Информатика. – 2012. – № 1(120), вып. 21. – С. 135-143.
16. **Vasilyev, V.I.** Identification of the psychophysiological state of the user based on hidden monitoring in computer systems / V.I. Vasilyev, A.E. Sulavko, S.S. Zhumazhanova, R.V. Borisov // Scientific and Technical Information Processing. – 2018. – Vol. 45, Issue 6. – P. 398-410.
17. **Sulavko, A.E.** Subjects authentication based on secret biometric patterns using wavelet analysis and flexible neural networks / A.E. Sulavko, D.A. Volkov, S.S. Zhumazhanova, R.V. Borisov // XIV International Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering (APEIE). – 2018. – P. 218-227.
18. **Sulavko, A.E.** Perspective neural network algorithms for dynamic biometric pattern recognition in the space of interdependent features / A.E. Sulavko, S.S. Zhumazhanova, G.A. Fofanov // Dynamics of Systems, Mechanisms and Machines. – 2018. – P. 1-12.
19. **Ivanov, A.I.** Comparable estimation of network power for chi-squared Pearson functional networks and Bayes hyperbolic functional networks while processing biometric data / A.I. Ivanov, P.S. Lozhnikov, S.E. Vyatchanin. // Control and Communications. – 2017. – P. 1-3.
20. **Sulavko, A.E.** Biometric pattern recognition using wide networks of gravity proximity measures / A.E. Sulavko, S.S. Zhumazhanova // Journal of Physics: Conference Series. – 2018. – Vol. 1050. – 012082.
21. **Vasilyev, V.I.** Flexible fast learning neural networks and their application for building highly reliable biometric cryptosystems based on dynamic features / V.I. Vasilyev, P.S. Lozhnikov, A.E. Sulavko, G.A. Fofanov, S.S. Zhumazhanova // IFAC-PapersOnLine. – 2018. – Vol. 51, Issue 30. – P. 527-532.
22. **Larcher, A.** Text-dependent speaker verification: Classifiers, databases and RSR2015 / A. Larcher, K.A. Lee, B. Ma, H. Li // Speech Communication. – 2014. – Vol. 60. – P. 56-77.
23. **Diaz, M.** A perspective analysis of handwritten signature technology / M. Diaz, M.A. Ferrer, D. Impedovo, M.I. Malik, G. Pirlo, R. Plamondon // ACM Computing Surveys. – 2019. – Vol. 51, Issue 6. – 117.
24. **Lozhnikov, P.** Cloud biometrical system identification through handwriting dynamics “SignToLogin” / P. Lozhnikov, A. Sulavko. – Certificate of registration No. TX 7-640-429. – Date of registration 18.12.2012.

### *Сведения об авторе*

**Сулавко Алексей Евгеньевич** (Омск, Россия) – кандидат технических наук, доцент кафедры комплексной защиты информации ФГБОУ ВО ОмГТУ. E-mail: *sulavich@mail.ru* .

*ГРНТИ: 28.23.37*
*Поступила в редакцию 9 мая 2019 г. Окончательный вариант – 16 октября 2019 г.*

# Highly reliable two-factor biometric authentication based on handwritten and voice passwords using flexible neural networks

*A.E. Sulavko* [1]
[1] *Omsk State Technical University, Omsk, Russia*

***Abstract***

The paper addresses a problem of highly reliable biometric authentication based on converters of secret biometric images into a long key or password, as well as their testing on relatively small samples (thousands of images). Static images are open, therefore with remote authentication they are of a limited trust. A process of calculating the biometric parameters of voice and handwritten passwords is described, a method for automatically generating a flexible hybrid network consisting of various types of neurons is proposed, and an absolutely stable algorithm for network learning using small samples of "Custom" (7-15 examples) is developed. A method of a trained hybrid "biometrics-code" converter based on knowledge extraction is proposed. Low values of FAR (false acceptance rate) are achieved.

*Keywords*: hybrid networks, quadratic forms, Bayesian functionals, handwritten passwords, voice parameters, wide neural networks, biometrics-code converters, protected neural containers.

*Acknowledgements*: This work is supported by the Russian Science Foundation under grant №17-71-10094.

*Citation*: Sulavko AE. Highly reliable two-factor biometric authentication based on handwritten and voice passwords using flexible neural networks. Computer Optics 2020; 44(1): 82-91. DOI: 10.18287/2412-6179-CO-567.

***References***

[1] Ivanov AI, Lozhnikov PS, Sulavko AE. Evaluation of signature verification reliability based on artificial neural networks, Bayesian multivariate functional and quadratic forms. Computer Optics 2017; 41(5): 765-774.

[2] Hafemann LG, Sabourin R, Oliveira LS. Writer-independent feature learning for offline signature verification using deep convolutional neural networks. International Joint Conference on Neural Networks 2016: 2576-2583.

[3] Souza VLF, Oliveira ALI, Sabourin R. A writer-independent approach for offline signature verification using deep convolutional neural networks features. 7th Brazilian Conference on Intelligent Systems 2018: 212-217.

[4] Tachibana H, Uenoyama K, Aihara Sh. Efficiently trainable text-to-speech system based on deep convolutional networks with guided attention. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2018: 4784-4788.

[5] Mai G, Cao K, Yuen PC, Jain AK. On the reconstruction of face images from deep face templates. Trans Patt Anal Machine Intell 2019; 41(5): 1188-1202.

[6] Hafemann LG, Sabourin R, Oliveira LS. Characterizing and evaluating adversarial examples for offline handwritten signature verification. IEEE Transactions on Information Forensics and Security 2019; 14(8): 2153-2166. DOI: 10.1109/TIFS.2019.2894031.

[7] Gulov VP, Ivanov AI, Yazov YuK, Korneev OV. Perspective of neuro network protection of cloud services through biometric deployment of personal information on the example of medical electronic history of disease (Brief review of the literature) [In Russian]. Journal of New Medical Technologies 2017; 24(2): 220-225.

[8] Ahmetov BS, Volchihin VI, Ivanov AI, Malygin AYu. Algorithms for testing biometric-neural network information protection mechanisms [In Russian]. Almaty: "KazNTU imeni K I Satpaeva" Publisher; 2013.

[9] Lozhnikov PS. Hybrid workflow biometric protection [In Russian]. "SO RAN" Publisher; 2017.

[10] Torfi A, Dawson J, Nasrabadi NM. Text-independent speaker verification using 3D convolutional neural networks. IEEE International Conference on Multimedia and Expo (ICME) 2018: 1-6.

[11] Akhmetov, BS, Ivanov, AI, Alimseitova, ZK Training of neural network biometry-code converters. News of the National Academy of Sciences of the Republic of Kazakhstan, Series of Geology and Technical Sciences 2018: 61-68.

[12] Malygin A, Seilova N, Boskebeev K, Alimseitova Zh. Application of artificial neural networks forhandwritten biometric images recognition Computer Modelling and New Technologies, 2017, 21(1), 31-38.

[13] Gorshkov YuG. Wavelet-based speech and acoustic biomedical signal processing. Moscow: "Radiotekhnika" Publisher; 2017.

[14] Lukic Y, Vogt C, Dürr O, Stadelmann T. Speaker identification and clustering using convolutional neural networks. 26th International Workshop on Machine Learning for Signal Processing 2016: 1-6.

[15] Zhilyakov EG, Firsova AA, Chekanov NA. Algorithms for detecting the fundamental tone of speech signals. Belgorod State University Scientific Bulletin; Series Economics; Computer Science 2012; 1(120:21): 135-143.

[16] Vasilyev VI, Sulavko AE, Zhumazhanova SS, Borisov RV. Identification of the psychophysiological state of the user based on hidden monitoring in computer systems. Scientific and Technical Information Processing 2018; 45(6): 398-410.

[17] Sulavko AE, Volkov DA, Zhumazhanova SS, Borisov RV. Subjects authentication based on secret biometric patterns using wavelet analysis and flexible neural networks. XIV International Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering 2018: 218-227.

[18] Sulavko AE, Zhumazhanova SS, Fofanov GA. Perspective neural network algorithms for dynamic biometric pattern recognition in the space of interdependent features. Dynamics of Systems, Mechanisms and Machines 2018: 1-12.

[19] Ivanov, AI Lozhnikov, PS Vyatchanin SE. Comparable estimation of network power for chi-squared Pearson functional networks and Bayes hyperbolic functional networks while processing biometric data. Control and Communications 2017: 1-3.
[20] Sulavko AE, Zhumazhanova SS. Biometric pattern recognition using wide networks of gravity proximity measures. J Phys Conf Ser 2018; 1050: 012082.
[21] Vasilyev VI, Lozhnikov PS, Sulavko AE, Fofanov GA, Zhumazhanova SS. Flexible fast learning neural networks and their application for building highly reliable biometric cryptosystems based on dynamic features. IFAC-PapersOnLine 2018; 51(30): 527-532.
[22] Larcher A, Lee KA, Ma B, Li H. Text-dependent speaker verification: Classifiers, databases and RSR2015. Speech Communication 2014; 60: 56-77.
[23] Diaz M, Ferrer MA, Impedovo D, Malik MI, Pirlo G, Plamondon R. A perspective analysis of handwritten signature technology. ACM Computing Surveys 2019; 51(6): 117.
[24] Lozhnikov P, Sulavko A. Cloud biometrical system identification through handwriting dynamics "SignToLogin" Certificate of registration No. TX 7-640-429 from 18.12.2012.

***Authors' information***

**Alexey E. Sulavko,** Candidate of Technical Sciences, Assistant Professor of Complex Information Protection department of Omsk State Technical University E-mail: *sulavich@mail.ru* .

*Received May 9, 2019. The final version – October 16, 2019.*