МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ФГКОУ ВО «Воронежский институт МВД России», г. Воронеж;
ФГБОУ ВО «Вятский государственный университет», г. Киров;
ФАУ «ГНИИИ проблем технической защиты информации ФСТЭК России»,
г. Воронеж;
ФГБОУ ВО «Пензенский государственный университет», г. Пенза;
ФГБОУ ВО «Липецкий государственный педагогический университет», г. Липецк;
ФГБОУ ВО «Рязанский радиотехнический университет», г. Рязань;
ФГБОУ ВО «Оренбургский государственный университет», г. Оренбург;
ФГБОУ ВО «МИРЭА – Российский технологический университет», г. Москва;
ФГУП «18 ЦНИИ» МО РФ, г. Москва;
ФГАНУ «Институт сверхвысокочастотной полупроводниковой электроники имени
В. Г. Мокерова Российской академии наук» (ИСВЧПЭ РАН), г. Москва;
АО «Пензенский научно-исследовательский электротехнический институт», г. Пенза;
Пензенский филиал ФГУП «Научно-технический центр «Атлас», г. Пенза;
ООО «Научно-техническое предприятие «Криптософт», г. Пенза;
АО «Научно-производственное предприятие «Рубин», г. Пенза;
АО «Производственное объединение «Электроприбор», г. Пенза;
АО «Радиозавод», г. Пенза;
АО «Системы управления», г. Москва;
Общероссийская общественная организация «Российское научно-техническое
общество радиотехники, электроники и связи имени А. С. Попова», г. Тула;
«Научно-исследовательский и конструкторский институт радиоэлектронной
техники», филиал ФГУП НПЦ «ПО «Старт» имени М. В. Проценко»,
г. Заречный Пензенской обл.;
ФГБОУ ВО «Петербургский государственный университет путей сообщения
Императора Александра I», г. Санкт-Петербург;
Филиал АО «ПНИЭИ» «Научно-исследовательское предприятие "Аргус"», г. Пенза;
ООО «Научно-производственная фирма "Кристалл"», г. Пенза;
Филиал Военной академии имени Петра Великого, г. Серпухов;
Обособленное подразделение ОАО «Инфотекс», г. Пенза;
ООО «НПФ «КРУГ», г. Пенза;
ООО «Научно-производственное предприятие «БиоКрипт», г. Пенза;
ООО «Биометрика», г. Пенза;
ООО «АЛГОМАТ», г. Калининград

Безопасность информационных технологий

Сборник научных статей по материалам
I Всероссийской научно-технической конференции
(24 апреля 2019 г.)

Пенза Издательство ПГУ 2019

УДК 681.322
Б39

Безопасность информационных технологий : сб. науч. ст. по Б39 материалам I Всерос. науч.-техн. конф. (24 апреля 2019 г.). – Пенза : Изд-во ПГУ, 2019. – 204 с.

ISBN 978-5-907185-48-7

Рассматриваются различные аспекты безопасности информационных технологий. Публикуемые материалы прошли рецензирование.

Издание предназначено для специалистов по безопасности информационных технологий, преподавателей и студентов вузов.

УДК 681.322

URL: http://пниэи.рф/activity/science/BIT/T11–p5.pdf

Приказ
о подготовке и проведении Всероссийской научно-технической конференции «Безопасность информационных технологий» № 525/о от 22.04.2019.

ISBN 978-5-907185-48-7

ПРОДОЛЖЕНИЕ СЛАВНЫХ ТРАДИЦИЙ ПЕНЗЕНСКОЙ НАУЧНОЙ ШКОЛЫ

В. И. Волчихин

Открывая I Всероссийскую научно-техническую конференцию «Безопасность информационных технологий», хочется отметить, что ее проведение стало возможным тесному взаимодействию российских предприятий, организаций и вузов в области вычислительной техники, информационной безопасности и ведущего в регионе вуза – Пензенского государственного университета.

Пенза в 1950–1960-е гг. стала одним из центров развития приборостроения и электроники, в частности, математического машиностроения. После войны здесь был построен самый крупный в стране на то время завод счетно-аналитических машин (САМ), – затем завод вычислительных электронных машин (ВЭМ). В 1948 г. на заводе САМ был создан отдел математических машин, разрабатывавший аналоговые вычислительные машины (АВМ). После ряда преобразований отдел влился в созданный на заводе САМ в 1953 г. Пензенский филиал Московского СКБ-245, который считается родоначальником Пензенского НИИ математических машин (ПНИИММ) – ОАО «НПП «Рубин». Это одно из лидирующих научно-исследовательских предприятий России в области вычислительной техники.

В 1955 г. в Пензу переезжает Б. И. Рамеев, возглавивший разработку цифровых ЭВМ серии «Урал». Выпуск ЭВМ «Урал-1» начался в 1957 г. Была выпущена целая серия ламповых ЭВМ «Урал». В 1960-х гг. начался выпуск полупроводниковых «Уралов», представлявших собой серию программно-совместимых компьютеров. Одновременно в Пензенском НИИ математических машин под руководством Н. С. Николаева и Э. С. Козлова продолжалась

разработка аналоговых ЭВМ, прежде всего сеточных моделей для решения дифференциальных уравнений в частных производных.

Интересно отметить, что сеточные модели были прямыми предшественниками клеточных нейронных сетей для решения дифференциальных уравнений в частных производных. По этому направлению следует отметить самый крупный в мире сеточный электроинтегратор ЭИ-С, созданный в 1957 г., и универсальную сеточную электромодель УСМ-1, созданную в 1960-х гг., выпускавшуюся серийно и использовавшуюся во многих научных и проектных организациях.

Принципиально новые возможности решения сложных краевых задач открылись после создания под руководством Н. С. Николаева и Э. С. Козлова аналого-цифровых вычислительных комплексов (АЦВК) «Сатурн» (1967) и «Сатурн-2» (1973). Важным направлением деятельности ПНИИММ, сейчас ОАО «НПП "Рубин"», является разработка специализированных информационных систем и автоматизированных систем управления.

Еще на целом ряде предприятий города Пензы вычислительная техника стала основным объектом и инструментом производства. Завод точной электромеханики освоил выпуск аналоговой вычислительной техники и устройств подготовки данных. Пензенское конструкторское бюро моделирующих приборов и машин – Пензенское КБ моделирования – решало проблемы цифрового моделирования в тренажеростроении. Завод «Счетмаш» на базе управляющих мини-ЭВМ выпускал системы управления атомными электростанциями. Основной тематикой НИИ вычислительной техники была разработка запоминающих устройств.

Центром подготовки специалистов в области вычислительной техники явился Пензенский государственный университет (ранее Пензенский индустриальный институт, Пензенский политехнический институт, Пензенский государственный технический университет). Еще в 1947 г. здесь образовалась кафедра вычислительного профиля «Счетно-решающие и аналитические машины».

В 1953 г. специальным Постановлением Совета Министров СССР с целью подготовки специалистов для вновь образованного в Пензе научно-исследовательского электротехнического института в Пензенском индустриальном институте была открыта специальность «Электромеханическая аппаратура» для подготовки разработчиков средств защиты информации, которые в настоящее время именуются специалистами по защите информации.

Начиная с 1961 г., в Пензенском политехническом институте проводились исследования алгоритмов обучения искусственных нейронных сетей. С 1998 г. в тематику научных исследований научно-педагогической школы «Исследование и разработка интеллектуальных устройств управления средствами поражения и систем обнаружения, классификации и идентификации объектов» включено исследование и практическое применение интеллектуальных устройств, использующих искусственные нейронные сети большого и сверхбольшого размера. Базой для формирования научного направления стали результаты исследований, проведенные В. И. Волчихиным, А. И. Ивановым, А. Ю. Малыгиным, В. А. Фунтиковым. Ими впервые предложены быстрые неитерационные алгоритмы обучения искусственных нейронных сетей. Применение этих алгоритмов позволило решить задачу увеличения входного массива данных для обучения нейронной сети и значительно повысить качество выходного решения при минимальных временных затратах на сам процесс обучения искусственной нейронной сети. Принципиальным технологическим прорывом является переход к полностью автоматизированному (абсолютно устойчивому) обучению искусственных нейронных сетей неограниченного размера. Многолетние усилия пензенской научной школы в этом направлении исследований завершились в 2011 г. введением в действие на территории РФ национального стандарта ГОСТ Р 52633.5–2011. Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа. Абсолютная устойчивость алгоритмов ГОСТ Р 52633.5 позволяет работать с «плохими» и «очень плохими» исходными биометрическими данными и обучать нейронные сети, способные обрабатывать неограниченное число входных параметров. Фактически национальный стандарт ГОСТ Р 52633.5–2011 снимает любые ограничения на размерность решаемых искусственными нейронными сетями задач.

С вводом в действие данного стандарта нейросетевое подсознание искусственного интеллекта становится способным решать задачи любой размерности (любой вычислительной сложности). Многолетние усилия пензенской научной школы автоматического обучения искусственных нейронных сетей через ГОСТ Р 52633.5–2011 снимают барьер «проклятия размерности». Искусственный интеллект нейросетевого подсознания роботов по своим возможностям оказывается сопоставимым с естественным интеллектом людей–

экспертов, что подтверждено пока только для задач биометрико-нейросетевой аутентификации.

По результатам работ, проведенных в научной школе под руководством профессоров А. И. Иванова, А. Ю. Малыгина, в 2009 г. Пензенский государственный университет был принят в члены ТК362 «Защита информации».

Последние результаты исследований были использованы при разработке пакета из пяти стандартов в рамках национальной Программы стандартизации по высоконадежной биометрии, дополняющих и расширяющих действия ГОСТ Р 52633.0–2006. То, что искусственный интеллект нейросетевого подсознания биометрической аутентификации становится очень эффективным, потребовало разработки специальных методов тестирования. Эти методы также разрабатывались в рамках пензенской научной школы и легли в основу национального стандарта ГОСТ Р 52633.3–2010. Защита информации. Техника защиты информации. Тестирование стойкости средств высоконадежной биометрической защиты к атакам подбора. Без эффективных методов тестирования нейросетевых решений применение их в ответственных приложениях невозможно.

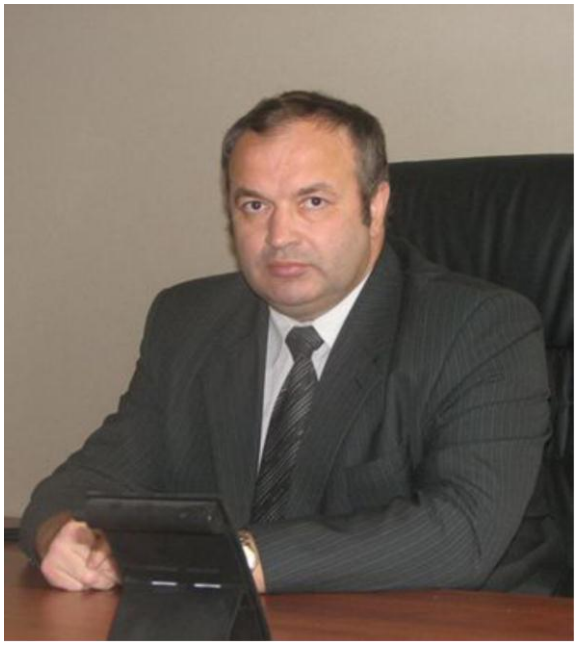
В настоящее время ПГУ в кооперации с рядом организаций разработали проект национального стандарта ГОСТ Р 52633.xx-2-xx. Защита информации. Техника защиты информации. Автоматическое обучение сетей квадратичных нейронов с многоуровневым квантованием биометрических данных, являющийся продолжением пакета стандартов по высоконадежной биометрико-нейросетевой аутентификации.

Проблема информационной безопасности, по словам Президента России В. Путина, стала глобальной: «Стратегической целью обеспечения информационной безопасности в области стратегической стабильности и равноправного стратегического партнерства является формирование устойчивой системы неконфликтных межгосударственных отношений в информационном пространстве».

Вопросы, которые будут рассмотрены на сегодняшней конференции, неразрывно связаны со словами, сказанными Президентом России.

В. И. Волчихин, председатель оргкомитета конференции, президент Пензенского государственного университета, заслуженный деятель науки РФ, доктор технических наук, профессор

ЭКСКУРС В ИСТОРИЮ БЕЗОПАСНОСТИ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ



В. А. Фунтиков

В настоящее время мы являемся свидетелями процессов активной информатизации общества. К сожалению, эти процессы приводят не только к положительным изменениям в нашей жизни. Одной из проблем является защита как самой информации, так и персональных биометрических данных пользователей.

Рассмотрению вышеуказанной проблемы посвящена I Всероссийская научно-техническая конференция «Безопасность информационных технологий», которая основывается на опыте десятилетнего проведения научно-технических конференций кластера пензенских предприятий, обеспечивающих безопасность информационных технологий. На сегодняшний день 27 организаций из разных городов выразили свое желание участвовать в работе Всероссийской конференции.

Одним из направлений, рассматриваемых в течение 10 лет, является биометрическая идентификация и аутентификация личности. По прогнозам отечественных и зарубежных ученых, ветвь искусственного интеллекта, занимающегося синтезом биометрических автоматов для распознавания людей, будет активно развиваться. Появление международных паспортно-визовых документов с биометрией, локальная и дистанционная биометрическая идентификация и аутентификация личности в банковской, образовательной, правоохранительной и других сферах нашей жизни – это уже реалии сегодняшнего дня.

При разработке интеллектуальных автоматов, способных надежно узнавать человека, крайне важно обеспечить информационную безопасность не только систем, но и самих пользователей этих систем. Принципиально важно обеспечить условия, при которых злоумышленники технически не могут извлечь из памяти

автомата биометрический образ человека. Кража биометрических образов людей или их компрометация должны быть исключены. Каждый человек, доверивший свою биометрию автомату, должен быть уверен в том, что его личная биометрическая информация (рисунок отпечатка пальца, динамика автографа, радужная оболочка глаза и т.п.) надежно хранится и не может быть извлечена из памяти и похищена.

Очевидно, что существующая сегодня простейшая биометрия, разработанная по стандарту BioAPI и использующая обычные шаблоны биометрических образов, не способна обеспечить конфиденциальность хранимой биометрической информации. Сегодняшние простейшие биометрические устройства должны рано или поздно уступить свое место более сложным и более интеллектуальным биометрическим устройствам следующего поколения. Как нельзя лучше для реализации интеллекта биометрических устройств следующего поколения подходят искусственные нейронные сети. Это – технологии, способные самообучаться. Если использовать нейросетевой преобразователь биометрия-код доступа к информационному ресурсу, то биометрический шаблон человека исчезает. Вместо шаблона биометрического образа человека биометрические устройства следующего поколения будут хранить таблицы связей и таблицы значений весовых коэффициентов обученной искусственной нейронной сети. Свойство нейросетевых технологий – обеспечивать анонимность биометрических данных распознаваемых людей и их кодов доступа – имеет первостепенное значение. Только обеспечивая конфиденциальность, анонимность, обезличенность биометрии граждан и их кодов доступа, мы можем построить действительно демократическое, информационное общество.

Первый российский стандарт по защите биометрических данных ГОСТ Р 52633–2006 вступил в действие на пять лет раньше соответствующих ему международных аналогов. За истекший период принято шесть стандартов по высоконадежной биометрии. Отрадно отметить, что Пензенский государственный университет принимал в их разработке непосредственное участие. В 2018 г. университет совместно с АО «ПНИЭИ», Пензенским филиалом ФГУП «НТЦ «Атлас», ФГУП «18 ЦНИИ» МО РФ разработали проект национального стандарта ГОСТ Р 52633.xx-2-xx. Защита информации. Техника защиты информации. Автоматическое обучение сетей квадратичных нейронов с многоуровневым квантова-

нием биометрических данных. Предположительно, новый стандарт на территории РФ будет действовать параллельно со стандартом ГОСТ Р 52633.5–2011.

С целью повышения качества подготовки студентов Пензенского государственного университета в сфере разработки технических средств информационной безопасности телекоммуникационных и автоматизированных систем, повышения результативности научно-исследовательской и учебно-методической работы, усиления практической направленности образовательного процесса 1 сентября 2014 г. на базе площадей акционерного общества «Пензенский научно-исследовательский электротехнический институт» создана кафедра «Технические средства информационной безопасности».

Накопленный опыт и знания в области обучения и тестирования обученных нейросетевых преобразователей могут быть использованы в различных областях и сферах нашей жизни. О важности работы в данном направлении Президент России Владимир Путин отметил еще в сентябре 2017 г. в Ярославле, на Всероссийском форуме «Проектория»: «Искусственный интеллект – это будущее не только России, это будущее всего человечества. Здесь колоссальные возможности и трудно прогнозируемые сегодня угрозы. Тот, кто станет лидером в этой сфере, будет властелином мира. И очень бы не хотелось, чтобы эта монополия была сосредоточена в чьих-то конкретных руках, поэтому мы, если мы будем лидерами в этой сфере, также будем делиться этими технологиями со всем миром, как мы сегодня делимся атомными технологиями, ядерными технологиями».

В. А. Фунтиков, сопредседатель оргкомитета конференции, генеральный директор АО «Пензенский научно-исследовательский электротехнический институт», кандидат технических наук

УДК 519.2

А. П. Юнин, А. И. Иванов, К. А. Ратников

ОЦЕНКА КАЧЕСТВА «БЕЛОГО ШУМА»: РЕАЛИЗАЦИЯ ТЕСТА «СТАИ ОБЕЗЬЯН» ЧЕРЕЗ МНОЖЕСТВО СВЕРТОК ХЭММИНГА ДЛЯ РАЗНЫХ СИСТЕМ СЧИСЛЕНИЯ

Аннотация. Рассматривается проблема повышения корректности вычисления энтропии длинных кодов со слабо зависимыми разрядами, порождаемыми нейросетевыми преобразователями биометрия-код или хэшированием биометрических данных средствами облегченной криптографии. Предложено рассматривать портрет «белого шума» в системе сверток Хэмминга, полученных в разных системах счисления.

A. P. Junin, A. I. Ivanov, K. A. Ratnikov

QUALITY ASSESSMENT OF "WHITE NOISE": THE IMPLEMENTATION OF THE TEST "FLOCKS OF MONKEYS" THROUGH A LOT OF HAMMING CONVOLUTIONS, BUILT FOR DIFFERENT NUMBER SYSTEMS

Abstract. This article deals with the problem of increasing the correctness of calculating the entropy of long codes with weakly dependent bits by nejrosetevymi converters code or biometrics biometrics means hashing Lite Cryptography. Suggested that the portrait of white noise in the system folds Hamming obtained in different notations.

Проблема контроля уровня случайности данных, получаемых в биометрии

Активное развитие цифровой экономики в России и за рубежом делает необходимым повсеместное (массовое) использование облегченной криптографии и биометрии. Облегченная криптография применяется в том случае, когда стоимость операций невелика и они выполняются в доверенном процессоре SIM-карты, SD-карты, 4-битном процессоре RFID-карты с микропитанием от созданного считывателем электромагнитного поля.

Возможность ослабления криптографии обусловлена также тем, что она используется не самостоятельно, а в связке с биомет-

рией пользователя [1–3]. Биометрия и ослабленная криптография в этом случае дополняют и усиливают друг друга.

Однако проблема локального получения действительно случайных чисел (действительно «белого шума») существует. При этом, чем короче число, тем важнее становится контроль уровня его случайности. В этом отношении проблема контроля случайности чисел для отечественной нейросетевой биометрии оказывается менее острой, чем та же проблема для зарубежных «нечетких экстракторов» [4–7]. Зарубежные «нечеткие экстракторы» гораздо менее интеллектуальны нейросетевых преобразователей биометрия-код [8, 9]. «Нечеткие экстракторы» строятся на использовании кодов, обнаруживающих и исправляющих ошибки за счет высокой избыточности. Например, Даугман [7] использует самокорректирующийся код с 20-кратной избыточностью, т.е. длина выходного кода «нечеткого экстрактора» в 20 раз короче, чем число контролируемых биометрических параметров.

В свою очередь, число контролируемых биометрических параметров не может быть как угодно большим. Обычно производители средств биометрической защиты не указывают, сколько параметров и какие параметры они контролируют. Единственным исключением является среда моделирования «БиоНейроАвтограф» [10], эта среда преобразует рукописный автограф (любой рукописный знак) в 416 контролируемых биометрических параметров (416 коэффициентов двухмерного преобразования Фурье рукописного образа). Положительным свойством среды моделирования «БиоНейроАвтограф» является то, что на данный момент времени – это единственный достоверный источник биометрической информации для студентов, аспирантов, преподавателей университетов России, Беларуси, Казахстана, Армении, свободно читающих на русском языке. Все данные среды моделирования доступны для наблюдения [11], могут быть перегружены в иной математический редактор: MathCAD, MatLAB, Mathematica. Далее пользователь уже сам может написать собственную нейросетевую обработку биометрических данных в рамках курсового проекта, дипломной работы или диссертации.

Если вернуться к заявленной тематике, то для данных среды моделирования «БиоНейроАвтограф» «нечеткий экстрактор» будет иметь выходной личный ключ длинной 416/20 = 21 бит. Такая длина кода несопоставимо меньше, чем 256 бит ключей, получаемых

штатным нейросетевым преобразователем среды «БиоНейроАвтограф». При использовании ключа длиной в 21 бит требования к уровню его случайности на много выше, чем к уровню случайности ключа в 256 бит. Если уровень не случайности ключа в 21 бит составит 10 % (реальная энтропия генератора может составить 19 бит) – это намного хуже для биометрических приложений, чем 10 %-я детерминированность генератора ключа в 256 бит, полученного от генератора, имеющего реальную энтропию – 230 бит. Все это следствие того, что собственная энтропия биометрического рукописного образа может находиться в интервале от 11 до 97 бит. Для слабых биометрических образов энтропии в 19 бит генератора ключа «нечетких экстракторов» вполне достаточно. Атаковать нужно со стороны слабой биометрии с 11-битной собственной энтропией. Для сильной биометрии с собственной энтропией в 97 бит атаковать нужно, подбирая 21 выходной бит «нечеткого экстрактора».

Синтез личного криптографического ключа из неоднозначных компонент биометрических данных

Особенностью биометрии является то, что каждый пример реализации одного и того же биометрического образа отличается от других примеров. Для выделения из них нестабильной компоненты достаточно взять два примера биометрического образа, осуществить центрирование и нормирование их данных и получить их разность. Эта ситуация отображена на рис. 1. Если теперь проквантовать эту разность, мы получим случайную последовательность.

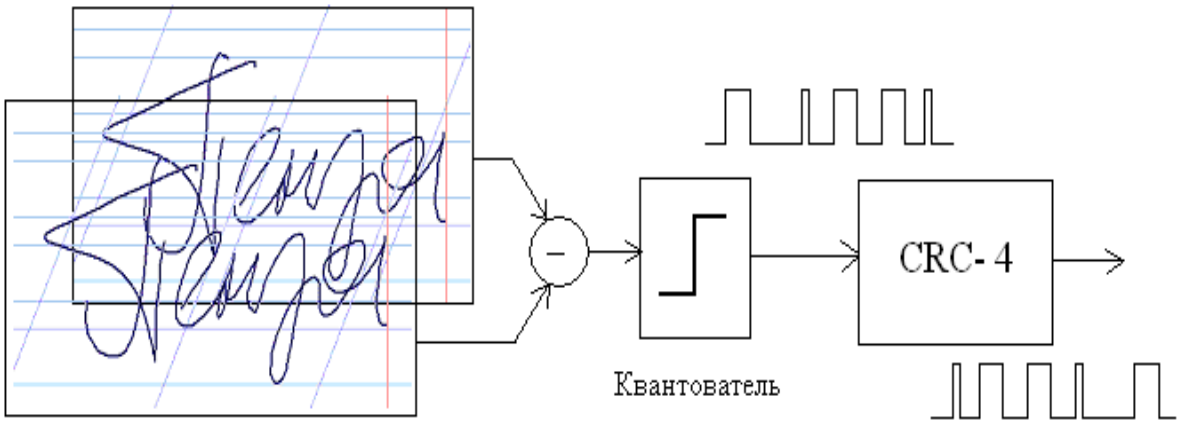


Рис. 1. Получение случайной последовательности из двух примеров одного биометрического образа

При вычитании стабильная часть примеров рукописного биометрического образа «*Пенза*» устраняется, подчеркивается его нестабильная часть. Появление нестабильной части биометрических примеров обусловлено влиянием множества факторов, – точно повторить рукописный образ человек не может даже в случае попытки его обвода по шаблону.

Для того, чтобы усилить полученную случайную последовательность, достаточно получить другую псевдослучайную последовательность, от программного генератора и сложить их по модулю два. Можно поступить иначе и усилить случайную последовательность пропустив ее через нелинейную рекурсивную свертку, например CRC-4 (подсчет контрольных сумм).

Если бы мы вычислили криптографическую хэш-функцию от полученной случайной последовательности [3], то можно было бы ее использовать для получения личного ключа без исследования ее качества. То, что мы, ориентируясь на малые вычислительные ресурсы, применили некриптографическое хэширование, в форме рекурренты CRC-4, приводит к необходимости оценки энтропии ключа. Заранее неизвестно, достаточно ли стандартная рекуррента CRC-4 увеличивает энтропию естественной нестабильности примеров биометрического образа.

Контроль энтропии в пространстве однобитных сверток Хэмминга

Если бы нами был создан ключ в 256 бит (например, для последующего обучения нейросети среды «БиоНейроАвтограф»), то оценить его энтропию по Шеннону технически невозможно. В связи с этим приходится от обычных кодов переходить в пространство сверток Хэмминга [2]:

$$h = 256 - \sum_{i=1}^{256} ("c_i") \oplus ("x_i"), \qquad (1)$$

где "c_i" – i-й разряд проверяемого разряда кода синтезированного ключа "x_i" – эталонная последовательность «не белого шума».

Если мы будем иметь ключ "$\bar{c}$" с почти независимыми разрядами, то множество сверток Хэмминга (1) с фрагментами «не белого шума» должно дать следующие статистические моменты:

$$\begin{cases} E(h) = 128 \ \ бит \\ \sigma(h) = 8 \ \ бит \end{cases}. \qquad (2)$$

В случае, если математическое ожидание Хэмминга E(h) будет существенно отличаться от значения 128 бит, мы получим отсутствие баланса равновероятных состояний «0» и «1» во всех разрядах кода. Если стандартное отклонение σ(h) будет существенно больше 8 бит, то мы должны сделать вывод о сильной корреляционной связи разрядов синтезированного кода ключа.

Контроль энтропии в пространстве 8-битных кодировок

В качестве эталонных последовательностей «не белого» шума, например, могут быть взяты фрагменты текста на русском языке [12–14]. Каждый из фрагментов текста на русском языке при вычислении расстояний Хэмминга фактически накрывается гаммой синтезированного ключа. Выполнение условия (2) эквивалентно покрытию текста идеальной гаммой, после чего защищенный гаммой текст становится «белым шумом».

Следует отметить, что снижение криптографических свойств синтезированного ключа $"\overline{c}"$ возникает в очевидном случае замены его на код осмысленного пароля [13, 14]. Естественно, что улучшение криптографических качеств ключа $"\overline{c}"$ будет монотонно приближаться к выполнению условий (2).

В работах [13, 14] показано, что при вычислениях энтропии русского языка приходится учитывать 8-битную кодировку русских и английских текстов. В этом случае расстояние Хэмминга будет вычисляться следующим образом:

$$h_8 = \sum_{i=1}^{32} \left| "c_i, c_{i+1}, ..., c_{i+7}" - "x_i, x_{i+1}, ..., x_{i+7}" \right| , \qquad (3)$$

где $"c_i, c_{i+1}, ..., c_{i+7}"$ – 8 бит кода в *i*-м фрагменте кода ключа; $"x_i, x_{i+1}, ..., x_{i+7}"$ – 8 бит кода образцового текста, принадлежащего i-му знаку эталонной последовательности.

Очевидно, что для свертки Хэмминга по модулю 256 (3) могут быть найдены ограничения на математическое ожидание и стандартное отклонение типа (2). Пользуясь близостью к этим ограничениям, мы всегда сможем оценить то, насколько тестируемая последовательность $"\overline{c}"$ близка к «белому» шуму.

Для нас принципиально важно то, что результаты тестирования через однобитные свертки Хэмминга (1) и тестирование в 8-битных свертках Хэмминга по модулю 256 сильно различаются [14].

Множество сверток Хэмминга, вычисленных обычными процессорами в системах счисления, кратных двум

Фактически мы имеем две свертки Хэмминга по модулю 2 (выражение (1)) и по модулю 256 (выражение (3)). Свертку по модулю два (1) можно рассматривать как эквивалент теста «стаи обезьян», набирающих текст на русском языке до его полного совпадения с контрольным фрагментом на однобитной печатающей машинке. Свертку по модулю 256 32 знаков образцового русского текста (3) следует рассматривать как обычную печатающую машинку в руках «стаи обезьян».

Исходя из этой интерпретации, можно предложить промежуточные свертки, для других печатающих машинок:

$$h_2 = \sum_{i=1}^{128} \left| "c_i, c_{i+1}" - "x_i, x_{i+1}" \right|, \tag{4}$$

$$h_4 = \sum_{i=1}^{64} \left| "c_i, c_{i+1}, ..., c_{i+3}" - "x_i, x_{i+1}, ..., x_{i+3}" \right|, \tag{5}$$

$$h_{16} = \sum_{i=1}^{32} \left| "c_i, c_{i+1}, ..., c_{i+15}" - "x_i, x_{i+1}, ..., x_{i+15}" \right|, \tag{6}$$

$$h_{32} = \sum_{i=1}^{16} \left| "c_i, c_{i+1}, ..., c_{i+31}" - "x_i, x_{i+1}, ..., x_{i+31}" \right|, \tag{7}$$

$$h_{64} = \sum_{i=1}^{8} \left| "c_i, c_{i+1}, ..., c_{i+63}" - "x_i, x_{i+1}, ..., x_{i+63}" \right|. \tag{8}$$

Множество процессоров с нечетным числом бит, на которых могут быть построены печатающие машинки для «стаи обезьян»

Все описанные выше конструкции легко согласуются с 2-, 4-, 8-, 16-, 32-, 64-битными процессорами под создаваемые для стаи обезьян печатающие машинки с конвертируемой друг в друга арифметикой. Однако «белый шум» не знает о людях и о том, что для «стаи обезьян» людям удобно писать процессоры печатающих машинок с системами счисления, кратными двум. Люди вполне могут пойти на отбрасывание малой части анализируемого кода и написать еще одну стаю процессоров под обезьян с 3 битами, 5 битами, 7 битами и т.д.:

$$h_3 = \sum_{i=1}^{85} \left| "c_i, c_{i+1}, c_{i+2}" - "x_i, x_{i+1}, x_{i+2}" \right|, \quad (9)$$

$$h_5 = \sum_{i=1}^{51} \left| "c_i, c_{i+1}, ..., c_{i+4}" - "x_i, x_{i+1}, ..., x_{i+4}" \right|, \quad (10)$$

$$h_7 = \sum_{i=1}^{36} \left| "c_i, c_{i+1}, ..., c_{i+6}" - "x_i, x_{i+1}, ..., x_{i+6}" \right|, \quad (11)$$

::

Наряду с 9 печатающими машинками, имеющими процессоры с четным числом бит, мы имеем дополнительно 247 печатающих машинок с процессорами, имеющими нечетное число разрядов, вполне пригодными для вычисления еще 247 сверток Хэмминга.

Мы имеем ситуацию, когда каждая обезьяна будет вооружена не одной, а 256 печатающими машинками, каждая из которых проверяет, не совпал ли набираемый «обезьяной» текст с осмысливаемыми эталонами на русском. Применяя подобную технологию, мы повышаем надежность контроля качества ключа (близость к эталонному «белому шуму») в 256 раз надежнее, чем обычный тест стаи обезьян, рекомендуемый NIST. Вместо того, чтобы рассматривать уровень случайности по этому тесту со стороны только одной пишущей машинки в 8-битной кодировке ASCII, мы одновременно рассматриваем ситуацию с точки зрения 256 пишущих машинок, каждая из которых работает в своей кодировке.

Естественно, что столь большое число наблюдателей избыточно. По рекомендациям NIST США обычно применяют 16 тестов на случайность. Видимо, при анализе качества ключей, полученных из нестабильной части биометрических данных, будет достаточно применения первых 16 пишущих машинок под каждую из обезьян.

Следует также отметить, что этот тип задач по созданию быстрых алгоритмов оценки энтропии длинных кодов применим достаточно широко. В частности, такой подход может улучшить решение обратных задач нейросетевой биометрии [15].

Библиографический список

1. Волчихин, В. И. Быстрые алгоритмы обучения нейросетевых механизмов биометрико-криптографической защиты информации : монография / В. И. Волчихин, А. И. Иванов, В. А. Фунтиков. – Пенза : Изд-во ПГУ. – 2005. – 273 с.

2. Малыгин, А. Ю. Быстрые алгоритмы тестирования нейросетевых механизмов биометрико-криптографической защиты информации / А. Ю. Малыгин, В. И. Волчихин, А. И. Иванов, В. А. Фунтиков. – Пенза : Изд-во ПГУ, 2006. –161 с.

3. Техническая спецификация [проект, публичное обсуждение начато с 01.02.2017 членами ТК 26 «Криптографическая защита информации»] Защита нейросетевых биометрических контейнеров с использованием криптографических алгоритмов.

4. Dodis, Y. Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy / Y. Dodis, L. Reyzin, A. Smith // Proc. EUROCRYPT, 2004. – April 13. – P. 523–540.

5. Monrose, F. Cryptographic key generation from voice / F. Monrose, M. Reiter, Q. Li, S. Wetzel // Proc. IEEE Symp. on Security and Privacy. – 2001. – P. 202–213.

6. Ramírez-Ruiz, J. Cryptographic Keys Generation Using FingerCodes / J. Ramírez-Ruiz, C. Pfeiffer, J. Nolazco-Flores //Advances in Artificial Intelligence – IBERAMIA-SBIA 2006 (LNCS 4140). – 2006. – P. 178–187.

7. Hao, F. Crypto with Biometrics Effectively / F. Hao, R. Anderson, J. Daugman // IEEE TRANSACTIONS ON COMPUTERS, – 2006. – Vol. 55, № 9. – P. 1073–1074.

8. Иванов, А. И. Нечеткие экстракторы: проблема использования в биометрии и криптографии / А. И. Иванов // Первая миля. – 2015. – № 1. – С. 40–47.

9. Иванов, А. И. Сопоставительный анализ показателей конкурирующих технологий биометрико-криптографической аутентификации личности / А. И. Иванов // Защита информации. ИНСАЙД. – 2014. – № 3. – С. 32–39.

10. Иванов, А. И. Среда моделирования «БиоНейроАвтограф». [Программный продукт создан лабораторией биометрических и нейросетевых технологий, размещен с 2009 г. на сайте АО «ПНИЭИ»] / А. И. Иванов, О. С. Захаров. – URL: http://пниэи.рф/activity/science/noc./bioneuroau-tugraph.zi

11. Иванов А. И. Автоматическое обучение больших искусственных нейронных сетей в биометрических приложениях : учеб. пособие к пакету лабораторных работ, выполняемых в среде моделирования «БиоНейроАвтограф» / А. И. Иванов. – Пенза : ОАО «ПНИЭИ». – 2013. – 27 с. – URL: http://пниэи.рф/activity/science/noc/tm_IvanovAI. pdf

12. Иванов, А. И. Многомерная нейросетевая обработка биометрических данных с программным воспроизведением эффектов квантовой суперпозиции / А. И. Иванов. – Пенза : АО «ПНИЭИ», 2016. – 133 с. – URL: http://пниэи.рф/activity/science/BOOK16.pdf

13. Волчихин, В. И. Многомерный портрет цифровых последовательностей идеального «белого шума» в свертках Хэмминга / В. И. Волчихин, А. И. Иванов, А. П. Юнин, Е. А. Малыгина // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2017. – № 4 (44). – С. 4–13. DOI 10.21685/2072-3059-2017-4-1.

14. Юнин, А. П. Оценка энтропии легко запоминаемых, длинных паролей со смыслом в ASCII кодировке для русского и английского языков / А. П. Юнин, О. В. Корнеев // Безопасность информационных технологий : тр. науч.-техн. конф. кластера пензенских предприятий. Т. 10. – Пенза. – 2016. – С. 40–42. – URL: http://пниэи.рф/activity/science/BIT/ T10-p40.pdf

15. Волчихин, В. И. Нейросетевая молекула: решение обратной задачи биометрии через программную поддержку квантовой суперпозиции на выходах сети искусственных нейронов / В. И. Волчихин, А. И. Иванов // Вестник Мордовского университета. – 2017. – Т. 27, № 4. – С. 518–523.

Юнин, А. П. Оценка качества «белого шума»: реализация теста «стаи обезьян» через множество сверток Хэмминга для разных систем счисления / А. П. Юнин, А. И. Иванов, К. А. Ратников // Безопасность информационных технологий : тр. I Всерос. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2019. – С. 10–18.

УДК 519.6

А. Г. Банных

НОМОГРАММА РЕГУЛЯРИЗАЦИИ ВЫЧИСЛЕНИЯ ЭНТРОПИИ ДЛИННЫХ КОДОВ, ПОЛУЧЕННАЯ ЧЕРЕЗ ОПИСАНИЕ БЕТА-РАСПРЕДЕЛЕНИЯ СТАТИСТИК РАССТОЯНИЙ ХЭММИНГА

Аннотация. Целью работы является повышение устойчивости вычисления энтропии длинных кодов с зависимыми разрядами за счет возможности учета не только нормальных законов распределения значений расстояния Хэмминга. Предложено использовать приближение реальных данных значений расстояний Хэмминга бета-распределением. Показано, что наиболее устойчивым является вычисление математического ожидания и стандартного отклонения расстояний Хэмминга. Приводится связь этих двух статистических моментов с параметрами бета-распределения. Решение системы двух нелинейных уравнений, связывающих упомянутые выше параметры, менее устойчиво. Программная реализация решения системы из двух нелинейных уравнений бета-распределения нерациональна.

A. G. Bannykh

NOMOGRAM OF REGULARIZATION OF ENTROPY CALCULATION OF LONG CODES, OBTAINED THROUGH THE DESCRIPTION OF THE BETA DISTRIBUTION OF HAMMING DISTANCE STATISTICS

Abstract. The aim of the work is to increase the stability of the calculation of the entropy of long codes with dependent discharges due to the possibility of taking into account not only the normal laws of the distribution of the Hamming distance values. It is proposed to use the approximation of the real data of Hamming distances by beta distribution. It is shown that the most stable is the calculation of the mathematical expectation and standard deviation of Hamming distances. The article cited the relationship of these two statistical moments with the parameters of the beta distribution. The solution of the system of two nonlinear equations connecting the parameters mentioned above is less stable. Software implementation of the solution of a system of two nonlinear equations of beta distribution is not rational.

Быстрый алгоритм вычисления энтропии в пространстве расстояний Хэмминга

Активные процессы информатизации современного общества приводят к необходимости массового использования средств криптографической защиты информации. В связи с тем, что пользователи не способны запоминать длинные пароли доступа и криптографические ключи в США и Евросоюзе развиваются технологии «нечетких экстракторов», преобразующих биометрический образ в короткий код доступа [1–3]. В России развивается технология нейросетевого преобразования биометрии в длинный код доступа или длинный код личного криптографического ключа [4, 5].

В связи с развитием нейросетевых технологий встает задача вычисления энтропии длинных кодов с зависимыми разрядами. Например, нейросеть среды моделирования «БиоНейроАвтограф» [6] дает код длиной 256 бит под применение алгоритмов отечественной криптографии. То есть актуальной является задача вычисления энтропии кодов длиной 256 бит.

Решить задачу вычисления энтропии кодов длиной 256 бит по Шеннону технически невозможно. По Шеннону придется вычислять сумму из 2^{256} слагаемых:

$$H("x_1, x_2, ..., x_{256}") = -\sum_{i=1}^{2^{256}} P_i \cdot \log_2(P_i). \qquad (1)$$

где P_i – вероятность появления одного из 2^{256} возможных состояний.

Проблема вычисления энтропии длинных кодов решается за счет перехода из пространства обычных кодов в пространство сверток Хэмминга [5–7]:

$$h = 256 - \sum_{i-1}^{256} ("x_i") \oplus ("c_i"), \qquad (2)$$

где "x_i" – значение i-го разряда свертываемого кода примеров образа «Чужой», "$\overline{c}$" – значение i-го разряда кода примеров образа «Свой».

Принципиально важным является то, что свертка Хэмминга из исходных 2^{256} состояний длинного кода создает гораздо меньшее число всего в 256 состояний. Еще одним важным фактом является то, свертка Хэмминга эффективно нормализует данные. Суммирование 256 случайных слагаемых по центральной предельной теореме статистики является эффективной процедурой нормализации.

Именно на эффекте нормализации построен ГОСТ Р 52633.3 [7]. По этому стандарту следует использовать порядка 100 случайно выбранных образов «Чужой». Схема численного эксперимента с использованием расстояний Хэмминга приведена на рис. 1.

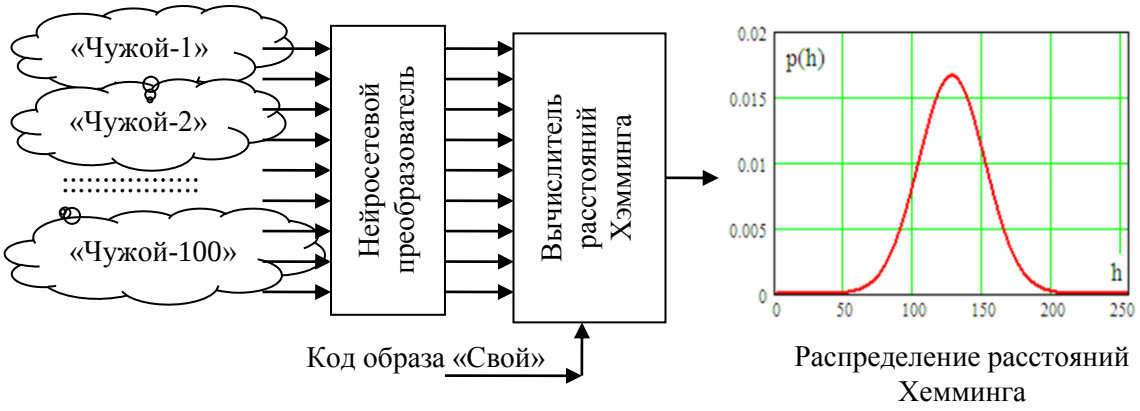


Рис. 1. Схема организации численного эксперимента по оценке вероятности ошибок второго рода (ошибочный пропуск «Чужого»)

Сотни расстояний Хэмминга вполне достаточно для вычисления достаточно надежной оценки математического ожидания – $E(h)$ и стандартного отклонения – $\sigma(h)$. По этим двум статистическим моментам мы можем оценить вероятность ошибок второго рода (ошибочное принятие «Чужого» как «Своего»).

$$P_2 = \frac{1}{\sigma(h)\sqrt{2\pi}} \int_{-\infty}^{1} \exp\left\{\frac{-(E(h)-u)^2}{2\cdot\sigma^2(h)}\right\}\cdot du\,. \qquad (3)$$

Если бы нейросетевой преобразователь биометрия-код был «идеальным», то энтропия его выходных кодов была бы точно ровна 256 битам (по длине кода). При полностью независимых разрядах кода $E(h)=128$, $\sigma(h)=8$, $P_2 = 2^{-256}$. По мере увеличения корреляционных связей между разрядами кода стандартное отклонение распределения расстояний Хэмминга увеличивается. При этом энтропия кода падает, а вероятность ошибок второго рода возрастает $P_2 << 2^{-256}$. Оценка энтропии может быть выполнена через значение вероятности ошибок второго рода:

$$H("x_1, x_2, ..., x_{256}") \approx -\log_2(P_2)\,. \qquad (4)$$

Использование аналитического описания бета-распределения для обработки статистик расстояний Хэмминга

Известно, что бета-распределение [8–10] может хорошо описывать не только нормальные распределения, но и иные асимметричные распределения. На рис. 2 даны примеры нормального (симметричного) и нескольких асимметричных распределений.

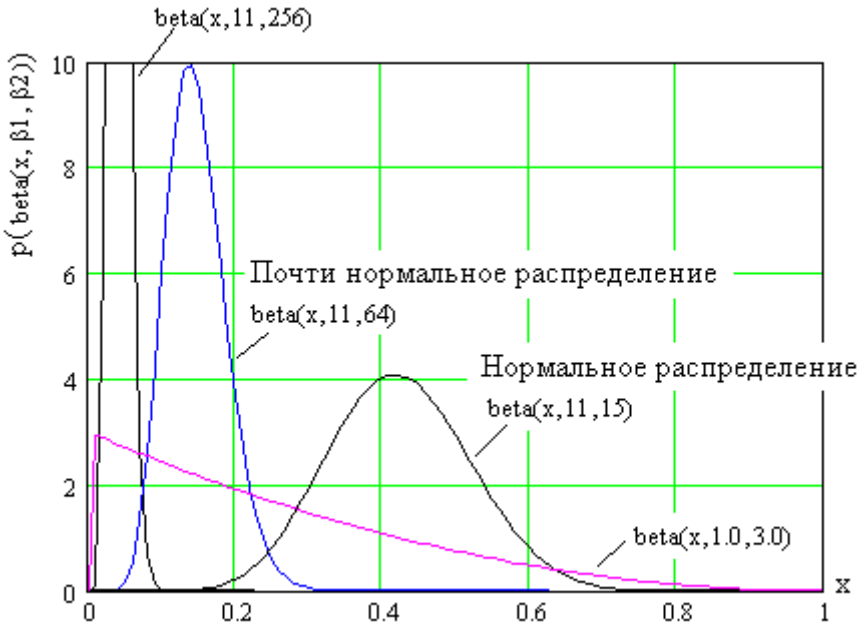


Рис. 2. Примеры симметричных и несимметричных распределений при разном соотношении двух параметрах бета-распределений

Из рис. 2 видно, что бета-распределение применимо только для значений переменной в интервале от 0 до 1. В случае использования расстояний Хэмминга переменная меняется в интервале от 0 до 256. То есть нам необходимо выполнить нормировку переменных:

$$x = \frac{h}{\max(h)}. \qquad (5)$$

В этом случае мы получим следующее аналитическое описание плотности распределения значений бета распределения:

$$p(beta(x, \beta 1, \beta 2)) = p(x, \beta 1, \beta 2) = \frac{(\beta 1 + \beta 2 + 1)!}{\beta 1! \cdot \beta 2!} \cdot x^{\beta 1} \cdot (1 - x)^{\beta 2}. \qquad (6)$$

Из статистической теории известно, что математическое ожидание бета распределений связано с его параметрами следующим соотношением:

$$E(x) = \frac{\beta 1 + 1}{\beta 1 + \beta 2 + 2}. \quad (7)$$

Стандартное отклонение бета распределения описывается иным уравнением:

$$\sigma^2(x) = \frac{(\beta 1 + 1)\cdot(\beta 2 + 1)}{(\beta 1 + \beta 2 + 2)\cdot(\beta 1 + \beta 2 + 3)}. \quad (8)$$

Воспользуемся уравнением (7), преобразуем его и выразим второй параметр бета-распределения через первый параметр и математическое ожидание:

$$\beta 2 = \frac{\beta 1 \cdot (1 - E(x)) - 1}{E(x)} - 2. \quad (9)$$

Подставив выражение (9) в (8), получим связь стандартного отклонения с первым параметром бета распределения:

$$\sigma(x) = \sqrt{\frac{(\beta 1 + 1)\cdot\left(\frac{\beta 1 \cdot (1 - E(x)) - 1}{E(x)} - 1\right)}{\left(\beta 1 + \frac{\beta 1 \cdot (1 - E(x)) - 1}{E(x)}\right)\cdot\left(\beta 1 + \frac{\beta 1 \cdot (1 - E(x)) - 1}{E(x)} + 1\right)}}. \quad (10)$$

По уравнению (10) может быть построена номограмма, связывающая между собой математическое ожидание и стандартное отклонение (рис. 3).

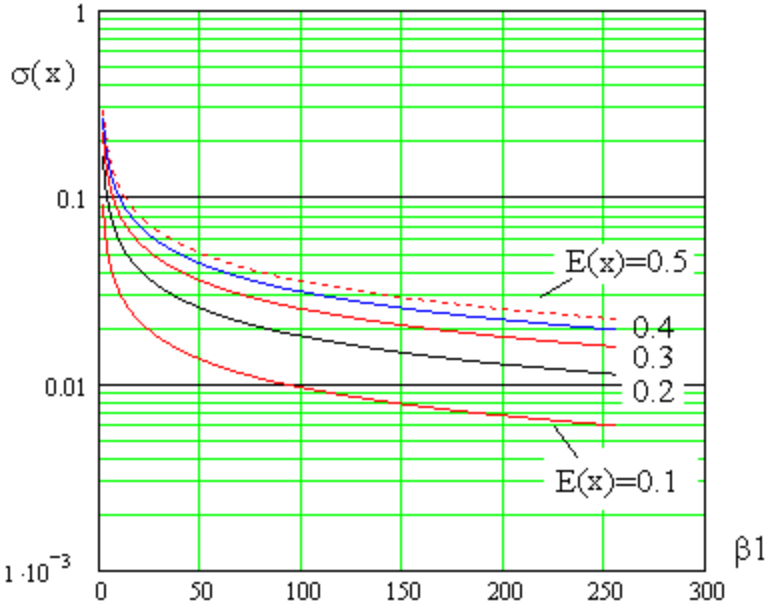


Рис. 3. Номограмма связи стандартного отклонения бета-распределения с его математическим ожиданием

Следует отметить, что номограмма рисунка 3 в реальном программном обеспечении должна быть представлена двухмерной таблицей значений первого параметра бета-распределения – $\beta 1$. Второй параметр бета-распределения вычисляется по формуле (9). После этого может быть оценена вероятность ошибок второго рода:

$$P_2 \approx \frac{(\beta 1+\beta 2+1)!}{\beta 1!\cdot\beta 2!}\cdot \int_0^{1/256} x^{\beta 1}\cdot(1-x)^{\beta 2}\cdot dx. \qquad (11)$$

Таким образом, мы получили номограмму, пользуясь которой можно вычислять энтропию длинных кодов с зависимыми разрядами для широкого класса асимметричных распределений расстояний Хэмминга. Точность вычислений и их устойчивость может быть высокой, так как опирается на точность, заранее вычисленных таблиц (номограммы).

Библиографический список

1. Juels, A. A Fuzzy Commitment Scheme / A. Juels, M. Wattenberg // Proc. ACM Conf. Computer and Communications Security. – Singapore, 1999. – November 1–4. – P. 28–36.

2. Ramírez-Ruiz, J. Cryptographic Keys Generation Using FingerCodes / J. Ramírez-Ruiz, r C. Pfeiffe, J. Nolazco-Flores // Advances in Artificial Intelligence – IBERAMIA-SBIA 2006 (LNCS 4140). – 2006 . – P. 178–187.

3. Hao, F. Crypto with Biometrics Effectively / F. Hao, R. Anderson, J. Daugman // IEEE TRANSACTIONS ON COMPUTERS. – 2006. – Vol. 55, № 9.

4. ГОСТ Р 52633.0–2006. Защита информации. Техника защиты информации. Требования к средствам высоконадежной биометрической аутентификации.

5. Нейросетевая защита персональных биометрических данных / Ю. К. Язов, В. И. Волчихин, А. И. Иванов, В. А. Фунтиков, И. Г. Назаров. – Москва : Радиотехника, 2012. – 157 с.

6. Иванов, А. И. Среда моделирования «БиоНейроАвтограф» [Программный продукт создан лабораторией биометрических и нейросетевых технологий] / А. И. Иванов, О. С. Захаров. – URL: http://пниэи.рф/activity/science/noc./bioneuroautugraph.zi

7. ГОСТ Р 52633.3–2011. Защита информации. Техника защиты информации. Тестирование стойкости средств высоконадежной биометрической защиты к атакам подбора.

8. Кобзарь, А. И. Прикладная математическая статистика. Для инженеров и научных работников / А. И. Кобзарь. – Москва : ФИЗМАТЛИТ, 2006. – 816 с.

9. Королюк, В. С. Справочник по теории вероятностей и математической статистике / В. С. Королюк, Н. И. Портенко, А. В. Скороход, А. Ф. Турбин. – Москва : Наука, 1985. – 640 с.

10. Иванов, А. И. Корректное квантово-континуальное преобразование данных, многократно ускоряющее оценку вероятности ошибок биометрической аутентификации личности / А. И. Иванов, А. В. Безяев, А. В. Елфимов, С. Е. Вятчанин // Специальная техника. – 2017. – № 1. – С. 48–51.

Банных, А. Г. Номограмма регуляризации вычисления энтропии длинных кодов, полученная через описание бета-распределением статистик расстояний Хэмминга / А. Г Банных // Безопасность информационных технологий : тр. I Всерос. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2019. – С. 19–25.

УДК 519.2; 519.6

Ю. И. Серикова

ДВОЙНАЯ РЕГУЛЯРИЗАЦИЯ ПРОЦЕДУР ОБУЧЕНИЯ НЕЙРОНОВ МАХАЛАНОБИСА ЗА СЧЕТ СИММЕТРИЗАЦИИ КОРРЕЛЯЦИОННЫХ СВЯЗЕЙ И КОМПЕНСАЦИИ ОШИБОК ВЫЧИСЛЕНИЯ КОЭФФИЦИЕНТОВ ПАРНОЙ КОРРЕЛЯЦИИ БИОМЕТРИЧЕСКИХ ДАННЫХ

Аннотация. Предложенные в работе процедуры регуляризации вычислений являются аналогами давно используемых процедур обучения нейронов. Регуляризация за счет симметризации корреляционных связей квадратичных форм является аналогом неитерационных алгоритмов автоматического обучения по ГОСТ Р 52633.5 нейронов с линейными свертками (персептронов).

Y. I. Serikova

DOUBLE REGULARIZATION OF LEARNING PROCEDURES FOR MAHALANOBIS NEURONS DUE TO SYMMETRIZATION OF CORRELATIONS AND COMPENSATION FOR ERRORS IN CALCULATING THE COEFFICIENTS OF PAIR CORRELATION OF BIOMETRIC DATA

Abstract. The proposed procedures for the regularization of computations are analogous to the long-used neuron learning procedures. Regularization due to symmetrization of correlation connections of quadratic forms is an analog of non-iterative algorithms of automatic learning in accordance with GOST R 52633.5 of neurons with linear convolutions (perceptrons). Iterative algorithms for learning neurons (perceptrons) for quadratic forms are the iterative algorithms for calculating correlation coefficients proposed in the article.

Нейроны с обогащением входных данных их линейным свертыванием

В настоящее время активно идут процессы информатизации общества. Предполагается, что значительный объем информации будет храниться с привлечением облачных сервисов. При этом важнейшим требованием к хранению электронных документов

в облаках является их надежная авторизация [1], а в некоторых случаях еще и обезличивание электронных документов [2]. Все эти требования удается выполнить, опираясь на использование нейросетевых преобразователей биометрических данных человека в код его личного криптографического ключа [3]. На данный момент нейросетевые преобразователи биометрия-код стандартизованы, для их автоматического обучения используется алгоритм, рекомендуемый ГОСТ Р 52633.5 [4]. Стандартизованный алгоритм обучения [4] абсолютно устойчив и имеет линейную вычислительную сложность, однако он ориентирован на использование нейронов с линейным обогащением входных биометрических данных. Формально все персептроны и обычные нейроны с иными гладкими функциями возбуждения следует рассматривать как линейную свертку пространства входных состояний с последующим нелинейным преобразованием уже свернутых (обогащенных) данных.

Нейроны с обогащением данных их свертыванием в квадратичном пространстве

Наряду с линейными свертками предварительного обогащения данных, в биометрии часто используются квадратичные свертки входных данных. За этими нейронами закрепилось название радиальных нейронов или радиально-базисных нейронов [5, 6]. Более корректно называть подобные конструкции нейронами квадратичных форм или нейронами Махаланобиса, эллиптическими нейронами. Описываются подобные конструкции следующей системой уравнений:

$$\begin{cases} e^2 = \left(\bar{\nu}\right)^T \cdot \left[R\right]^{=1} \cdot \bar{\nu} \\ z(e^2) = "0" \ \ при \ \ e^2 \le k \ , \\ z(e^2) = "1" \ \ при \ \ e^2 > k \end{cases} \qquad (1)$$

где k – порог срабатывания выходного квантователя нейрона; $\bar{\nu}$ – вектор нормированных и центрированных биометрических данных; $\left[R\right]$ – матрица корреляционных связей, контролируемых биометрических данных.

Проблема обучения нейронов Махаланобиса (1) связана с тем, что приходится вычислять коэффициенты корреляции на ма-

лых обучающих выборках в 20 примеров образа «Свой». Подобная задача корректна для нейронов Махаланобиса с 2-4 входами. Попытки увеличения входной размерности нейрона приводят к утрате корректности вычислений, в этом случае ошибки вычисления элементов корреляционной матрицы $[\Delta R]$ оказывают большее влияние на результат, чем влияние реальных значений элементов корреляционной матрицы.

Задача становится плохо обусловленной, при ее решении необходимо пользоваться регуляризацией, например, по Тихонову [7].

Симметризация корреляционных связей квадратичных нейронов

Одним из эффективных путей регуляризации вычислений является симметризация корреляционных связей обращаемой матрицы [8, 9]. Общий подход к такой регуляризации построен на том, что для каждого из нейронов выбирают данные, имеющие одинаковые коэффициенты корреляции:

$$\begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix}, \begin{bmatrix} 1 & r & r \\ r & 1 & r \\ r & r & 1 \end{bmatrix}, \begin{bmatrix} 1 & r & r & r \\ r & 1 & r & r \\ r & r & 1 & r \\ r & r & r & 1 \end{bmatrix}, \begin{bmatrix} 1 & r & r & r & r \\ r & 1 & r & r & r \\ r & r & 1 & r & r \\ r & r & r & 1 & r \\ r & r & r & r & 1 \end{bmatrix}, \begin{bmatrix} 1 & r & r & r & r & r \\ r & 1 & r & r & r & r \\ r & r & 1 & r & r & r \\ r & r & r & 1 & r & r \\ r & r & r & r & 1 & r \\ r & r & r & r & r & 1 \end{bmatrix}, \ldots \quad (2)$$

Выбрать одинаково коррелированные биометрические параметры из полной не симметричной в выше и ниже диагонали корреляционной матрицы достаточно легко. Размерность полной корреляционной матрицы реальных биометрических данных высока. Так, если пользоваться средой моделирования «БиоНейроАвтограф» [10], то вне диагонали полной корреляционной матрицы размерности 416x416 будут находиться 86 112 коэффициента корреляции. Пример распределения значений коэффициентов корреляции реальных биометрических данных приведен на рис. 1.

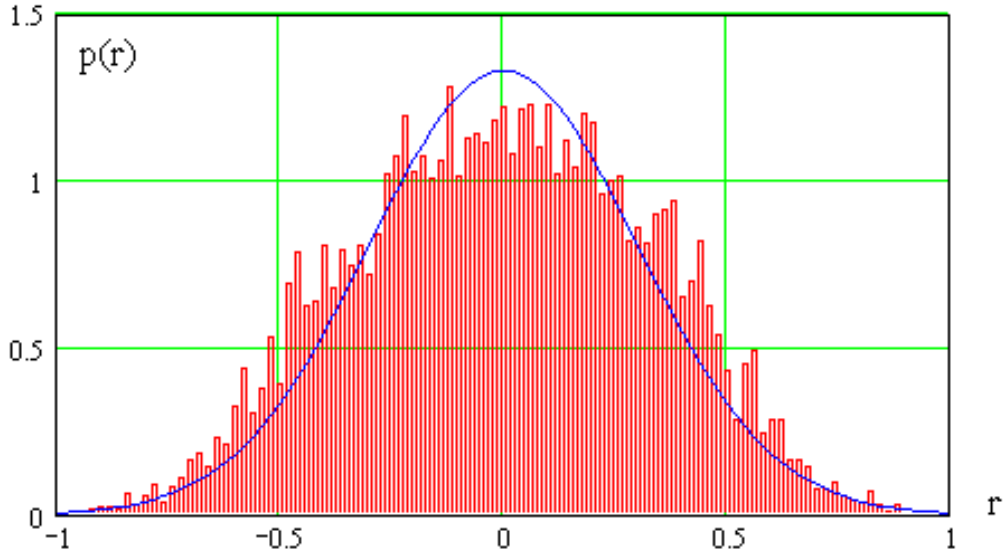


Рис. 1. Распределение значений коэффициентов корреляции биометрических данных, полученных в среде моделирования «БиоНейроАвтограф» [10] для рукописного слова *«Пенза»*

Идеальной является ситуация, когда выбираются данные, чьи коэффициенты корреляции близкие к нулю. В этом случае – мера Махаланобиса (1) будет иметь единичную корреляционную матрицу (проблема обращения корреляционной матрицы исчезает). Однако, по мере увеличения значений модулей коэффициентов равной корреляции вне диагонали матрицы ее коэффициент обусловленности увеличивается. Возникает иллюзия того, что обусловленность задачи обучения нейронов Махаланобиса быстро ухудшается (см. рис. 2), однако это не совсем так. Дело в том, что число обусловленности:

$$\mathrm{cond}[R] = \frac{\max(\lambda_i)}{\min(\lambda_i)}, \tag{3}$$

является отношением собственных чисел матрицы и отражает проблему неустойчивости вычислений только в первом приближении. При инженерных вычислениях число обусловленности можно рассматривать как коэффициент усиления ошибок исходных данных:

$$\Delta(e^2) \approx \mathrm{cond}\left[R\right] \cdot E\left(\left|\Delta r\right|\right). \tag{4}$$

Если число обусловленности растет (рис. 2), но одновременно снижается модуль ошибок вычисления исходных данных $\left|\Delta r\right| \to 0$, то устойчивость вычислений в целом сохраняется.

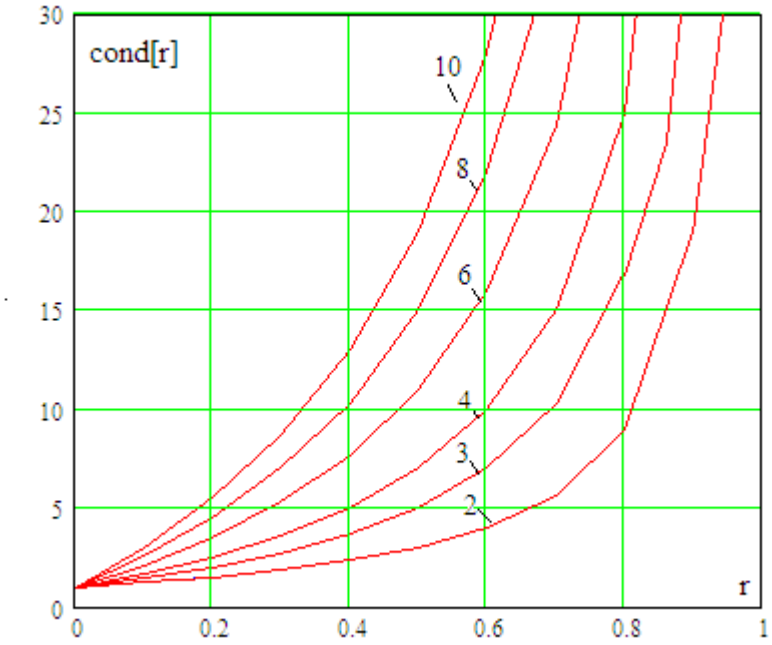


Рис. 2. Значения числа обусловленности симметричных матриц размерностей 2, 3, ..., 10, как функций равной коррелированности биометрических данных

Эта как раз та ситуация, которая наблюдается на практике. Ошибка вычислений коэффициентов корреляции из-за малого объема обучающей выборки оказывается наибольшей, когда данные слабо коррелированы. На рис. 3 приведено распределения значений коэффициентов корреляции на малых выборках из 7, 9, …, 21 примеров.

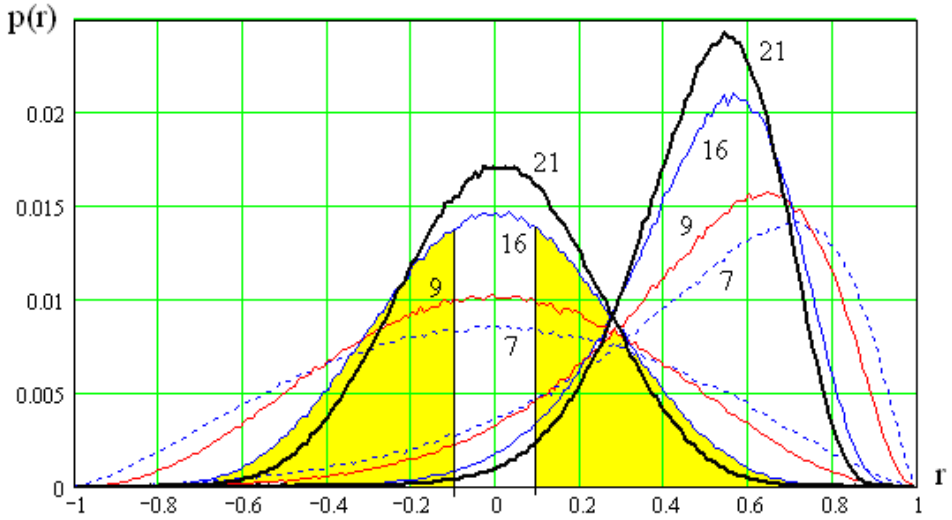


Рис. 3. Распределение значений коэффициентов корреляции на малых выборках для слабо коррелированных данных $E(r) = 0.0$ и сильно коррелированных данных $E(r) = 0.5$ при разных размерах обучающей выборки 7, 9,.., 21

Из рис. 3 видно, что при обучающей выборке в 16 примеров коэффициент корреляции попадает в интервал от –0.1 до +0.1 с вероятностью только 0.2. Однако интервал неопределенности Δr снижается по мере увеличения модуля коэффициента корреляции. Более того, при росте $|r| \to 1.0$ происходит монотонное снижение неопределенности ошибки вычисления $|\Delta r| \to 0.0$. Именно эти тенденции и являются тенденциями регуляризации вычислений в целом (4).

Регуляризация обучения за счет компенсации ошибок вычисления коэффициентов парной корреляции

Наиболее устойчивыми являются вычисления, когда из общей корреляционной матрицы выбираются значения $r_{1,i} = 0$. Например, эти значения могут выбираться из первой строки корреляционной матрицы, то есть мы выбираем параметры слабо коррелированные с первым биометрическим параметром. Так как коэффициенты корреляции вычислены с ошибкой, нет смысла находить их точно нулевое значение. Вполне достаточно группировать данные, для которых коэффициенты корреляции первой строки попадают в интервал от -0.05 до + 0.05. Если считать распределение коэффициентов корреляции биометрических данных нормальным (смотри рис. 1), то в первой строке полной корреляционной матрицы будет обнаружено не менее

$$(\text{pnorm}(0.05,0,1)- \text{pnorm}(-0.05,0,1)\cdot 416=16.589. \qquad (5)$$

слабо коррелированных параметров. Вычисления по формуле (5) выполнено в среде моделирования MathCAD. Это означает, что мы можем использовать нейроны Махаланобиса 17-го порядка. Номера в группе, выделяемых биометрических параметров оказываются монотонно увеличивающимися со случайным интервалом между соседями, например, $\{\nu_1, \nu_{15}, \nu_{41}, ..., \nu_{387}\}$. Следует упростить задачу, осуществив смену нумерации параметров, добившись обычного приращения номера параметра на единицу $\{\nu_1, \nu_3, \nu_3, ..., \nu_{17}\}$. В этом случае формальная запись меры Махаланобиса (1) не меняется, и будет выглядеть следующим образом:

$$e^2 = [\nu_1\ \nu_2\ ...\ \nu_{17}] \times \begin{bmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \end{bmatrix} \times \begin{bmatrix} \nu_1 \\ \nu_2 \\ ... \\ \nu_{17} \end{bmatrix}. \qquad (6)$$

Однако мы знаем, что все коэффициенты корреляции вычислены с погрешностью, то есть единичная обратная корреляционная матрица в выражении (6) должна иметь некоторые ошибки в элементах, находящихся вне диагонали. Мы имеем право, стабилизировать вычисления, скомпенсировав ошибки вычисления коэффициентов корреляции:

$$e^2 = \left[\nu_1\ \nu_2\ ....\ \nu_{17}\right] \times \begin{bmatrix} 1 & \Delta_{1,2} & ... & \Delta_{1,17} \\ \Delta_{1,2} & 1 & ... & \Delta_{2,17} \\ ... & ... & ... & ... \\ \Delta_{1,17} & \Delta_{2,17} & ... & 1 \end{bmatrix} \times \begin{bmatrix} \nu_1 \\ \nu_2 \\ ... \\ \nu_{17} \end{bmatrix}. \qquad (7)$$

Для того, что бы скомпенсировать ошибки вычислений коэффициентов корреляции необходимо подобрать значения 128-ти мерного компенсатора $\left\{\Delta_{1,2,},\Delta_{1,3},...,\Delta_{1,17},\Delta_{2,3},\Delta_{2,4},...,\Delta_{2,17},\Delta_{3,4},...,\ ,...,\Delta_{16,17}\right\}$. Эта операция выполняется одной из итерационных процедур направленного подбора, осуществляющей поиск максимума следующего показателя качества:

$$\max\left\{q = \frac{\left|E(e^2(\bar{\nu})) - E(e^2(\bar{\xi}))\right|}{\sqrt{\sigma(e^2(\bar{\nu}))\cdot\sigma(e^2(\bar{\xi}))}}\right\}, \qquad (8),$$

где $\bar{\nu}$ – векторы всех примеров обучающей выборки образа «Свой», $\bar{\xi}$ – векторы всех примеров обучающей выборки образов «Чужие».

В случае если мы выбираем биометрические данные с равной коррелированностью r = 0.05, то мы получим похожую метрику Махаланобиса:

$$e^2 = \left[\nu_1\ \nu_2\ ...\ \nu_{17}\right] \times \begin{bmatrix} 1 & 0.05 & ... & 0.05 \\ 0.05 & 1 & ... & 0.05 \\ ... & ... & ... & ... \\ 0.05 & 0.05 & ... & 1 \end{bmatrix}^{-1} \times \begin{bmatrix} \nu_1 \\ \nu_2 \\ ... \\ \nu_{17} \end{bmatrix}. \qquad (9)$$

Обращение симметричной корреляционной матрицы в (9) приводит к получению матрицы с аналогичной симметрией:

$$\begin{bmatrix} 1 & r & ... & r \\ r & 1 & ... & r \\ ... & ... & ... & ... \\ r & r & ... & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & \pi & ... & \pi \\ \pi & 1 & ... & \pi \\ ... & ... & ... & ... \\ \pi & \pi & ... & 1 \end{bmatrix}. \qquad (10)$$

Это означает, что компенсацию ошибок коэффициентов корреляции общей формы меры Махаланобиса, следует выполнять для следующей записи:

$$e^2 = \left[\nu_1\ \nu_2\ ...\ \nu_{17}\right] \times \begin{bmatrix} 1 & \pi+\Delta_{1,2} & ... & \pi+\Delta_{1,17} \\ \pi+\Delta_{1,2} & 1 & ... & \pi+\Delta_{2,17} \\ ... & ... & ... & ... \\ \pi+\Delta_{1,17} & \pi+\Delta_{2,17} & ... & 1 \end{bmatrix} \times \begin{bmatrix} \nu_1 \\ \nu_2 \\ ... \\ \nu_{17} \end{bmatrix}. \quad (11)$$

При использовании любых симметризованных корреляционных матриц суть регуляризации через подбор стабилизирующей вычисления матрицы не меняется. Для любой симметризованной матрицы одинаковых коэффициентов корреляции – r следует искать максимум функционала качества (8) вблизи значений элементов обратной корреляционной матрицы – π.

Из приведенных выше аргументов следует, что регуляризация обучения нейронов Махаланобиса должна быть двухступенчатой. Первая ступень регуляризации сводится к выбору одинаково коррелированных между собой биометрических параметров. Только эта процедура уже позволяет значительно снизить число обусловленности, решаемой задачи. Вторым этапом обучения является компенсация ошибок вычисления коэффициентов корреляции, обусловленная недостаточным объемом выборки примеров образа «Свой». Данную регуляризацию можно рассматривать как попытку скомпенсировать ошибки вычисления коэффициентов корреляции из-за малого объема примеров обучающей выборки образа «Свой».

Следует подчеркнуть, что первый этап регуляризации имеет квадратичную вычислительную сложность, так как построен на вычислении полной корреляционной матрицы. Оценка вычислительной сложности второго этапа регуляризации не определена, так как все итерационные процедуры подбора весовых коэффициентов нейронов имеют меняющуюся в процессе обучения вычислительную сложность. В начале обучения вычислительная сложность близка к линейной, однако ее показатель быстро увеличивается в плоть до экспоненциальной вычислительной сложности. Вторая часть регуляризации по своей сути является одним из вариантов «обычного» подбора весовых коэффициентов у нейрона.

Библиографический список

1. Ложников, П. С. Биометрическая защита гибридного документооборота / П. С. Ложников. – Новосибирск : Изд-во СО РАН, 2017. – 130 с.

2. Гулов, В. П. Перспектива нейросетевой защиты облачных сервисов через биометрическое обезличивание персональной информации на примере медицинских электронных историй болезни / В. П. Гулов, А. И. Иванов, Ю. К. Язов, О. В. Корнеев // Вестник новых медицинских технологий (JORNAL OF NEW MEDICAL TECHNOLOGIEC). – 2017. – Т. 24, № 2 (июнь), 2017. – С. 220–225.

3. Язов, Ю. К. Нейросетевая защита персональных биометрических данных / Ю. К. Язов, В. И. Волчихин, А. И. Иванов, В. А. Фунтиков, И. Г. Назаров. – Москва : Радиотехника, 2012. – 157 с.

4. ГОСТ Р 52633.5–2011. Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа.

5. Хайкин, С. Нейронные сети : полный курс / Саймон Хайкин. – Москва : Вильямс, 2006. – С. 1104.

6. Руд, Б. Руководство по биометрии : пер. с англ. / Болл Руд, Коннел Джонатан Х., Панканти Шарат, Ратха Налини К., Сеньор Эндрю У. – Москва : Техносфера, 2007. – 368 с.

7. Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач. Москва : Наука, 1979. – 248 с.

8. Волчихин, В. И. Фрактально-корреляционный функционал, используемый при поиске пар слабо зависимых биометрических данных в малых выборках / В. И. Волчихин, А. И. Иванов, Б. Б. Ахметов, Ю. И. Серикова // Вестник высших учебных заведений. Поволжский регион. Технические науки. – 2016. – № 4. – С. 25–31.

9. Ivanov, A. I. Reducing the Size of a Sample Sufficient for Learning Due to the Symmetrization of Correlation Relationships Between Biometric Data / A. I. Ivanov, P. S. Lozhnikov, Yu. I. Serikova // Cybernetics and Systems Analysis. – 2016. – № 3, May–June. – P. 49–56. – URL: http://link.springer.com/article/10.1007/s10559-016-9838-x

10. Иванов, А. И. Среда моделирования «БиоНейроАвтограф» [Программный продукт размещен с 2009 года на сайте АО «ПНИЭИ»] / А. И. Иванов, О. С. Захаров. – URL: http://пниэи.рф/activity/science/noc./bio-neuroautugraph.zi

Серикова, Ю. И. Двойная регуляризация процедур обучения нейронов Махаланобиса за счет симметризации корреляционных связей и компенсации ошибок вычисления коэффициентов парной корреляции биометрических данных / Ю. И. Серикова // Безопасность информационных технологий : тр. I Всерос. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2019. – С. 26–34.

УДК 519.2**;** 519.6

И. Г. Монахова, А. В. Майоров

ПОЛОЖИТЕЛЬНЫЙ ОПЫТ СОЗДАНИЯ ПЕРВОГО В МИРОВОЙ ПРАКТИКЕ БИОМЕТРИЧЕСКОГО УДОСТОВЕРЯЮЩЕГО ЦЕНТРА ПО ГРАНТУ ПОДДЕРЖКИ ПРАВИТЕЛЬСТВА ПЕНЗЕНСКОЙ ОБЛАСТИ В 2012–2013 гг.

Аннотация. Рассматривается проблема документирования последовательности событий инициировавших существенное развитие биометрических технологий по гранту правительства Пензенской области. Решить проблему возможно в случае перехода к биометрическим удостоверяющим центрам, которые дополнительно к криптографическим протоколам поддерживают протоколы биометрической аутентификации личности. Приводится описание использованных технических решений. Практическая реализация первого в мировой практике биометрического удостоверяющего центра в 2012 г. подтвердила теоретические положения об усилении уровня авторизации криптографических процедур формирования цифровой подписи.

G. Monakhova, A. V. Mayorov

POSITIVE EXPERIENCE OF CREATING THE WORLD'S FIRST BIOMETRIC CERTIFICATION CENTER FOR A GRANT TO SUPPORT THE GOVERNMENT OF THE PENZA REGION IN 2012–2013

Abstract. The problem of documenting the sequence of events that initiated a significant development of biometric technologies under the grant of the Government of the Penza Region is considered. It is possible to solve the problem in case of transition to biometric certification centers, which, in addition to cryptographic protocols, support biometric identity authentication protocols. A description of the used technical solutions. The practical implementation of the first in world practice biometric certification center in 2012 confirmed the theoretical provisions on strengthening the level of authorization of cryptographic procedures for the formation of digital signatures.

Существующая система поддержки электронной цифровой подписи

В настоящее время формирование электронной подписи под электронным документом является базовым технологическим элементом электронного документооборота и значимого Интернет взаимодействия. Сегодня электронная подпись нужна для того, чтобы:

- участвовать в закупках по 44-ФЗ, 223-ФЗ;
- получить доступ на коммерческие площадки B2B-center, Газпромбанк, ТЭК-торг, uTender, Фабрикант и др.;
- предоставлять отчетность в электронной форме в Федресурс, ЦБ РФ, ЛесЕГАИС, ФГИС ЦС и др.;
- взаимодействовать с государством через Госуслуги, портал государственной экспертизы, Росаккредитация и др.;
- зарегистрировать онлайн-кассу в ФНС.

Список услуг, предполагающих у их потребителей электронной подписи, будет только расширяться. Для того чтобы стать потребителем подобных услуг необходимо получить лицензию на использование программного продукта КриптоПро, желательно иметь носитель личного ключа (например, РутокенЛайт) и обязательно нужно зарегистрировать свой открытый ключ в Удостоверяющем Центре (получить сертификат своего открытого ключа). Все эти услуги платные.

И это только открытая часть айсберга значимого для пользователей электронного документооборота. Когда мы выходим через Интернет на сервер значимой для нас организации (например, на сервер электронной торговли или на сервер своего банка), то мы пользуемся защищенным Интернет соединением по протоколам SSL/TLS [1], которые организованы с использованием ассиметричной криптографии. Асимметричная криптография предполагает использование сертификатов открытых ключей, ранее зарегистрированных Интернет провайдерами в соответствующих международных удостоверяющих центрах.

Получается, что уже сейчас мы незаметно для себя активно используем многократно вложенную друг в друга асимметричную криптографию, опирающуюся на систему Удостоверяющих Центров. Когда нам требуется использование защищенных каналов Интернет соединений, то мы пользуемся протоколам SSL/TLS и сертификатами открытых ключей Интернет оборудования, выдан-

ными международными удостоверяющими центрами. Когда речь идет о проверке цифровой подписи под электронным документом, то мы должны использовать отечественную криптографию [2] при формировании цифровой подписи и сертификаты открытых ключей, выданные одним из национальных удостоверяющих центров России.

Проблема усиления уровня авторизации цифровой подписи

Следует отметить, что Федеральный закон № 63 [3] «О цифровой подписи» и его криптографическая реализация [2] построены на предположениях о корректной мотивации владельца цифровой подписи и его окружения. Например, владелец корпоративной ЭЦП не передаст свои полномочия своему «сослуживцу» на время своего отсутствия на рабочем месте.

Для реализации подобного злоупотребления сегодня достаточно отдать «сослуживцу» носитель ключа (например, USB-токен и сообщить «сослуживцу» пароль доступа к личному ключу). При этом существующие средства контроля не могут обнаружить факт сговора (подмену легитимного владельца ЭЦП). Эта угроза устраняется, если привязать доступ к криптографическому ключу через предъявление биометрического образа легитимного владельца ЭЦП [4].

Из рис. 1 видно, что по патенту [4] получить личный ключ можно только в том случае, когда предъявлен не только короткий ПИН код (или пароль доступа), но и легальный пользователь предъявил свой биометрический образ (например, рукописный пароль «*Пенза*»). Внутри USB-токена должна находиться не только программа криптографического формирования ЭЦП пользователя, но и нейросетевой преобразователь биометрия-код фрагмента личного ключа.

Пользователь уже может не попытаться передать право формирования цифровой подписи своему сослуживцу, так как у «сослуживца» будет другая биометрия. В USB-токенах могут быть реализованы любые биометрические технологии:

- анализ рисунка отпечатка пальца;
- анализ голосового пароля;
- анализ геометрии лица пользователя;
- анализ контуров руки;
- анализ рисунка подкожных кровеносных сосудов;
- анализ рисунка радужной оболочки глаза.

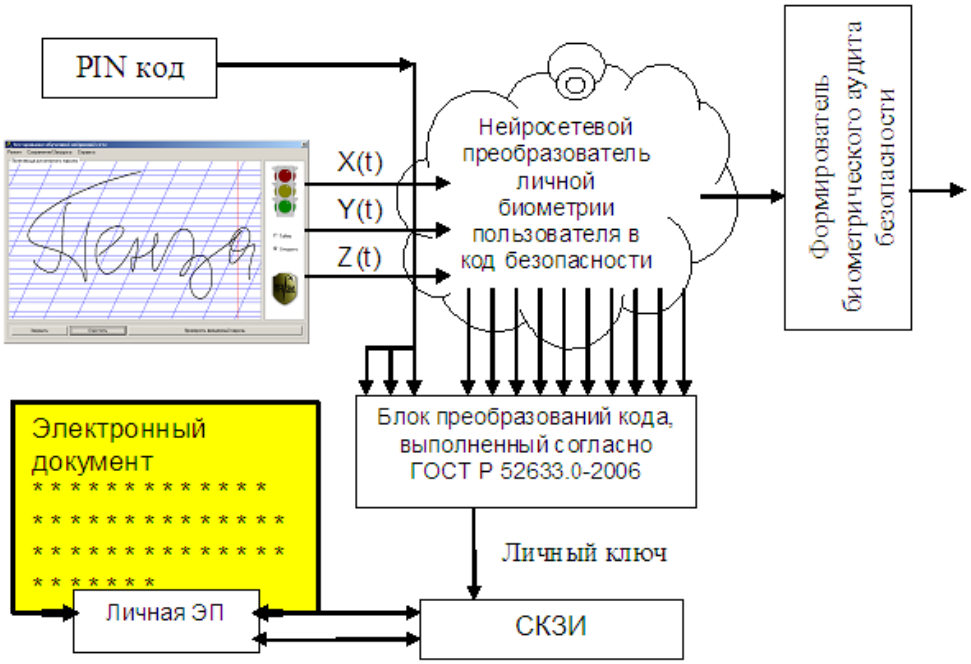


Рис. 1. Биометрико-нейросетевой криптоформирователь личной ЭП пользователя с высокой степенью авторизации

Важнейшим элементом информационной безопасности является непрерывный контроль за оборотом биометрических данных в корпоративной информационной системе [5]. Вторым важным элементом обеспечения биометрического контроля является устранение возможности сознательного ослабления своей биометрии пользователем, то есть обучать свою нейронную сеть по ГОСТ Р 52633.5 [6] он должен только под контролем. После обучения нейронной сети должно быть выполнено ее тестирование по ГОСТ Р 52633.3 [7]. То есть до размещения данных обученной нейронной сети в доверенную вычислительную среду USB-БиоТокен должны быть выполнены несколько важных операций. Выполнять их может только некоторый центр доверия. Например, таким центром доверия может стать БиоУЦ. То есть к функциям обычного Удостоверяющего Центра (УЦ) должны быть добавлены дополнительные функции.

Дополнительные функции биометрического удостоверяющего центра

В табл. 1 даны основные функции обычного удостоверяющего центра (левая часть таблицы) и дополнительные функции БиоУЦ (права часть таблицы).

Таблица 1

№	Функции обычного УЦ	№	Дополнительные функции БиоУЦ
1	Создание пары из открытого ключа и личного ключа	7	Контроль качества обучения нейросетевого преобразователя
2	Создание сертификата открытого ключа	8	Уничтожение личного ключа и примеров БиоОбраза «Свой», на которых обучалась нейросеть
3	Размещение личного ключа в контейнер парольной защиты доступа	9	Размещение обученного нейросетевого преобразователя в USB-БиоТокен
4	Уничтожение личного ключа	10	Осуществление сбора БиоАудита информационной безопасности значимых биометрических действий
5	Контроль действия сертификатов открытых ключей (запрет использования устаревших и аннулированных сертификатов)	11	Обнаружение атак на биометрическую защиту электронного документооборота
6	Обучение нейросетевого преобразователя биометрия-код выдавать код личного ключа При предъявлении БиоОбраза «Свой»	12	Поддержка режима обезличенной биометрической аутентификации, выпуск обезличенных сертификатов открытого ключа

Положительный опыт создания первого в мировой практике биометрического удостоверяющего центра емкостью до 300 тысяч пользователей по гранту правительства Пензенской области

Обычно совершенствования той или иной технологии осуществляются постепенно. Из табл. 1 мы видим, что переход от обычных УЦ к БиоУЦ приводит с одной стороны к повышению уровня авторизации доступа к личному ключу формирования электронной цифровой подписи, а с другой стороны является достаточно сложной технической задачей. О ее сложности можно судить хотя бы потому, что число поддерживаемых БиоУЦ функции увеличивается с 5 до 12 по сравнению с обычным УЦ.

Первый БиоУЦ был создан в 2013 г. ООО «Биометрика» [8], Заказчиком работы являлся АО «Пензенский научно-исследовательский электротехнический институт», Правительство Пензенской области выделило грант на покупку оборудования для

проведения опытно-конструкторских работ. Основные требования к опытному образцу приведены ниже.

1.1. Программное обеспечение должно обеспечивать выполнение следующих функций:

1) регистрация или перерегистрация пользователя удостоверяющего центра с использованием его биометрических данных и формированием биометрических контейнеров;

2) идентификация пользователя удостоверяющего центра в режиме открытого имени и в режиме обезличенности;

3) аутентификация пользователя удостоверяющего центра в режиме открытого имени и в режиме обезличенности с использованием биометрических технологий;

4) формирование базы биометрических контейнеров с гарантированной защитой от раскрытия обезличенности;

5) тестирование биометрических контейнеров в автоматическом и ручном режимах с учётом требований ГОСТ Р 52633.0–2006 с целью определения стойкости к атакам подбора при реализации различных видов атак и вывод полученных результатов;

1.2. Программное обеспечение должно обеспечивать следующие динамические характеристики работы:

1) время удалённой биометрической аутентификации пользователя при передаче им всех необходимых данных не должно превышать 3-х секунд;

2) время настройки параметров преобразования биометрия-код во время регистрации пользователя при передаче им всех необходимых данных не должно превышать 30 секунд;

3) время преобразования биометрия-код (без учёта времени ввода биометрических данных) не должно превышать 1 с.;

4) время настройки преобразователя биометрия-код (без учёта времени ввода биометрических данных) не должно превышать 5 с.

1.3. ПО СБА (Средства Биометрической Аутентификации) должно обеспечивать высоконадёжную биометрическую аутентификацию и идентификацию пользователя (оператора) путём анализа его рукописного пароля или рисунка отпечатка пальца, выполненных по требованиям ГОСТ Р 52633.

1.4. ПО СБА должно обеспечивать возможность использования нескольких биометрических технологий аутентификации, поддерживаемых ПО КД.

1.5. ПО КД должен поддерживать не менее двух технологий биометрической аутентификации (рукописный образ, отпечаток пальца), а также режим аутентификации по паролю.

1.6. ПО КД должно поддерживать схемы объединения нескольких биометрических образов. Объединение нескольких биометрических образов для связывания с заданным кодом должно выполняться с помощью логических операций «И» и «ИЛИ».

1.7. Программное обеспечение должно поддерживать следующие варианты взаимодействия:

1) обращения ПО СБА за информацией о сертификатах ключей ЭЦП на ПО УЦ;

2) обращения ПО КД к ПО СБА на биометрическую аутентификацию;

3) связывание личного ключа или биометрического ключа с биометрическими данными пользователя;

4) изменение биометрических данных пользователя, используемых для связи с выданными сертификатами ключей;

5) подтверждение пользователя в интересах третьих лиц;

6) сопоставление сертификата ключа и авторизованного лица;

7) выполнение операций по формированию ЭЦП с использованием высоконадёжной биометрической аутентификации.

1.8. Программное обеспечение должно обеспечивать безопасное хранение и использование конфиденциальной биометрической информации пользователя (оператора) в защищённом нейросетевом биометрическом контейнере и обеспечивать, при необходимости, анонимность или обезличенность пользователя.

1.9. Программное обеспечение должно автоматически «обучаться». Число примеров обучения (для одной биометрической технологии) должно составлять от 8 до 32 биометрических образов пользователя (оператора).

1.10. Число попыток аутентификации устанавливается политиками безопасности.

1.11. При использовании для ввода рукописного образа графического планшета с чувствительностью 2000 точек на дюйм программное обеспечение должно обеспечивать:

1) вероятность ошибочного пропуска «Чужого» с первой попытки:

– менее 10^{-4} для лиц с нестабильным почерком;

– менее 10^{-6} для лиц со стабильным почерком;

– менее 10^{-8} для лиц с высокостабильным почерком;

2) вероятность ошибочного отказа «Своему»:

– не более 0,4 с первой попытки;

– не более 0,2 при двух попытках;

– не более 0,1 при трёх попытках.

1.12. При использовании сканера отпечатков пальцев с разрешающей способностью изображения не менее 480x320 для ввода рисунков отпечатка пальцев программное обеспечение должно обеспечивать:

1) вероятность ошибочного пропуска «Чужого»” с первой попытки не менее 10^{-6} для среднестатистического пользователя.

2) вероятность ошибочного отказа «Своему»:

– не более 0,4 с первой попытки;

– не более 0,2 при двух попытках;

– не более 0,1 при трёх попытках.

1.13. Срок службы программного обеспечения должен составлять не менее 5 лет.

Заключение

Проведенная опытно-конструкторская работа (шифр ОКР «Удостоверение-Э») позволила создать первый в мировой практике БиоУЦ. Его опытная эксплуатация подтвердила основные теоретические положения о значительном повышении уровня авторизации доступа к личному ключу формирования электронной цифровой подписи. Электронная Цифровая Подпись с применением биометрической поддержки перестает быть Цифровой Печатью, которую один пользователь при желании может передать другому пользователю. Видимо, биометрическая поддержка доступа к личному ключу формирования корпоративной электронной цифровой подписи будет в ближайшем будущем одной из основных тенденций развития электронного корпоративного документооборота.

Библиографический список

1. Смит, Р. Э. Аутентификация от паролей до открытых ключей / Р. Э. Смит. – Москва : Вильямс, 2002. – 423 с.

2. ГОСТ Р 34.10–2012. Информационная технология. Криптографическая защита информации. Процессы формирования и проверки электронной цифровой подписи.

3. Об электронной подписи : федер. закон № 63 от 06.04.2011.

4. Пат. 2365047 Российская Федерация. Способ формирования электронных документов и устройство для его реализации / Иванов А. И., Фунтиков В. А. ; приоритет от 04.06.2007 ; опубл. 20.08.2009, Бюл. № 23.

5. Пат. 2427921 Российская Федерация. Способ формирования аудита персональной биометрической информаци / Язов Ю. К., Иванов А. И., Наза-

ров И. Г., Фунтиков В. А., Иванов С. М., Трифонов С. Е., Ефимов О. В. ; приоритет от 26.01.2010 ; опубл. 27.08.2012, Бюл. № 24.

6. ГОСТ Р 52633.5–2011. Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа.

7. ГОСТ Р 52633.3–2011. Защита информации. Техника защиты информации. Тестирование стойкости средств высоконадежной биометрической защиты к атакам подбора.

8. ОКР «Разработка удостоверяющего центра емкостью до 300 тысяч пользователей и клиентского программного обеспечения с предоставлением расширенных услуг по биометрической аутентификации» (шифр ОКР «Удостоверение-Э») выполнен по гранту правительства Пензенской области в 2012–1013 гг. Разработчик ООО «Биометрика».

Монахова, И. Г. Положительный опыт создания первого в мировой практике биометрического удостоверяющего центра по гранту поддержки правительства Пензенской области в 2012–2013 гг. / И. Г. Монахова, А. В. Майоров // Безопасность информационных технологий : тр. I Всерос. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2019. – С. 35–43.

УДК 519.2; 519.6

Е. А. Малыгина, С. Е. Вятчанин, А. И. Солопов

УСИЛЕНИЕ УРОВНЯ ЗАЩИЩЕННОСТИ БИОМЕТРИЧЕСКИХ ТЕХНОЛОГИЙ ЗА СЧЕТ ИСПОЛЬЗОВАНИЯ НЕЙРОНОВ КРАМЕРА – ФОН МИЗЕСА

Аннотация. Рассматривается перспектива использования нейронов Крамера – фон Мизеса в биометрико-нейросетевых механизмах защиты в современных инфокоммуникационных системах обработки данных в облачных сервисах. Показано, что мощность критерия Крамера – фон Мизеса на малых выборках примеров биометрических данных оказывается существенно выше, чем мощность аналогичного критерия хи-квадрат, что делает их крайне перспективными для применения в нейросетевых преобразователях биометрия-код.

E. A. Malygina, S. E. Vyatchanin, A. I. Solopov

STRENGTHENING THE LEVEL OF PROTECTION OF BIOMETRIC TECHNOLOGIES THROUGH THE USE OF KRAMER – VON MISES NEURONS

Abstract. The prospect of using Kramer – von Mises neurons in biometric-neural network protection mechanisms in modern infocommunication data processing systems in cloud services is considered. It is shown that the power of the Kramer-von Mises criterion on small samples of examples of biometric data is significantly higher than the power of the analogous chi-square criterion, which makes them extremely promising for use in neural network biometrics-code converters.

Становление, развитие и защита российской цифровой экономики невозможна без активного использования отечественной криптографии в современных инфокоммуникационных системах обработки данных в облачных сервисах.

К сожалению, пользователи современных инфокоммуникационных систем не в состоянии запоминать длинные пароли доступа к информационным ресурсам. Эту проблему предложено решить при помощи биометрических данных самого пользователя данных сервисов.

В США, Канаде, Евросоюзе развивается технология так называемых «нечетких» экстракторов [1–3]. Технология сводится к тому, что из биометрического образа извлекаются его параметры, далее их значения квантуются. Как отмечено в зарубежных научных публикациях полученный БиоКод может содержать от 20 до 30 % ошибок. Для их устранения используют маскирование наиболее нестабильных бит БиоКода, не устраненные ошибки обнаруживаются и исправляются классическими самокорректирующимися кодами, обладающими примерно 20 кратной избыточностью. Таким образом длина БиоКода оказывается примерно в 20 меньше числа биометрических параметров, извлеченных из образа. Поэтому БиоКоды на выходе «нечетких» экстракторов оказываются недопустимо короткими для того, чтобы далее использовать полноценные криптографические механизмы.

В России создается, стандартизуется и поддерживается технология нейросетевого преобразования биометрических данных в код заранее полученного криптографического ключа [4, 5]. Основной проблемой применения искусственных нейронных сетей является то, что они очень медленно учатся и для их обучения [6] необходимы огромные обучающие выборки, содержащие порядка 100 000 примеров биометрических образов. Медленное обучение и большие выборки обучающих примеров для обучения «глубоких» нейронных сетей [6] необходимы из-за того, что итерационный алгоритм «обратного распространения ошибки» имеет экспоненциальную вычислительную сложность.

Выходом из тупика экспоненциальной вычислительной сложности обучения «глубоких» нейронных сетей стало решение использовать однослойные искусственные нейронные сети с большим числом выходов. Теоретически и практически получены результаты, показывающие, что для однослойных искусственных нейронных сетей вычислительная сложность обучения оказывается линейной [7]. Это позволило их обучение на выборке из 20 примеров образа «Свой». При этом время обучения оказывается незначительным (от 0.5 до 0.05 секунды) даже при использовании ПЭВМ с процессорами низкой производительности.

Основная идея быстрых стандартизованных алгоритмов не итерационного обучения сводится к тому, что обучение линейной части нейрона (сумматора) и нелинейной части нейрона (квантователя) разделены. Декомпозиция задачи на две простых подзадачи (линейную свертку в пространстве входных состояний и нелиней-

ное преобразование выходных данных свертки [8]) всегда приводит к значительному упрощению вычислений. В такой декомпозиции задачи построен алгоритм обучения, регламентированный стандартом ГОСТ Р 52633.5 [8]. Прозрачная парадигма разделения задачи обучения нейронной сети на две составляющих (линейную и нелинейную) дает очень хорошие результаты по сокращению размеров обучающей выборки и сведению задачи до линейной вычислительной сложности. Тем не менее, этот конструктивный прием, не является универсальным, вполне могут существовать и другие варианты декомпозиции задачи, приводящие к близким или даже более качественным результатам.

В настоящее время существует множество статистических критериев для поверки гипотезы нормального распределения значений биометрических данных. Наиболее часто на практике используется хи-квадрат критерий [9, 10] в силу того, что для этого критерия Пирсон в 1900 году построил аналитическое описание хи-квадрат распределения.

К большому сожалению, хи-квадрат критерий оказывается наиболее точным при выборках в 200 опытов и больше. При обработке биометрических данных приходится иметь выборки в 20 примеров, что не дает использовать критерий хи-квадрат.

Такая же ситуация возникает и случаи использования других не параметрических статистических критериев [11], все они так же дают надежные оценки при выборках существенно более 20 примеров.

Теоретически доказано, что для биометрии данные одного контролируемого параметра примеров образа «Свой» имеют распределение близкое к нормальному. Для данных случайно выбранных образов «Чужие» распределение оказывается близко к равномерному. В связи с этим есть возможность исследуемые оценивать критерии по отношению к друг другу сравнивая между собой, даваемые ими равные ошибки первого и второго рода $P_1 = P_2 = P_{EE}$ при использовании малых выборок нормального и равномерного законов распределения. На рис. 1 приведены функции снижения вероятностей ошибок по мере роста размеров тестовой выборки.

Из данных рис. 1 видно, что мощность критерия KfM при малых выборках оказывается выше, чем у хи-квадрат критерия. Это делает критерий KfM более перспективным для биометрии в сравнении с хи-квадрат критерием, критерием Фроцини и критерием Джини.

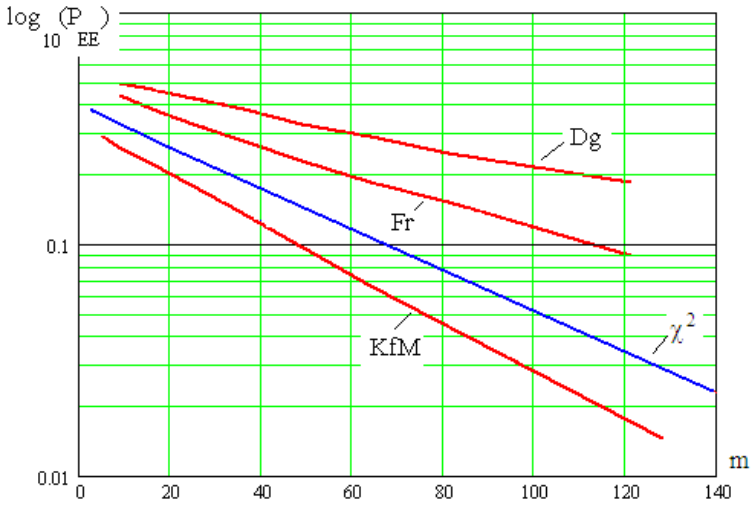


Рис. 1. График сопоставления мощностей интегральных статистических критериев Джини (Dg), Фроцини (Fr), хи-квадрат (χ^2) и Крамера – фон Мизеса (KfM)

По всей вероятности, основной причиной высокой мощности «Крамера – фон Мизеса» (KfM) критерия является то, что у него оказываются минимальные шумы квантования континуальных данных.

Исходный критерий «Крамера – фон Мизеса» вычисляется через интеграл квадрата разницы между гипотетической функцией вероятности – $P(\nu)$ эмпирической функции вероятности $\hat{P}(\nu)$:

$$KfM = \int_{-\infty}^{+\infty} (P(\nu) - \hat{P}(\nu))^2 \, d\nu. \qquad (1)$$

Формальный переход от интеграла (1) к операции суммирования нейроном с n входами не сложен:

$$\begin{cases} KfM(\nu) = \sum_{i=1}^{n} (P(\nu_i) - \hat{P}(\nu_i))^2 \\ KfM(\xi) = \sum_{i=1}^{n} (P(\nu_i) - \hat{P}(\xi_i))^2 \end{cases}, \qquad (2)$$

где ν_i пример образа «Свой» по n-контролируемым нейроном биометрическим параметрам; ξ_i пример образа «Чужой» по n-контролируемым нейроном биометрическим параметрам.

Эталонная функция вероятности – $P(\nu)$ оказывается легко вычислима только в том случае, когда выполнено центрирование и

нормирование, контролируемых биометрических параметров образа «Свой»:

$$\tilde{\nu}_i = \frac{\nu_i - E(\nu_i)}{\sigma(\nu_i)}. \qquad (3)$$

В новой центрированной и нормированной системе биометрических координат эталонная функция вероятности имеет нулевое математическое ожидание и единичное стандартное отклонение:

$$P_\nu(\tilde{\nu}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\tilde{\nu}} \exp\left\{\frac{-u^2}{2}\right\} du. \qquad (4)$$

На рис. 2 дан пример эталонной функции вероятности для нормального закона.

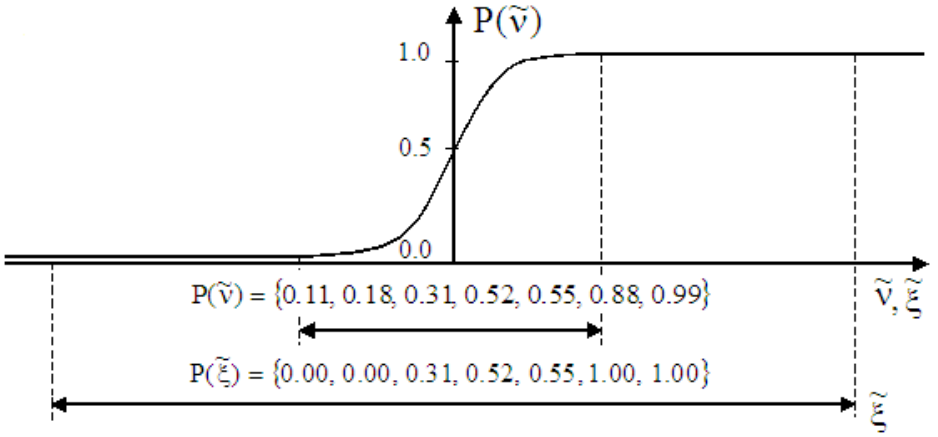


Рис. 2. Графический пример работы функции преобразования вектора значений переменных в вектор значений их вероятности

Из графика на рисунке 2 видно, что данные примеров образа «Свой» после их преобразования в вероятность должны будут оказаться в интервале от 0.0 до 1.0. Причем появление предельных значений 0.0 и 1.0 маловероятно.

Совершенно иная ситуация возникает, если вероятностному нейрону «Крамера – фон Мизеса» будут предъявлены данные примера образа «Чужой». Как видно из рис. 2 динамический диапазон данных – $\tilde{\xi}_i$ примерно в три раза больше, чем данных образа «Свой». Это означает, что на входы нейрона KfM будут поступать несколько предельных минимальных значений вероятности {0.0, 0.0,….} в начале упорядоченного списка и несколько предельных значений в конце упорядоченного списка {..., 1.0, 1,0}.

В целом вероятностный нейрон KfM работает за счет того, что выявляет неоправданно большое число предельных состояний вероятности в наблюдаемом биометрическом образе.

Нейрон KfM и структура нейронной сети, сформированной из нейронов KfM представлены на рис. 3.

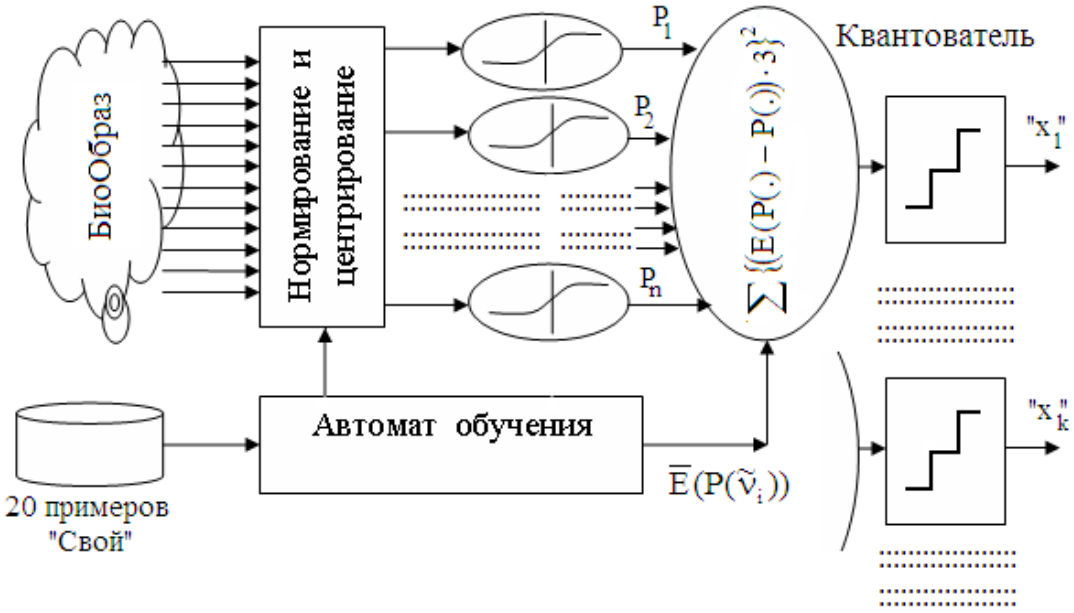


Рис. 3. Обобщенная структура нейросетевого преобразователя биометрических данных в код, построенная на использовании вероятностных нейронов «Крамера – фон Мизеса»

Полностью автоматическое обучение сети нейронов «Крамера – фон Мизеса»

Нейронных сетей может быть много, так же, как и типов нейронов из которых они собраны. Для нейросетевой биометрии принципиально важным является то, чтобы нейронная сеть преобразователя биометрия-код была способна быстро обучаться на малой выборке примеров «Свой» в полностью автоматическом режиме. В частности, этими свойствами обладает сеть персептронов, обученных алгоритмом ГОСТ Р 52633.5 [8].

В соответствии с требованиями алгоритма ГОСТ Р 52633.5 [8] нейроны подключены ко входам нейросети случайным образом. При синтезе нейронной сети, связи каждого нейрона задаются таблицей, заполняемой от генератора псевдослучайных чисел. Для сети нейронов KfM связи каждого нейрона должны формироваться также.

Важнейшим аспектом автоматического обучения сетей KfM является вычисление вектора математических ожиданий – $\overline{E}(\nu_i)$ и вектора стандартных отклонений – $\overline{\sigma}(\nu_i)$ параметров 20-ти примеров образа «Свой». Эти данные запоминаются и используются для центрирования и нормирования (3) всех входных параметров нейронной сети.

Вторым важнейшим аспектом является вычисление средних значений вероятности для упорядоченной выборки биометрических параметров образа «Свой» $E(P(\tilde{\nu}_1)),\ E(P(\tilde{\nu}_2)),.....,E(P(\tilde{\nu}_n))$. Эти параметры далее используются как эталоны для сравнения с предъявленными нейрону данными:

$$\begin{cases} KfM(\nu) = \sum_{i=1}^{n} \left\{ З\left(E(P(\tilde{\nu}_i) - P(\tilde{\nu}_i))\right)\right\}^2 \\ KfM(\xi) = \sum_{i=1}^{n} \left\{ З\left(E(P(\tilde{\nu}_i) - P(\tilde{\xi}_i))\right)\right\}^2 \end{cases}. \qquad (5)$$

Система (5) содержит два уравнения, так как для данных образа «Свой» и для данных образа «Чужой» нейрон работает совершенно по-разному. Для данных образа «Свой», каждое из которых относительно мало происходит их квадратичное сжатие и накопление сумматором.

Для данных образа «Чужой» Происходит не только линейное обогащение данных за счет суммирования, но и его нелинейное усиление за счет того, что нормировка биометрических данных «Чужой» выполняется на базе статистических данных «Свой»:

$$\tilde{\xi}_i = \frac{\xi_i - E(\nu_i)}{\sigma(\nu_i)}. \qquad (6)$$

Именно по этой причине с вероятностью выше 0.5 данные образа «Чужой» оказываются больше 1.0 и усиливаются квадратичной функцией. В конечном итоге выполняется условие:

$$KfM(\xi) > KfM(\nu). \qquad (7)$$

Так как нейроны «Крамера – фон Мизеса» работают с вероятностями, на каждом входе сумматора таких нейронов должен стоять функциональный преобразователь континуумов входных состояний в вероятности их появления. Схема процедуры нейросетевой аутентификации БиоОбразов, построенна на использовании множества вероятностных нейронов «Крамера – фон Мизеса».

По такой схема каждый нейрон отвечает за один или два выходных разрядов кода аутентификации.

Принципиально важным свойством вероятностных нейронов KfM является очень высокая устойчивость их обучения на малых выборках. Для того, чтобы выполнить нормирование и центрирование биометрических данных под БиоОбраз «Свой» достаточно базы из 20 примеров. Как показала практика на базе обучающей выборки из 20 примеров образа «Свой» можно с достаточной точностью вычислить вектор математических ожиданий – $\overline{E}(\nu_i)$ и вектор стандартных отклонений $\overline{\sigma}(\nu_i)$ всех контролируемых биометрических параметров.

Единственной не тривиальной функцией автомата обучения является вычисление по 20 примерам вектора математических ожиданий вероятности появления сортировки по возрастанию значений контролируемых биометрических параметров. При обучении нет переборов и запоминания промежуточных данных. Это означает, что обучение больших сетей искусственных нейронов «Крамера – фон Мизеса» имеет линейную вычислительную сложность, так же, как и алгоритм ГОСТ Р 52633.5 [8].

Усиление хэширующих свойств данных образов «Чужой» сетью нейронов Крамера – фон Мизеса

Основным свойством всех преобразователей биометрия-код является устранение почти до нуля естественной энтропии биометрических данных образа «Свой»:

$$H(\overline{\nu}) >> H(\text{"}\overline{c}\text{"}) \cong 0.03 \text{ бита.} \quad (8)$$

Для образов «Чужой» все нейросетевые преобразователи должны выполнять обратную функцию, усиливая естественную энтропию биометрических образов «Чужие»:

$$H(\overline{\xi}) < H(\text{"}\overline{x}\text{"}) \cong 45.03 \text{ бита.} \quad (9)$$

Проведенные исследования показали, что сети нейронов «Крамера – фон Мизеса» при двухуровневых квантователях значительно уступают по их хэширующим свойствам, стандартизованным нейросетям, обученным по ГОСТ Р 52633.5. Это происходит из-за того, что все нейросети, обученные по ГОСТ Р 52633.5 сбалансированы по вероятности появления выходных состояний «0» и «1» в разрядах выходного кода для образов «Чужие». В сети

нейронов KfM нет сбалансированности по вероятности появления в разрядах разных состояний. На рис. 4 приведены распределения вероятностей образов «Чужой» и наиболее вероятные положения порогов срабатывания квантователей.

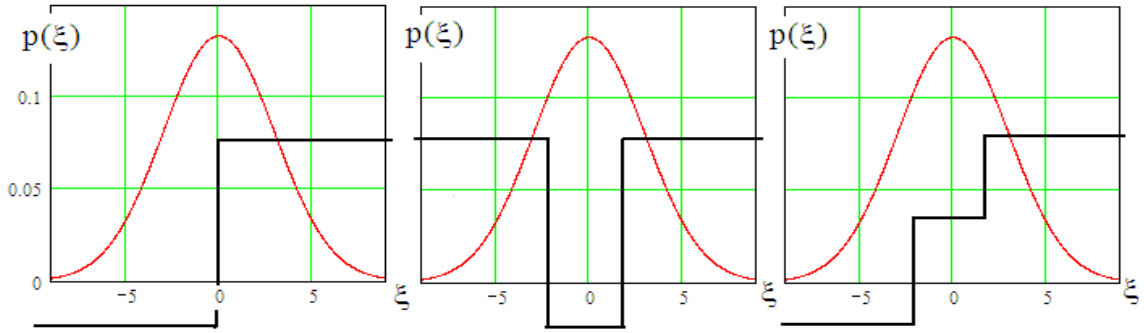


Рис. 4. Графики примеров использования разного типа квантователей для линейных нейронов

При обучении линейных нейронов по ГОСТ Р 52633.5 пороги всех квантователей настроены на срабатывание в центе распределения параметров образов «Все «Чужие» как это показано на в левой части рис. 4.

В центральной части рис. 4 отображена ситуация, когда в место нейронов с линейным накоплением используется одна из возможных квадратичных форм. Тогда превышение порога сравнения дает двустороннее ограничение данных, как это показано в центральной части рис. 4. В этом случае утрачивается баланс состояний «0» и «1» на выходах квадратичных нейронов. Этот эффект наблюдается и для нейронов KfM, как следствие наблюдается эффект снижения энтропии (9) квадратичных нейронов в сравнении с линейными нейронами, обученными по ГОСТ Р 52633.5 [8].

Для устранения этого нежелательного эффекта в нейронах KfM предложено использовать квантователи с тремя выходными состояниями [12–15], как это показано в правой части рис. 4. Для этой цели производится контроль значений математических ожиданий – $\mathrm{E}(\mathrm{P}(\tilde{\xi}))$ на входе нейрона KfM:

$$\begin{cases} \mathrm{z(Kfm)} = \text{"01"} \quad \text{if} \quad \mathrm{KfM} > \gamma \wedge \mathrm{E}(\mathrm{P}(\tilde{\xi})) > 0.5, \\ \mathrm{z(Kfm)} = \text{"00"} \quad \text{if} \quad \mathrm{KfM} < \gamma, \\ \mathrm{z(Kfm)} = \text{"10"} \quad \text{if} \quad \mathrm{KfM} > \gamma \wedge \mathrm{E}(\mathrm{P}(\tilde{\xi})) < 0.5 \end{cases}, \qquad (10)$$

где γ – порог срабатывания компаратора, находящегося на выходе кадратичного нейрона KfM, узнающего примеры образа «Свой» с высокой вероятностью в форме выходного состояния «00».

При использовании трехуровнего квантования (10) вероятности состояний «0» и «1» на выходе разрядов сети нейронов KfM выравниваются. Наблюдается рост энтропии выходных кодов «Чужой» до величины порядка 45 бит, что в полтора раза больше, чем тот же показатель для гостовских линейных нейронов. За счет использования механизмов трехуровнего квантования устраняется проведение «Атаки Маршалко» по извлечению данных из обученной нейронной сети [10, 11].

Заключение

Действующий в России ГОСТ Р 52633.5 распространяется только на автоматическое обучение сетей персептронов. Как было показано в этой статье сети KfM также имеют полностью автоматическое обучение с линейной вычислительной сложностью. Фактически действующий стандарт ГОСТ Р 52633.5 может быть дополнен таким же стандартом для сетей нейронов «Крамера – фон Мизеса».

Ранее было известно множество квадратичных нейросетевых функционалов [6, 7], однако они не использовались в нейросетевых преобразователях биометрия-код. Это было обусловлено низким уровнем энтропии их выходных кодов. Переход к троичным нейронам решает эту задачу, увеличивает энтропию выходных кодов по сравнению с сетями ГОСТ Р 52633.5 [5] и устраняет проведение «Атаки Маршалко» по извлечению данных из обученной нейронной сети [10, 11].

Библиографический список

1. Dodis, Y. Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy / Y. Dodis, L. Reyzin, A. Smith // Proc. EUROCRYPT. – 2004. – April 13. – P. 523–540.

2. Monrose, F. Cryptographic key generation from voice / F. Monrose, M. Reiter, Q. Li, S. Wetzel // Proc. IEEE Symp. on Security and Privacy. – 2001. – P. 202–213.

3. Hao, F. Crypto with Biometrics Effectively / F. Hao, R. Anderson, J. Daugman // IEEE TRANSACTIONS ON COMPUTERS, 2006. – Vol. 55, № 9. – September. – P. 1073–1074.

4. Язов, Ю. К. Нейросетевая защита персональных биометрических данных / Ю. К. Язов, В. И. Волчихин, А. И. Иванов, В. А. Фунтиков, И. Г. Назаров. – Москва : Радиотехника. – 2012. – 157 с.

5. Иванов, А. И. Биометрическая идентификация личности по динамике подсознательных движений : монография / А. И. Иванов. – Пенза : Изд-во ПГУ. – 2000. – 178 с.

6. Гудфеллоу, Я. Глубокое обучение / Я. Гудфеллоу, И. Бенджио, А. Курвиль. –Москва : ДМК Пресс, 2017. – 652 с.

7. Иванов, А. И. Нейросетевые технологии биометрической аутентификации пользователей открытых систем : автореф. дис. … д-ра техн. наук по специальности 05.13.01 «Системный анализ, управление и обработка информации» / Иванов А. И. – Пенза. – 2002. – 34 с.

8. ГОСТ Р 52633.5–2011. Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа.

9. Кобзарь, А. И. Прикладная математическая статистика для инженеров и научных работников / А. И. Кобзарь. – Москва : ФИЗМАТЛИТ, 2006. – 816 с.

10. ГОСТ Р 50.1.037–2002. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. – Москва : Госстандарт России, 2001. – Ч. 1. Критерии типа хи-квадрат. – 140 с.

11. ГОСТ Р 50.1.037–2002. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. – Москва : Госстандарт России, 2002. – Ч. 2. Непараметрические критерии. – 123 с.

12. Иванов, А. И. Подавление шумов квантования биометрических данных при использовании многомерного критерия Крамера – фон Мизеса / А. И. Иванов, А. И. Газин, С. Е. Вятчанин и др. // Проблемы информационной безопасности. Компьютерные системы. – Санкт Петербург : Изд-во ПТУ, 2016. – № 2. – С. 21–28.

13. Иванов, А. И. Сравнение мощности критерия среднего геометрического и Крамера – фон Мизеса на малых выборках биометрических данных / А. И. Иванов, А. Ю. Малыгин, П. А. Перфилов, С. Е. Вятчанин // Модели, системы, сети в экономике, технике, природе и обществе. – 2016. – № 2. – С. 155–158.

14. Волчихин, В. И. Перспективы использования искусственных нейронных сетей с многоуровневыми квантователями в технологии биометрико-нейросетевой аутентификации / В. И. Волчихин, А. И. Иванов, В. А. Фунтиков, Е. А. Малыгина // Известия высших учебных заведений. Поволжский регион. Технические науки. – Пенза : Изд-во ПГУ, 2013. – № 4 (28). – С. 88–99.

15. Алгоритмы обучения и тестирования нейросетевых преобразователей биометрия-код : монография / Б. С. Ахметов, В. И. Волчихин, А. И. Иванов, В. А. Фунтиков, Е. А. Малыгина. – Алматы : Изд-во КазНТУ, 2014. – 136 с.

Малыгина, Е. А. Усиление уровня защищенности биометрических технологий за счет использования нейронов Крамера – фон Мизеса / Е. А. Малыгина, С. Е. Вятчанин, А. И. Солопов // Безопасность информационных технологий : тр. I Всерос. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2019. – С. 44–54.

УДК 519.2, 612.087, 621.319.7

С. Е. Вятчанин, А. И. Иванов, Е. А. Малыгина, А. И. Солопов

СНИЖЕНИЕ ТРЕБОВАНИЙ К ОБЪЕМУ ОБУЧАЮЩЕЙ ВЫБОРКИ ЗА СЧЕТ СИММЕТРИЗАЦИИ СЕТЕЙ КВАДРАТИЧНЫХ ФОРМ

Аннотация. В отдельных биометрических приложениях вместо сетей нейронов могут быть использованы сети квадратичных форм, что закономерно ставит вопрос об их корректном статистическом описании, построенном на малой обучающей выборке примеров образа «Свой».

S. E. Vyatchanin, A. I. Ivanov, E. A. Malygina, A. I. Solopov

REDUCING THE REQUIREMENTS FOR THE TRAINING SAMPLE DUE TO THE SYMMETRIZATION OF NETWORKS OF QUADRATIC FORMS

Abstract. In individual biometric applications, instead of networks of neurons, networks of quadratic forms can be used, which naturally raises the question of their correct statistical description, built on a small training sample of examples of the Own image.

В настоящее время нейросетевые преобразователи биометрия-код [1, 2] строятся исходя из применения в нейроне входного сумматора (линейного элемента) для обогащения входных данных перед процедурой квантования. Это связано с относительной простотой и абсолютной устойчивостью используемого алгоритма обучений нейронов с линейными функционалами предварительного обогащения входных данных [3].

Однако из теории известно, что квадратичные функционалы повышения качества данных так же имеют простые и устойчивые алгоритмы обучения [4–6]. То есть в ряде биометрических приложений вместо сетей нейронов вполне могут быть использованы сети квадратичных форм.

В свою очередь, перспектива применения в будущем сетей квадратичных форм ставит вопрос о их корректном статистическом описании, построенном на малом объеме данных из 10…16 примеров образа «Свой».

Рассмотрим ситуацию, когда используется простейшая квадратичная форма с 8 входами, выбранными случайно из 416 контролируемых биометрических параметров [7]:

$$y_j = \sum_{i=1}^{8} \frac{(E(\nu_i) - \nu_{i,j})^2}{(\sigma(\nu_i))^2}, \qquad (1)$$

где ν_i – значения контролируемых биометрического параметра примеров образа «Свой» с нумерацией, соответствующей нумерации входов сумматора (связь номера входа сумматора – i с номером, контролируемого параметра – k с номерами от 1 до 416, задана таблицей связей); j – номер биометрического примера в обучающей или тестовой выборке.

Очевидно, что описание статистических свойств функционала (1) является тривиальным, если биометрические параметры являются независимыми. Положение меняется, когда связь, контролируемых биометрических параметров, оказывается существенной. Для зависимых биометрических данных происходит смещение моды распределения значений в левую сторону, как это показано на рис. 1.

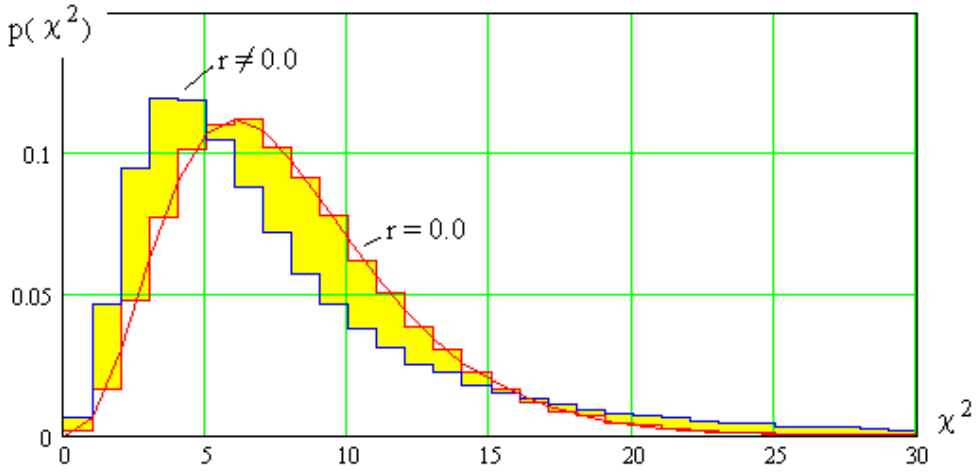


Рис. 1. Различие (отмечено заливкой) между идеальными данными без корреляции и реальными биометрическими данными со значительными корреляционными связями

Отметим, что смещение вершины гистограммы (моды) в левую сторону можно наблюдать, анализируя экспериментальные данные. Так, если мы имеем сеть из 256 квадратичных форм и образ «Свой» будет представлен 16 примерами, то общий массив данных статистического анализа составит 256x16= 4096 опытов. Этого объема вполне достаточно для гистограммы с единичными интервалами, которая приведена на рисунке 1 (r ≠ 0.0).

Аналогичную гистограмму мы можем построить, привлекая имитационное моделирование. Так как мы знаем, математические ожидания – $E(\nu_i)$ и стандартные отклонения – $\sigma(\nu_i)$ мы можем создать вектор независимых данных и для него построить гистограмму r =0.0. Естественно, что эти две гистограммы должны иметь существенное расхождение.

На рис. 1 расхождение гистограммы зависимых и независимых данных помечено заливкой.

Очевидно, что, увеличивая корреляционные связи синтезируемых данных, будем получать уменьшение ошибки между площадями двух гистограмм. Добиться роста корреляционных связей между синтезированными данными удается, если умножить вектор псевдослучайных данных на связывающую матрицу:

$$R := \begin{pmatrix} 1 & a & a & a & a & a & a & a \\ a & 1 & a & a & a & a & a & a \\ a & a & 1 & a & a & a & a & a \\ a & a & a & 1 & a & a & a & a \\ a & a & a & a & 1 & a & a & a \\ a & a & a & a & a & 1 & a & a \\ a & a & a & a & a & a & 1 & a \\ a & a & a & a & a & a & a & 1 \end{pmatrix}$$

Вне диагонали матрицы [1, 6] находятся одинаковые элементы «a», что обеспечивает равную коррелированность вектора случайных данных после умножения его на связывающую матрицу – R. В случае, когда параметр регуляризации оказывается нулевым, то выходные данные оказываются независимыми. Если же регулируемый параметр «a» отличен от нуля, то выходные данные связывающего преобразования $\bar{z} = R \cdot \bar{x}$ оказываются равно коррелированными.

Для осуществления многомерной статистической регуляризации биометрических данных нам достаточно осуществлять монотонное увеличение параметра регуляризации «a», добиваясь минимума ошибки расхождения гистограммы распределения реальных биометрических данных и гистограммы распределения синтезированных, равнокоррелированных данных. Точка минимума ошибки (минимума площади заливки) даст значение коэффициента коррелированности многомерной статистической модели образа «Свой».

Можно показать, что перечисленные выше условия всегда корректны в контексте минимизации модуля ошибки расхождения двух гистограмм или дифференциального критерия Джини [8, 9]. В этом отношении можно говорить о корректности регуляризационных вычислений по дифференциальному критерию Джини.

Библиографический список

1. Язов, Ю. К. Нейросетевая защита персональных биометрических данных / Ю. К. Язов, В. И. Волчихин, А. И. Иванов, В. А. Фунтиков, И. Г. Назаров. – Москва : Радиотехника, 2012. – 157 с.

2. ГОСТ Р 52633.0–2006. Защита информации. Техника защиты информации. Требования к средствам высоконадежной биометрической аутентификации.

3. ГОСТ Р 52633.5–201.1. Защита информации. Техника защиты инфор-мации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа.

4. Хайкин, С. Нейронные сети: полный курс / С. Хайкин. – Москва : Вильямс, 2006. – С. 1104.

5. Галушкин, А. И. Нейронные сети: история развития / А. И. Галушкин, Я. З. Цыпкин. – Москва : Радиотехника, 2001. – 840 с.

6. Ахметов, Б. Б. Многомерные статистики существенно зависимых биометрических данных, порождаемые нейросетевыми эмуляторами квадратичных форм : монография / Б. Б. Ахметов, А. И. Иванов. – Алматы : LEM, 2016. –86 с.

7. Иванов, А. И. Среда моделирования «БиоНейроАвтограф» [Программный продукт создан лабораторией биометрических и нейросетевых технологий, размещен с 2009 г. на сайте АО «ПНИЭИ»] / А. И. Иванов, О. С. Захаров. – URL: http://пниэи.рф/activity/science/noc.htm

8. Серикова, Н. И. Биометрическая статистика: сглаживание гистограмм, построенных на малой обучающей выборке / Н. И. Серикова, А. И. Иванов, С. В. Качалин // Вестник СибГАУ, 2014. – № 3 (55). – С. 146–150.

9. Серикова, Н. И. Оценка правдоподобия гипотезы о нормальном распределении по критерию Джини для сглаженных гистограмм, построенных на малых тестовых выборках / Н. И. Серикова, А. И. Иванов, Ю. И. Серикова // Вопросы радиоэлектроники. – Москва : ЦНИИ «Электроника», 2015. – Вып. 1. – С. 85–94. – Сер.: СОИУ.

Вятчанин, С. Е. Снижение требований к объему обучающей выборки за счет симметризации сетей квадратичных форм / С. Е. Вятчанин, А. И. Иванов, Е. А. Малыгина, А. И. Солопов // Безопасность информационных технологий : тр. I Всерос. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2019. – С. 55–58.

УДК 519.2, 612.087, 621.319.7

А. П. Карпов, А. П. Юнин

УСЛОВИЯ КОРРЕКТНОГО ВЫЧИСЛЕНИЯ ЭНТРОПИИ ОСМЫСЛЕННЫХ ДЛИННЫХ ПАРОЛЕЙ В ПРОСТРАНСТВЕ СВЕРТОК ХЭММИНГА С ЭТАЛОННЫМИ ТЕКСТАМИ НА РУССКОМ И АНГЛИЙСКОМ ЯЗЫКАХ

Аннотация. Рассматривается задача повышения корректности вычисления энтропии длинных кодов с зависимыми разрядами, являющимися осмысленными легко запоминаемыми паролями на родном языке пользователя. Показано, что результаты вычислений являются более корректными, если отказаться от побитного сложения по модулю 2 при вычислении сверток Хэмминга. Предложено использовать свертывание данных по модулю 8, так как кодирование паролей и эталонных текстов выполняются в 8-битной кодировке. Более того, корректное преобразование данных может быть выполнено только при использовании кода длинного пароля, свертываемого с эталонным текстом на родном языке пользователя.

A. P. Karpov, A. P. Junin

CONDITIONS FOR CORRECTLY CALCULATING THE ENTROPY OF MEANINGFUL LONG PASSWORDS IN HAMMING'S CONVOLUTIONS SPACE WITH REFERENCE TEXTS IN RUSSIAN AND ENGLISH LANGUAGES

Abstract. The problem of increasing the correctness of the calculation of the entropy of long codes with dependent digits, which are meaningful easily remembered passwords in the user's native language, is considered. It is shown that the results of calculations are more correct if we refuse bitwise modulo two when calculating Hamming convolutions. It is proposed to use data folding in modulo 8, since the coding of passwords and reference texts is performed in 8-bit encoding. Moreover, the correct data conversion can be performed only when using a long password code rolled up with the reference text in the user's native language.

Проблема вычисления энтропии длинных кодов с зависимыми разрядами

Если пытаться вычислять энтропию длинных кодов по Шеннону, то мы сталкиваемся с задачей экспоненциальной вычислительной сложности. Так для кодов длинной 256 бит, полученных от

программного генератора псевдослучайных чисел, возникает 2^{256} состояний. Произведение «Война и мир» в 4 томах Льва Толстого имеет1640 страниц, 2000 знаков на странице дает 2^{22} знаков. Пользуясь как эталонным текстом русского языка произведением «Война и мир» по Шеннону, мы можем оценивать пароли длинной до 176 бит или 22 знака. Для оценки пароля длинной в 32 случайных знака потребуется 2^{130} произведений на русском языке размерами сопоставимыми с 4-мя томами «Войны и мира». Все оцифрованные русскоязычные источники на содержат такой объем информации. Даже если бы такой эталон русскоязычного текста существовал, его анализ на обычном современном компьютере может занять тысячи лет машинного времени.

Проблема состоит в том, что, руководствуясь Шенноном, приходится обрабатывать большие массивы данных и ждать появления редких событий. Положение меняется, если мы из пространства обычных кодов переходим в пространство расстояний Хэмминга [1–3]. Для кодов длинной 256 бит расстояний Хэмминга меняется
в интервале $0 \le h \le 256$, итого 257 состояний:

$$h = 256 - \sum_{i=1}^{256} ("c_i") \oplus ("x_i"), \tag{1}$$

где "c_i" – разряд кода длинного пароля, "x_i" – этот же разряд кода эталонного текста.

В работах [4–6] показано, что свертка Хэмминга может быть выполнена не только по модулю два. Для того, что бы обобщить результаты сверток и сделать их сопоставимыми, нормируем интервал, в котором могут меняться расстояния Хэмминга:

$$\tilde{h} = \frac{h}{\max(h)}. \tag{2}$$

В этом случае нормированные расстояния всех сверток Хэмминга всегда будут находиться в интервале от 0 до 1. Для примера на рис. 1 даны распределения нормированных расстояний Хэмминга для эталонных текстов на русском и английском языках.

Из рис. 1 видно, что распределение расстояний Хэмминга при тестирование пароля на русском языке ближе к состоянию $\tilde{h} = 0$, то есть, подбирая пароль сочетаниями фраз на русском мы получим меньшую вероятность ошибок второго рода. В итоге оценка энтропии пароля в среде MathCAD дает величину:

$$-\log\left(\text{pnorm}\left(\frac{1}{256},0.357,0.032\right),2\right) = 92.628 \text{ бит}$$

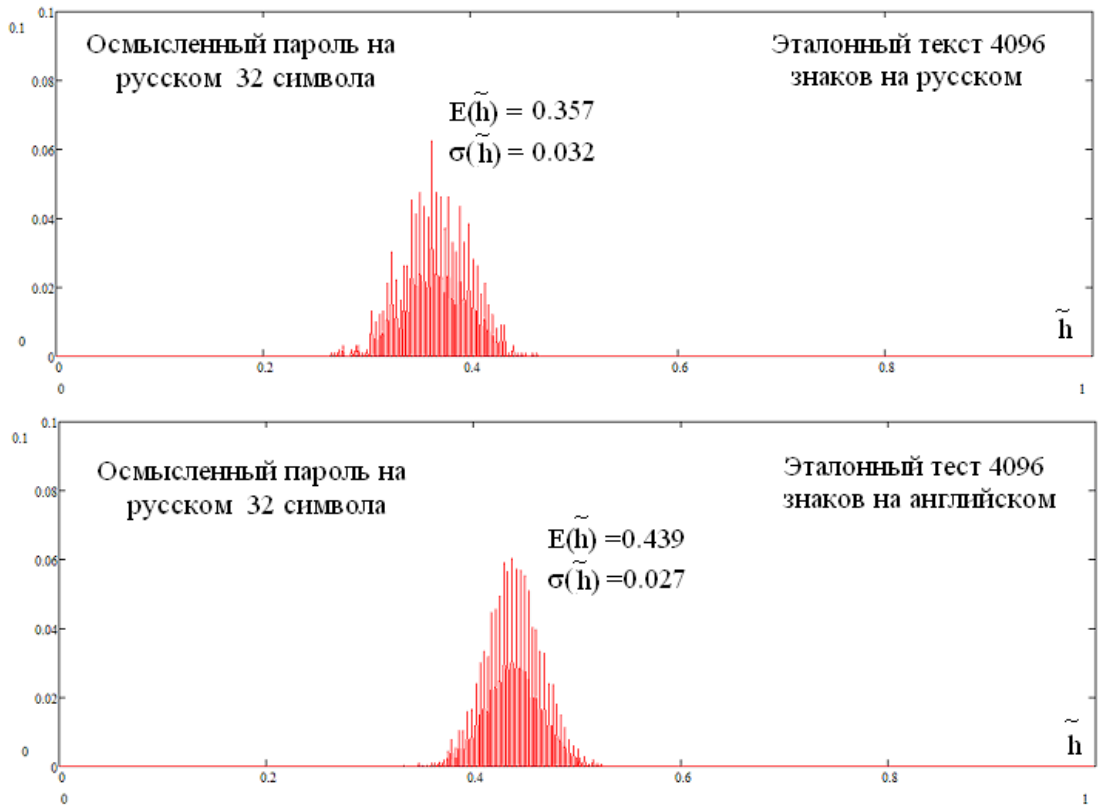


Рис. 1. Распределения расстояний Хэмминга при свертывании кода длинного осмысленного пароля с эталонными текстами на русском и английском языках

Если мы будем пытаться осуществить атаку, подбирая пароль на русском английскими фразами, то получим очень большую оценку энтропии:

$$-\log\left(\text{pnorm}\left(\frac{1}{256},0.439,0.027\right),2\right) = 192.661 \text{ бит}$$

Смысл подобных оценок понятен, пароль на русском языке следует подбирать, пользуясь фрагментами текстов на русском языке.

Следует отметить, что приведенные выше оценки являются слишком оптимистичными. Это обусловлено тем, что при вычислениях мы не принимали в расчет 8-битную кодировку символов. Учет 8-ми битной структуры кодов ASCII приводит к необходимости вычислять свертки Хэмминга по модулю восемь:

$$h_8 = 256 \cdot 32 - \sum_{i=1}^{32} \left("c_i, c_{i+1}, ..., c_{i+8}"\right) \oplus_8 \left("x_i, x_{i+1}, ..., x_{i+8}"\right). \qquad (3)$$

В итоге получаем более реалистичные распределения расстояний Хэмминга, приведенные на рис. 2.

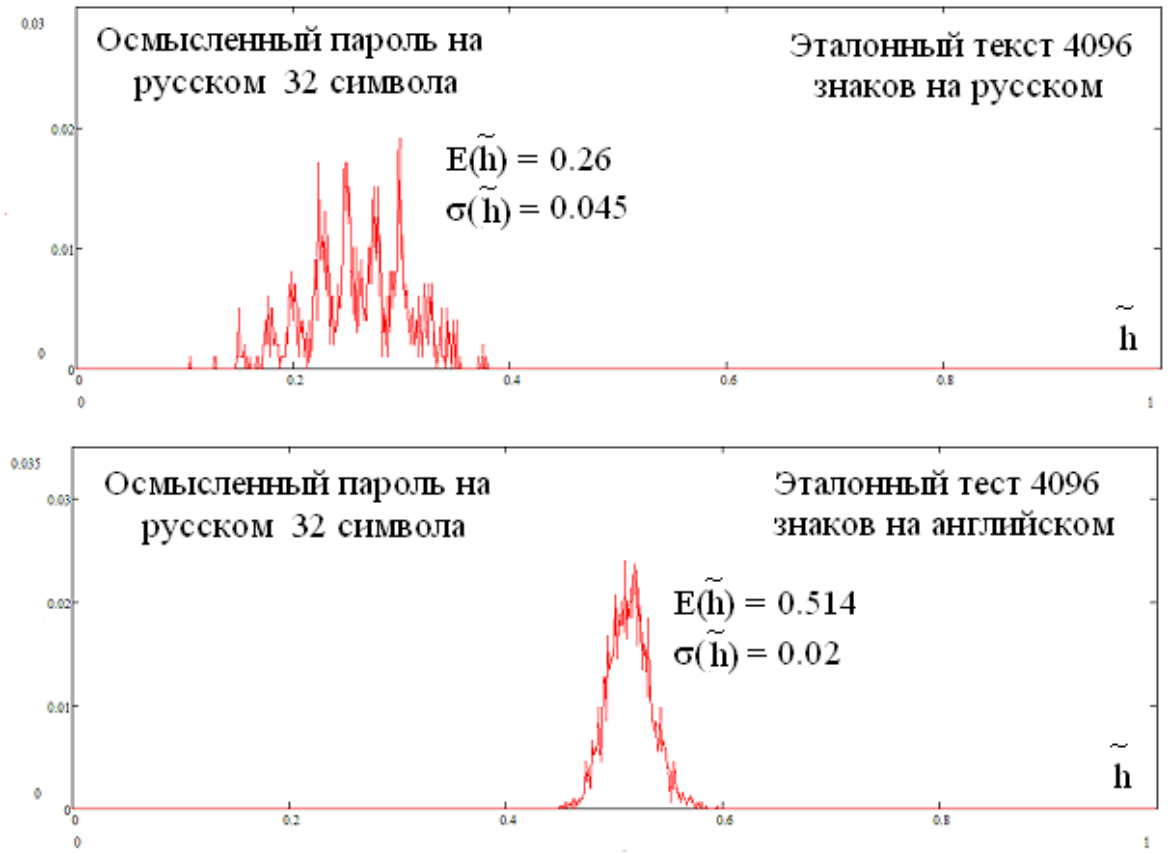


Рис. 2. Распределение расстояний Хэмминга в 8-битной системе счисления со свертыванием данных по модулю 8

Оценка энтропии для осмысленного пароля на русском при его тестировании тоже на русском получается ниже.

$$-\log\left(\mathrm{pnorm}\left(\frac{1}{256\cdot 32},0.26,0.045\right),2\right) = 27.954 \text{ бит}$$

Если мы такую же оценку выполняем, применяя сочетания слов на английском, то получаем увеличение энтропии:

$$-\log\left(\mathrm{pnorm}\left(\frac{1}{256\cdot 32},0.514,0.02\right),2\right) = 482.228 \text{ бит}$$

И в том и в другом случае получаются гораздо более реалистичные оценки энтропии. И в двоичной и восьмеричной системах сверток Хэмминга мы наблюдаем дефект вычислений (неустойчивость метода) когда тестируем пароль на другом языке. В восьмеричной системе счисления этот дефект усилился.

Наличие этого дефекта связано с тем, что ASCII кодировки имеют компактное расположение кодов букв латиницы и кодов букв кириллицы. Оба этих алфавита имеют расстояние между центрами групп «латиницы» и «кириллицы» 224 – 96 = 128 (7 бит). Именно это обстоятельство и приводит к расхождению математических ожиданий расстояний Хэмминга распределений рис. 1 и 2. Эту ситуацию иллюстрирует рис. 3 и 4.

Первая половина таблицы кодов ASCII

символ	10-й код	2-й код	символ	10-й код	2-й код	символ	10-й код	2-й код	символ	10-й код	2-й код
	32	00100000	8	56	00111000	P	80	01010000	h	104	01101000
!	33	00100001	9	57	00111001	Q	81	01010001	i	105	01101001
"	34	00100010	:	58	00111010	R	82	01010010	j	106	01101010
#	35	00100011	;	59	00111011	S	83	01010011	k	107	01101011
$	36	00100100	<	60	00111100	T	84	01010100	l	108	01101100
%	37	00100101	=	61	00111101	U	85	01010101	m	109	01101101
&	38	00100110	>	62	00111110	V	86	01010110	n	110	01101110
'	39	00100111	?	63	00111111	W	87	01010111	o	111	01101111
(	40	00101000	@	64	01000000	X	88	01011000	p	112	01110000
)	41	00101001	A	65	01000001	Y	89	01011001	q	113	01110001
*	42	00101010	B	66	01000010	Z	90	01011010	r	114	01110010
+	43	00101011	C	67	01000011	[	91	01011011	s	115	01110011
,	44	00101100	D	68	01000100	\	92	01011100	t	116	01110100
-	45	00101101	E	69	01000101	]	93	01011101	u	117	01110101
.	46	00101110	F	70	01000110	^	94	01011110	v	118	01110110
/	47	00101111	G	71	01000111	_	95	01011111	w	119	01110111
0	48	00110000	H	72	01001000	`	96	01100000	x	120	01111000
1	49	00110001	I	73	01001001	a	97	01100001	y	121	01111001
2	50	00110010	J	74	01001010	b	98	01100010	z	122	01111010
3	51	00110011	K	75	01001011	c	99	01100011	{	123	01111011
4	52	00110100	L	76	01001100	d	100	01100100	|	124	01111100
5	53	00110101	M	77	01001101	e	101	01100101	}	125	01111101
6	54	00110110	N	78	01001110	f	102	01100110	~	126	01111110
7	55	00110111	O	79	01001111	g	103	01100111	□	127	01111111

Рис. 3. Первая половина вариантов 8-битной ASCII кодировки символов (заливкой отмечены группы символов, используемых при кодировке текстов на английском)

На величину стандартного отклонения распределения расстояний Хэмминга прежде всего влияет компактность кодировки групп символов (отсутствие разрывов между кодами). Как следствие, сделать процедуры вычисления сверток Хэмминга более устойчивыми удается перекодировками, которые ликвидируют пробелы между кодами в группах «латиница» для текстов на английском и «кириллица» для текстов на русском. Часто используе-

мые в текстах знаки препинания должны иметь коды в группе символов в соответствии с вероятностью их появления в тексте. Группировка кодов и их упорядочивание по частоте появления символов являются мощными методами структурной регуляризации вычислений энтропии.

Вторая половина таблицы кодов ASCII

символ	10-й код	2-й код	символ	10-й код	2-й код	символ	10-й код	2-й код	символ	10-й код	2-й код
Ђ	128	10000000		160	10100000	А	192	11000000	а	224	11100000
Ѓ	129	10000001	Ў	161	10100001	Б	193	11000001	б	225	11100001
‚	130	10000010	ў	162	10100010	В	194	11000010	в	226	11100010
ѓ	131	10000011	J	163	10100011	Г	195	11000011	г	227	11100011
„	132	10000100	¤	164	10100100	Д	196	11000100	д	228	11100100
…	133	10000101	Ґ	165	10100101	Е	197	11000101	е	229	11100101
†	134	10000110	¦	166	10100110	Ж	198	11000110	ж	230	11100110
‡	135	10000111	§	167	10100111	З	199	11000111	з	231	11100111
€	136	10001000	Ё	168	10101000	И	200	11001000	и	232	11101000
‰	137	10001001	©	169	10101001	Й	201	11001001	й	233	11101001
Љ	138	10001010	Є	170	10101010	К	202	11001010	к	234	11101010
‹	139	10001011	«	171	10101011	Л	203	11001011	л	235	11101011
Њ	140	10001100	¬	172	10101100	М	204	11001100	м	236	11101100
Ќ	141	10001101	-	173	10101101	Н	205	11001101	н	237	11101101
Ћ	142	10001110	®	174	10101110	О	206	11001110	о	238	11101110
Џ	143	10001111	Ї	175	10101111	П	207	11001111	п	239	11101111
ђ	144	10010000	°	176	10110000	Р	208	11010000	р	240	11110000
‘	145	10010001	±	177	10110001	С	209	11010001	с	241	11110001
’	146	10010010	І	178	10110010	Т	210	11010010	т	242	11110010
“	147	10010011	і	179	10110011	У	211	11010011	у	243	11110011
”	148	10010100	ґ	180	10110100	Ф	212	11010100	ф	244	11110100
•	149	10010101	µ	181	10110101	Х	213	11010101	х	245	11110101
–	150	10010110	¶	182	10110110	Ц	214	11010110	ц	246	11110110
—	151	10010111	·	183	10110111	Ч	215	11010111	ч	247	11110111
□	152	10011000	ё	184	10111000	Ш	216	11011000	ш	248	11111000
™	153	10011001	№	185	10111001	Щ	217	11011001	щ	249	11111001
љ	154	10011010	є	186	10111010	Ъ	218	11011010	ъ	250	11111010
›	155	10011011	»	187	10111011	Ы	219	11011011	ы	251	11111011
њ	156	10011100	j	188	10111100	Ь	220	11011100	ь	252	11111100
ќ	157	10011101	S	189	10111101	Э	221	11011101	э	253	11111101
ћ	158	10011110	s	190	10111110	Ю	222	11011110	ю	254	11111110
џ	159	10011111	ї	191	10111111	Я	223	11011111	я	255	11111111

Рис. 4. Вторая половина 8-битной ASCII кодировки символов (заливкой отмечены группы символов, используемых при кодировке текстов на русском)

Таким образом, оценку энтропии в пространстве сверток Хэмминга можно сделать еще более устойчивой, если осуществлять предварительную перекодировку символов ASCII кодировки по специальную кодировку, обеспечивающую минимизацию значения математического ожидания расстояний Хэмминга и их стандартного отклонения.

Библиографический список

1. Иванов, А. И. Оценка усиления стойкости коротких цифровых паролей (PIN кодов) при их рукописном воспроизведении / А. И. Иванов, О. В. Ефимов, В. А. Фунтиков // Защита информации. INSIDE. – 2006. – № 1. – С. 55–57.

2. Малыгин, А. Ю. Быстрые алгоритмы тестирования нейросетевых механизмов биометрико-криптографической защиты информации / А. Ю. Малыгин, В. И. Волчихин, А. И. Иванов, В. А. Фунтиков. – Пенза : Изд-во ПГУ. –161 с.

3. ГОСТ Р 52633.3–2011. Защита информации. Техника защиты информации. Тестирование стойкости средств высоконадежной биометрической защиты к атакам подбора.

4. Юнин, А. П. Оценка энтропии легко запоминаемых, длинных паролей со смыслом в ASCII кодировке для русского и английского языков / А. П. Юнин, О. В. Корнеев // Тестирование стойкости средств высоконадежной биометрической защиты к атакам подбора : тр. науч.-техн. конф. кластера пензенских предприятий // Безопасность информационных технологий. – Пенза, 2016. – Т. 10. – С. 40–42. – URL: http://пниэи.рф/activity/science/BIT/T10-p40.pdf

5. Волчихин, В. И. Многомерный портрет цифровых последовательностей идеального «белого шума» в свертках Хэмминга / В. И. Волчихин, А. И. Иванов, А. П. Юнин, Е. А. Малыгина // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2017. – № 4 (44). – С. 4–13. DOI 10.21685/2072-3059-2017-4-1.

6. Иванов, А. И. Многомерная нейросетевая обработка биометрических данных с программным воспроизведением эффектов квантовой суперпозиции / А. И. Иванов. – Пенза : ПНИЭИ. – 2016. – 133 с. – URL: http://пниэи.рф/activity/science/BOOK16.pdf

Карпов, А. П. Условия корректного вычисления энтропии осмысленных длинных паролей в пространстве сверток Хэмминга с эталонными текстами на русском и английском языках / А. П. Карпов, А. П. Юнин // Безопасность информационных технологий : тр. I Всерос. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2019. – С. 59–65.

УДК 519.2, 612.087, 621.319.7

О. В. Корнеев

ОБЕЗЛИЧИВАНИЕ МЕДИЦИНСКИХ ЭЛЕКТРОННЫХ ДОКУМЕНТОВ ПРИ ИХ ХРАНЕНИИ С ПРИВЛЕЧЕНИЕМ ОБЛАЧНЫХ СЕРВИСОВ (ТЕХНОЛОГИЯ SAFENET, РАЗРАБОТКА НАЦИОНАЛЬНОЙ БИОМЕТРИЧЕСКОЙ ПЛАТФОРМЫ)

Аннотация. Показано, что один из эффективных способов защиты персональных данных при их размещении в облачных хранилищах является обезличивание электронных документов, например медицинских электронных историй болезни. При этом злоупотребления своим обезличиванием со стороны пациентов исключается за счет использования биометрической аутентификации личности. В этом случае пациент не может подменить себя другим человеком с другой биометрией, а врач не имеет технических возможностей дезавуировать обезличивание электронных медицинских документов. Защита больших объемов персональной медицинской информации строится на том, что злоумышленник, получивший незаконный доступ к облачному хранилищу не может узнать, какому человеку принадлежит та или иная история болезни.

Технология обезличивания электронных медицинских документов на текущий момент наиболее глубоко проработана. Она может быть использована как образец облачной организации национальной биометрической платформы (технология SafeNet). Ожидается, что следующее поколение электронных паспортов граждан РФ должно храниться в облаках в обезличенной форме, при этом владелец электронного биометрического паспорта знает его параметры хранения и легко может им воспользоваться. Злоумышленник же, напротив, для того, чтобы добраться до электронного паспорта нужного ему человека, будет вынужден перебирать очень большое число возможных вариантов, вскрывая защиту каждого из проверяемых вариантов.

O. V. Korneev

DEPRECIATION OF MEDICAL ELECTRONIC DOCUMENTS DURING THEIR STORAGE WITH THE INVOLVEMENT OF CLOUD SERVICES (SAFENET TECHNOLOGY, DEVELOPMENT OF A NATIONAL BIOMETRIC PLATFORM)

Abstract. It is shown that one of the most effective ways to protect personal data when they are placed in cloud storages is the de-identification of electronic documents, for example, medical electronic medical records. At the same time, the

abuse of their depersonalization by patients is excluded due to the use of biometric authentication of the person. In this case, the patient cannot replace himself with another person with a different biometrics, and the doctor does not have the technical ability to disavow the de-identification of electronic medical documents. Protection of large amounts of personal medical information is based on the fact that an attacker who received illegal access to the cloud repository can not find out who the person owns one or another medical history.

The technology of anonymity of electronic medical documents is currently the most deeply developed. It can be used as a sample of the cloud organization of the national biometric platform (SafeNet technology). It is expected that the next generation of electronic passports of citizens of the Russian Federation should be stored in the clouds in an impersonal form, while the owner of the electronic biometric passport knows its storage parameters and can easily use it. The attacker, on the other hand, will have to go through a very large number of possible options in order to get to the e-passport of the person he needs, revealing the protection of each of the checked options.

Введение

В настоящее время активно идут процессы информатизации современного общества, как итог, наша персональная информация постепенно перемещается в Интернет облака. Ярким примером общего вектора развития являются медицинские информационные системы. В 2006 году в России был введен в действие отечественный стандарт, регламентирующий требования к электронной истории болезни [1], это позволило разработать типовую медицинскую информационную систему [2]. Далее, встал вопрос о переходе к использованию типового электронного места врача [3]. В итоге, информационная технология уже позволяет создавать интегрированные электронные медицинские карты [4], которые могут размещаться как на локальном сервере медицинской информационной системы, так и на Интернет серверах поставщиков облачных услуг.

Естественно, что эта общая тенденция порождает новые угрозы информационной безопасности, которые должны быть устранены с учетом уже сложившейся технической практики [5] и национального законодательства [6]. За рубежом проблема решается через биометрическую аутентификацию личности человека с использованием, так называемых, «нечетких экстракторов» [7]. В России для этих же целей используются искусственные нейронные сети [8], применяя которые можно осуществлять обезличивание [9] медицинских электронных документов, в случае их размещения в облачных хранилищах.

Очевидным является так же то, что массовое использование биометрического обезличивания персональной информации облачных сервисов подпадает под юрисдикцию государственных регуляторов рынка средств информационной безопасности. От производителей потребитель в праве потребовать сертификаты ФСБ России на соответствующие криптографические модули защиты, а также собственник облачного сервера в праве потребовать у производителя сертификат ФСТЭК России на соответствие нейросетевых преобразователей биометрии в код пароля доступа (если таковые используются в продукте).

В первом случае, экспертная организация ФСБ России (испытательная лаборатория) должна оценить корректность программной реализации модулей криптографической защиты информации, использованных для защиты облачных хранилищей. Во втором случае экспертная организация (испытательная лаборатория, аккредитованная ФСТЭК России) должна оценить корректность реализации нейросетевых преобразований на соответствие требованиям пакета национальных стандартов ГОСТ Р 52633.xx и, в том числе, стойкость к атакам подбора нейросетевой биометрической защиты по алгоритмам тестирования ГОСТ Р 52633.3–2011 [10].

Важнейшим технологически-правовым моментом всех облачных сервисов является то, что облачные услуги не подпадают под действие ФЗ № 152 «О персональных данных» [6], если персональные данные зашифрованы (имеется сертификат ФСБ России на средство шифрования) или персональные данные надежно обезличены (имеется сертификат ФСТЭК России на средство биометрической аутентификации, исключающей подмену обезличенного пользователя).

То есть, владелец облачного сервиса не является «оператором» персональных данных и не осуществляет «обработку» персональных данных. От владельца облачного сервиса требуется не более, чем физическое размещение его оборудования (серверов и дата центров) на территории Российской Федерации.

Оператором персональных данных, в нашем случае является, медицинская организация, владеющая медицинской информационной системой (МИС). Именно медицинская организация должна уведомить Уполномоченный орган Роскомнадзора по защите прав субъектов персональных данных для последующей регистрации в качестве оператора по обработке персональных данных. Именно медицинская организация должна подготовить техническое зада-

ние по созданию системы защиты персональных данных и затем убедить Уполномоченный орган Роскомнадзора о достаточном уровне защиты персональных данных в медицинской информационной системе, опираясь на наличие сертификатов ФСБ России и/или ФСТЭК России.

Актуальная модель угроз для персональных данных пациентов медицинской информационной системы

В прошлом веке регистратура поликлиник имела стеллажи, заполненные медицинскими картами пациентов. К врачу можно было попасть только после того, как в регистратуре найдут твою медицинскую карту. Похитить несколько тонн бумаги незаметно было технически невозможно. Сегодня технология изменилась, личная электронная медицинская карта хранится на сервере медицинской информационной системы. Системный администратор медицинской информационной системы, увольняясь с очередного места работы, легко может прихватить с собой архив медицинских карт всех пациентов с актуальной информацией за несколько лет.

Более того, в рамках наиболее продвинутых медицинских учреждений США отмечены случаи, когда в их информационной системе появлялись вирусы, уничтожающие всю актуальную информацию и параллельно шифрующие архивы на новом ключе. Администрация любого респектабельного лечебного учреждения, как правило, не стремиться к поиску и публичному наказанию виновных. Для нее куда более важным является восстановление медицинского технологического процесса.

Следует подчеркнуть, что столь радикальные технические решения, как тотальное шифрование не всегда оправдано. Даже в специализированных медицинских учреждениях, ориентированных на лечение социально-значимых заболеваний, диагнозы и методы лечения повторяются. Вполне достаточно изъять из всех электронных медицинских документов персональные данные пациентов, позволяющие найти человека помимо его воли [11].

Злоумышленник, похитивший архив электронных медицинских документов, сможет узнать цену всех медицинских услуг, оказанных пациенту, он будет иметь полную информацию о ходе лечения и его результатах. Однако, он не сможет узнать самого главного – кто является пациентом, какой номер страхового медицинского полиса у пациента. Этих данных нет в медицинской информационной системе, они хранятся отдельно.

Для технологии обезличивания особой роли не играет, где будут размещены данные (обезличенные архивы). После обезличивания электронные истории болезни могут открыто храниться, как на локальном сервере медицинской информационной системе без доступа к сети Интернет, так и на сервере провайдера облачных услуг. Эти две ситуации отображены на рис. 1.

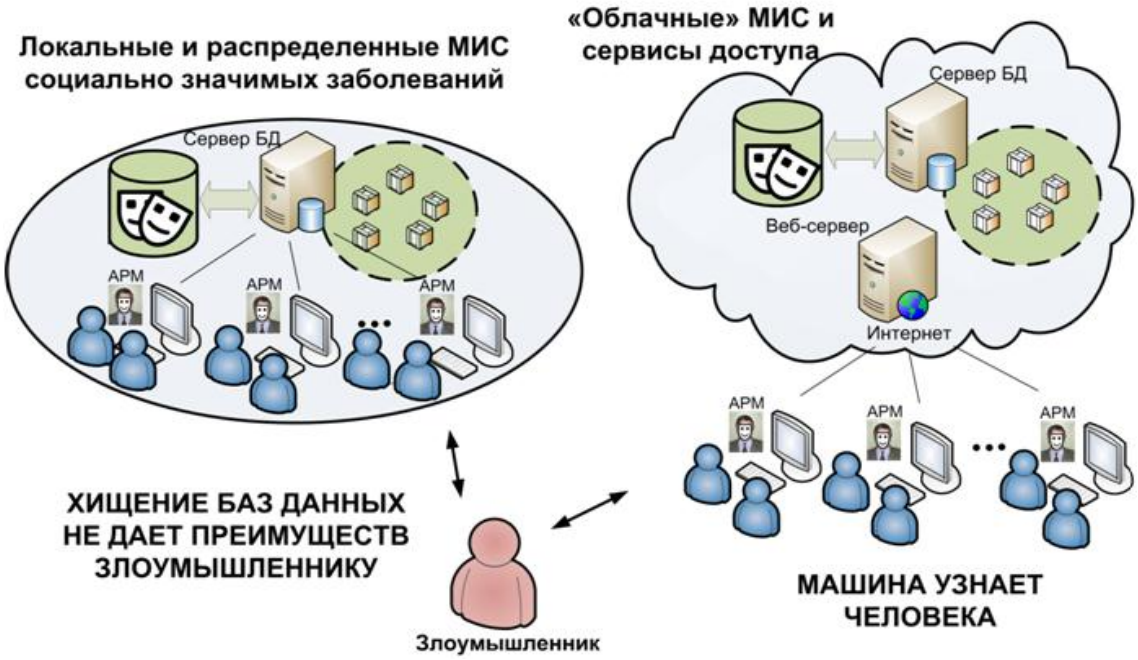


Рис. 1. Два варианта размещения обезличенных электронных медицинских документов на локальном сервере медицинской информационной системы и в облачном хранилище

И в том и в другом случае, злоумышленник не имеет возможности добраться до информации о полноценных персональных данных человека. Чем чаще встречается заболевание, тем надежнее оказывается защита наших персональных данных через обезличивание электронных медицинских документов.

Фактически, обезличивание информации медицинских электронных документов является защитой от злоупотреблений со стороны персонала, обслуживающего пациентов (системного администратора МИС, провизора аптеки, техников-лаборантов при исследовании биологических проб, группы лечащих врачей по нескольким медицинским специализациям). При этом, опираться на «клятву Гиппократа» можно только рассматривая категорию – «врачи», иной медико-технический персонал будет связан не более чем некоторыми обязательствами о неразглашении сведений «Для

служебного пользования», принятыми по отношению к работодателю при трудоустройстве.

Новая информационная технология позволяет перейти от обычных организационных мероприятий по защите персональных данных к отсутствию технических возможностей у всего медико-технического персонала медицинской организации. Естественно, что в этой новой ситуации возникают и новые угрозы.

Основной угрозой, порождаемой обезличиванием электронных медицинских документов является злоупотребления со стороны пациентов. Пациент может быть заинтересован в фальсификации объективных данных о своем здоровье. Кому-то хочется быть здоровым и такие люди могут попытаться подменить себя здоровым человеком. Возможна противоположная ситуация, когда здоровый человек, по каким-то причинам хочет казаться больным.

Биометрическая поддержка обезличивания медицинских электронных документов

Люди способны практически безошибочно узнавать друг друга. Эту способность долгое время принято было рассматривать как прерогативу естественного интеллекта. Однако, начиная с конца прошлого века по всему миру активно стали развиваться биометрические технологии [12, 13]. Очевидным лидером этих процессов были США [12], поставившие цель создания и продвижение нового поколения биометрических паспортно-визовых документов. Сформулированная США цель была достигнута и сегодня большинство стран имеет биометрическое усиление своих международных паспортов. Российский международный паспорт, как и паспорта других стран, имеет встроенный радио-читаемый процессор с биометрическими данными своего владельца.

Одной из неприятностей биометрических технологий США, стандартизированных ISO/IEC JTC1 SC 37 (Biometrics), является полное отсутствие гарантий информационной безопасности. Так называемые «биометрические шаблоны», хранящиеся в радио-читаемых процессорах био-паспортов, полностью повторяют идеологию полицейской биометрии, ориентированную на поиски преступников. По этой причине открытые «биометрические шаблоны» ISO/IEC JTC1 SC 37 (Biometrics) нельзя размещать на Интернет серверах облачного хранилища.

В связи с этим, появилось специальное направление исследований, которое занимается защищенным вариантом исполнения средств биометрической аутентификации [7, 8, 13]. Защищенный вариант в форме зарубежных «нечетких экстракторов» [7] или отечественный вариант нейросетевых преобразователей биометрия-код [8, 13] изначально ориентированы под безопасное хранение персональных биометрических данных человека на сервере МИС или в облаках. На рис. 2 иллюстрируется технология нейросетевого преобразования того или иного биометрического образа в код, используемый при последующей криптографической аутентификации.

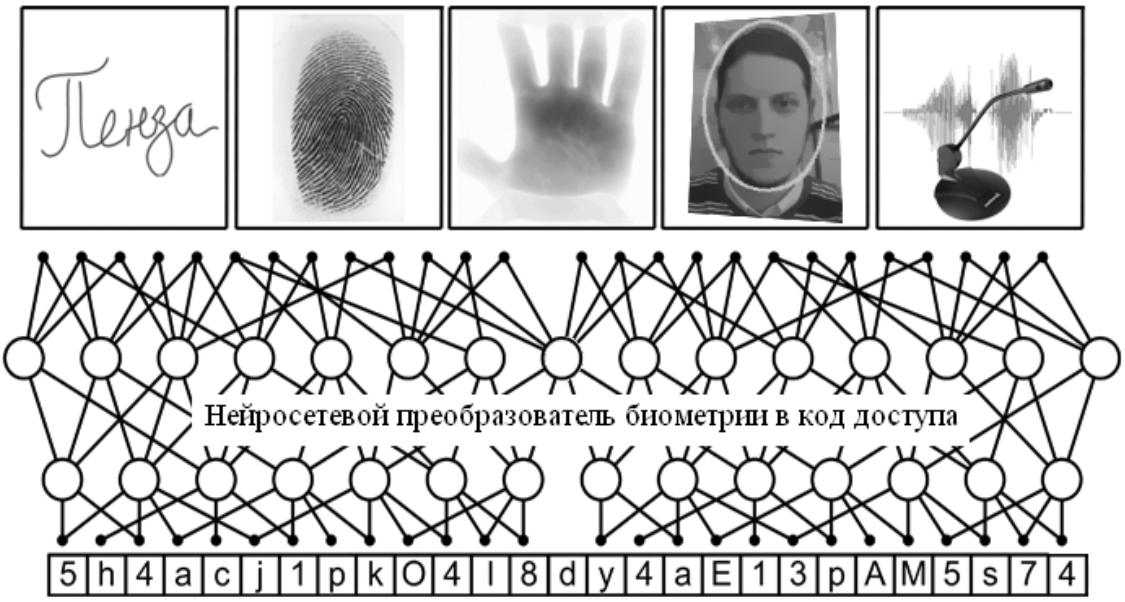


Рис. 2. Биометрические образы человека, преобразуемые искусственной нейронной сетью в код криптографического ключа или длинного пароля доступа

Каждый из биометрических образов перед его использованием должен быть отсканирован, далее, из биометрического образа должны быть извлечены контролируемые биометрические параметры. Эти вычисленные параметры подаются на входы заранее обученной искусственной нейронной сети, которая для любого примера образа «Свой» порождает на своих выходах длинный код аутентификации (код личного криптографического ключа или длинного пароля доступа). Если злоумышленник – «Чужой» будет подставлять случайные биометрические образы, то на выходах нейронной сети он будет получать случайные выходные коды.

Основная технологическая функция биометрической поддержки обезличивания состоит в том, что в место медицинского персонала машина узнает человека. Незнакомый врачу пациент своей биометрией подтверждает свои полномочия без предъявления паспорта и объявления своего подлинного имени. Предъявив медицинской информационной системе идентификационную карту (аналог радио-читаемой карты, например, используемой в метро) и рисунок своего отпечатка пальца пациент подтверждает то, что он ранее официально зарегистрирован. Как минимум, существует бумажный договор, где указаны действительные персональные данные пациента в связке с идентификационным номером его идентификационной карты.

Пациент не может обмануть информационную систему, отдав свой идентификатор другому человеку. Врач не может принести вред пациенту через разглашение тайны его диагноза, так как он не знает подлинных персональных данных пациента.

Технология извлечения знаний из преобразователей биометрия-код

Большинству специалистов кажется, что размещение персональных биометрических данных в зарубежном «нечетком экстракторе» или в отечественной нейронной сети, является достаточно надежной защитой. На самом деле это не так, времена Шеннона давно прошли, каждый из нас имеет вычислительную машину, а хакеры имеют возможность использования сотен зомбированных серверов. Это позволяет хакерам организовывать достаточно сложные атаки на биометрию.

Целью атак является попытка извлечения знаний из «нечеткого экстрактора» или нейросетевого контейнера. Оба типа этих преобразователей по своему определению должны иметь область примеров образов «Свой» с нулевой энтропией выходных кодовых состояний. В связи с этим, мы можем заранее создать базу из 1024 образов «Чужой», подавать эти образы на вход преобразователя биометрия-код. Далее, мы можем рассчитать энтропию каждого их полученных кодов по отношению ко всем иным кодам. Результат таких вычислений приведен на рис. 3.

Из рис. 3 видно, что энтропия всех кодов образов «Чужой» находится в интервале от 4 бит до 19 бит (данные отложены по вертикальной оси). При этом расстояние Хэмминга между выходными кодами меняется от 0 до 256 бит (данные горизонтальной

оси). С право и с лево находятся коды с минимальной энтропией, наиболее близкие к коду образа «Свой» и его инверсии.

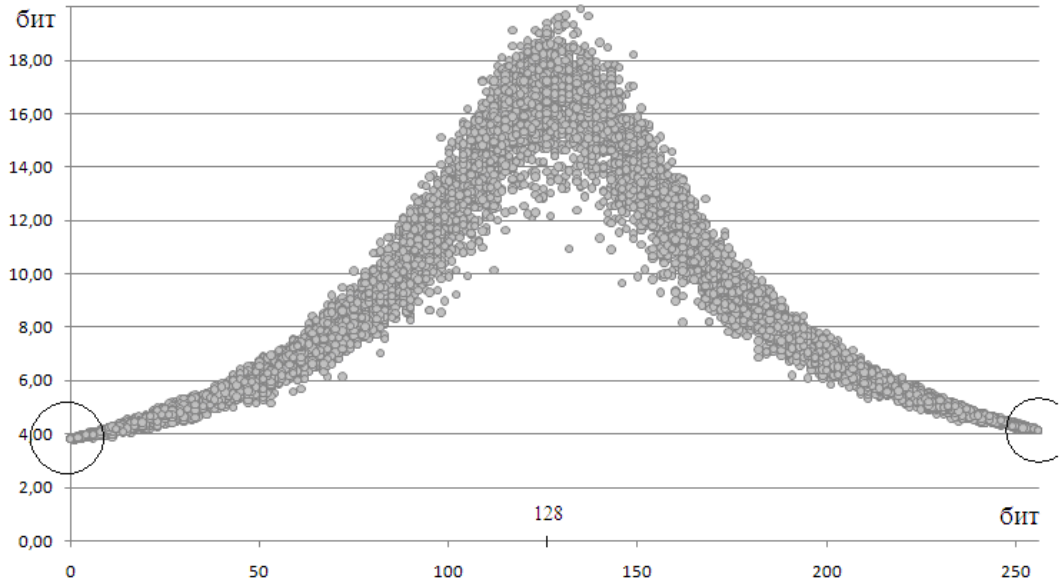


Рис. 3. Распределение частных значений энтропии выходных кодов 1024 биометрических образов «Чужой» для нейросетевого преобразователя биометрия-код

Мы не знаем, какой код дает образ «Свой», однако мы можем выделить две группы кодов-претендентов и, соответствующие им группы образов «Чужой». Обычно в первой и второй группах оставляют по 20 образов, что соответствует 2 % от исходной базы. Этот процесс можно повторить, восстановив численность образов «Чужой» до первоначальной численности. Восстановление численности выполняется скрещиванием двух образов-родителей и получением от них образов-потомков по ГОСТ Р 52633.2 [14].

Процедуры скрещивания и селекции следует повторять в нескольких поколениях, что позволяет извлечь порядка 9 7% достоверной информации об образе «Свой» и коде, им порождаемом. При этом для извлечения знаний из «нечеткого экстрактора» потребуется 5 поколений (10 минут машинного времени обычного компьютера), а для извлечения знаний из нейронной сети потребуется 50 поколений (50 минут машинного времени).

Гарантии безопасности хранения личной биометрии пациентов на облачном сервере

Очевидно, что 10 или 50 минут машинного времени для злоумышленников не является сколько-нибудь существенным препят-

ствием. То есть, «нечеткие экстракторы» и нейросетевые преобразователи нельзя хранить на облачных серверах. «Нечеткие экстракторы» и нейросетевые преобразователи следует рассматривать как аналоги обычных паролей, вводимых с клавиатуры. В вычислительной машине или на сервере нельзя хранить пароли – это опасно. В вычислительной машине и на сервере обычно хранят хэш-функции от пароля. То есть, безопасно хранить на сервере мы можем некоторую последовательность хэш-функций от нейросетевого преобразователя биометрия-код.

Для пояснения технологии защиты на рис. 4 приведена обобщенная структура нейросетевого преобразователя биометрия-код.

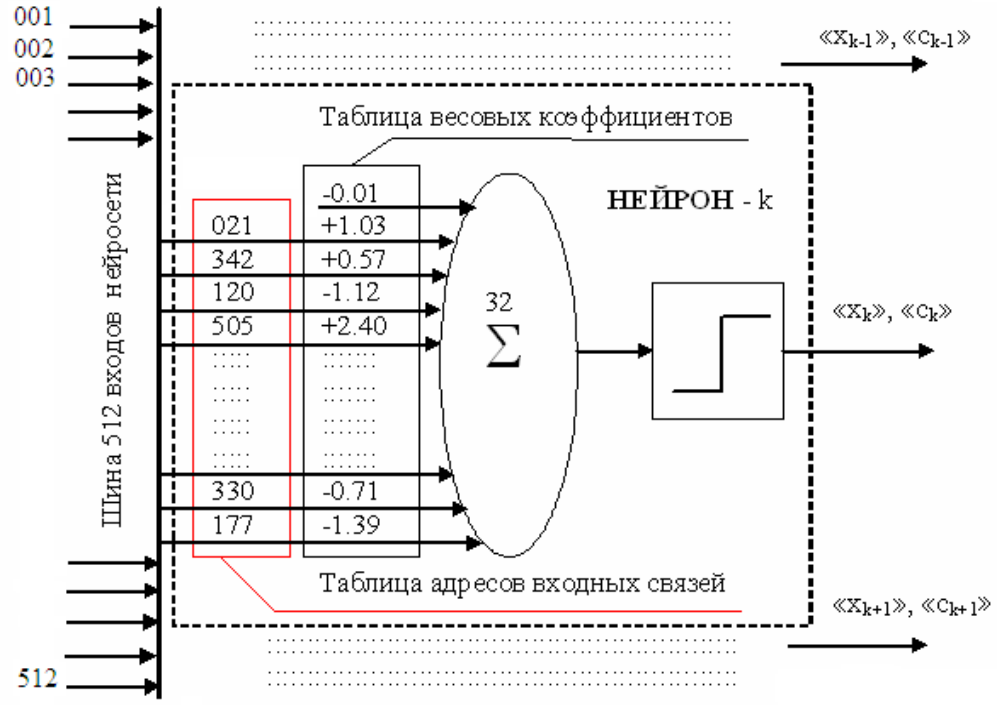


Рис. 4. Обобщенная структура нейросетевых преобразователей

Каждый нейрон обученного преобразователя имеет две таблицы. Первой является таблица связей входов нейрона со входами нейронной сети в целом. Второй является таблица весовых коэффициентов сумматора для каждого из нейронов. Этих двух таблиц достаточно.

Стандартный алгоритм обучения нейронной сети ГОСТ Р 52633.5–2011 [15] предполагает задание таблиц выходных связей от генератора псевдослучайных чисел. Для того, чтобы защитить нейросетевые контейнеры, достаточно выработать гамму, накрывающую первую и вторую таблицы. Для защиты таблиц пер-

вого нейрона гамма вырабатывается хэшированием конкатенации обычных данных аутентификации:

$$\Gamma_1 = \text{HASH}(\text{соль}|\text{пароль}). \quad (1)$$

Этого достаточно для снятия защиты с таблиц первого нейрона. После открытия таблиц может быть выполнена операция по преобразованию биометрического образа первым нейроном и получено его выходное состояние «c_1».

Гамма для таблиц второго нейрона вычисляется с учетом выходного состояния первого нейрона:

$$\Gamma_2 = \text{HASH}(\text{соль}|\text{пароль}|c_1). \quad (2)$$

На каждой итерации число учитываемых выходных состояний нейронов увеличивается. Так для i-го нейрона, используемая гамма будет строиться хэшированием конкатенации (i-1) состояния предшествующих нейронов:

$$\Gamma_i = \text{HASH}(\text{соль}|\text{пароль}|c_1|c_2|\ldots|c_{i-1}). \quad (3)$$

Очевидно, что любая ошибка при вводе пароля или ошибочная подстановка другого биометрического образа будет приводить к тому, что таблицы обученной нейронной сети восстановить не удастся. Как следствие, не удастся восстановить и выходной код нейросетевой биометрической аутентификации, поддерживающей обезличивание.

То, что таблицы нейросетевого контейнера защищены гаммированием, а гаммы создаются вызовом полноценных криптографических хэш-функций, является гарантией безопасного хранения биометрических данных на сервере провайдера облачных услуг. Из-за защиты таблиц обученных нейронных сетей гаммированием, хакерам уже не удается наблюдать неоднородности энтропии кодов нейросетевого преобразователя (рис. 3). Энтропия выходных кодов перестает иметь какие-либо особенности, атакующий вынужден решать вычислительно сложную задачу по восстановлению неизвестных данных после их хэширования (1)–(3).

Типовая структура МИС, использующая защиту электронных медицинских документов через их обезличивание

В целом структура медицинской информационной системы изменяется незначительно. Ее безопасность строится на применении асимметричной криптографии (поддерживается инфраструктура открытых ключей), что иллюстрируется рис. 5.

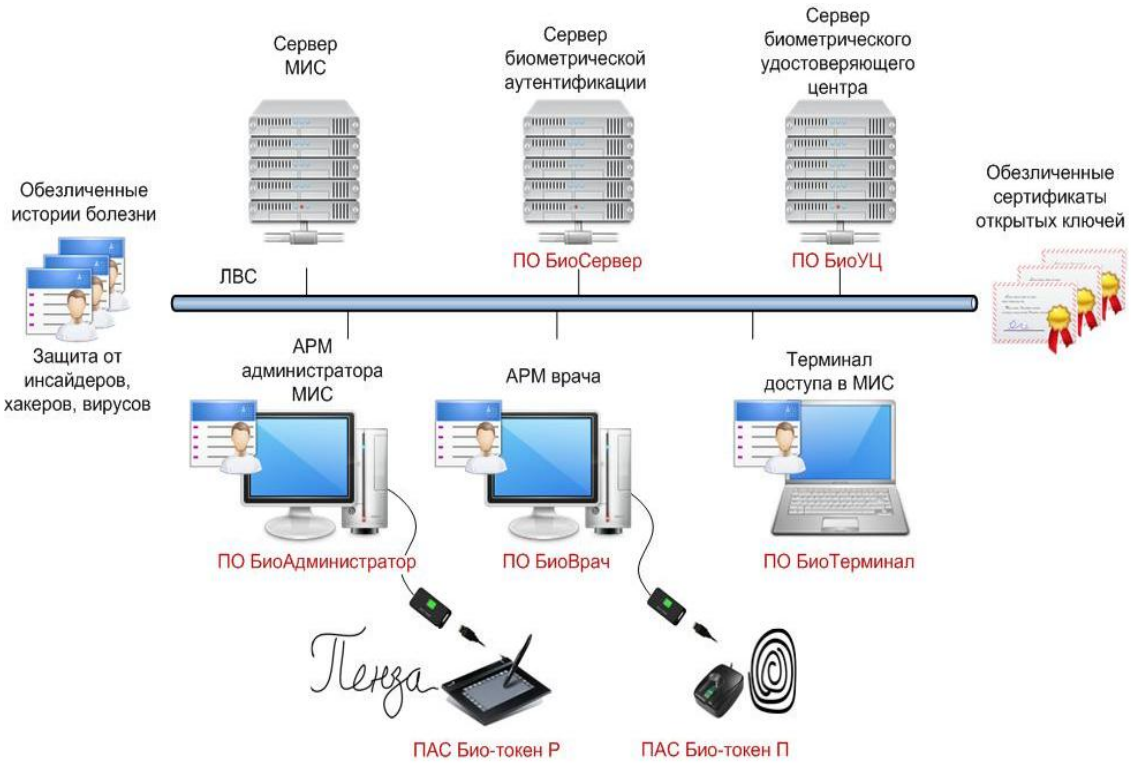


Рис. 5. Структура медицинской информационной системы, поддерживающей обезличенные электронные документы (обезличенные сертификаты открытых ключей)

По сравнению с традиционными техническими решениями в структуре МИС появляется сервер биометрической аутентификации и сервер биометрического удостоверяющего центра. В место обычных ЛИЧНОСТНЫХ сертификатов открытых ключей в МИС используются обезличенные сертификаты. Во всем остальном применяются типовые аппаратно-программные решения. Существующей сегодня продукт [11], пока ориентирован под две биометрические технологии (рукописный почерк и анализ рисунка отпечатка пальца). Датчики съема биометрической информации стандартные, сканер биометрии к вычислительной машине подключается через доверенную вычислительную среду USB «БиоТокен». В доверенной вычислительной среде USB «БиоТокен» выполняются все потенциально опасные операции:

- распаковка защищенного нейросетевого контейнера;
- обработка биометрических данных;
- формирование и проверка цифровой подписи.

Хищение обезличенных историй болезни и иных электронных документов, заражение вирусами МИС для описанного выше технического решения не опасны. МИС, построенные на подобных

принципах могут использовать типовое рабочее место врача [2, 3] с интегрированной электронной медицинской картой [4], хранящейся как на локальном сервере, так и на облачных серверах.

Технология SafeNet, разработка национальной биометрической платформы

Общей тенденцией развития информационного общества является создание больших облачных хранилищ данных. Облачный сервис удобен тем, что пользователь может обращаться к своим данным, оставаясь мобильным. Одним из перспективных направлений развития является переход к использованию электронных биометрических паспортов и удостоверений личности, хранящихся, например, на облачном сервере ФМС России.

У любого гражданина РФ появляется возможность не носить с собой паспорт или удостоверение личности. При необходимости гражданин РФ может доказать наличие у него прав, скачав свой электронный паспорт с облачного сервиса ФМС России. При этом электронные документы граждан должны храниться в обезличенной форме, что является одной из дополнительных степеней защиты информации. Настоящий владелец электронного биометрического паспорта знает его параметры (знает параметры его хранения) и легко может им воспользоваться. Злоумышленник же, напротив, для того, чтобы добраться до электронного паспорта нужного ему человека будет вынужден перебирать очень большое число возможных вариантов, вскрывая защиту каждого из проверяемых вариантов.

Естественно, что система обезличивание электронных паспортов может быть многоуровневой и иметь динамическую подсистему смены адресов хранения документов после каждого авторизованного обращения к облачному хранилищу. Принципиально важным является то, что обезличивание электронных документов и регулярная смена их адресов хранения в системе является достаточно эффективной дополнительной степенью защиты персональных данных.

Опытно-конструкторская работа «Лекарь» [11] (выполненная в 2014–2015 гг.) показала, что владелец электронного хранилища за счет обезличивания электронных документов многократно снижает требования к криптографическим механизмам защиты каждого отдельного электронного документа. Формально по ФЗ № 152

«О персональных данных» собственник облачного хранилища юридически не должен получать от пользователей их согласия на обработку персональных данных, содержащихся в обезличенных и криптографически защищенных нейросетевых контейнерах. В частности, ФМС России при обезличенном Интернет хранении электронных паспортов, защищенных отечественной криптографией перестает играть роль «оператора персональных данных» по ФЗ № 152. Оборудование, на котором развернуто облачное хранилище не нуждается в лицензировании.

Библиографический список

1. ГОСТ Р 52636–2006. Электронная история болезни. Общие положения.
2. Федеральная типовая медицинская информационная система (ФТМИС) [Разработчик «Крокус Консалдинг» 2008 г.] : государственный контракт по ФЦП «Электронная Россия» (2002–2010).
3. Электронное рабочее место врача. Руководство пользователя. – Москва, 2014.
4. Зингерман, Б. В. Интегрированная электронная медицинская карта: задачи и проблемы / Б. В. Зингерман, Н. Е. Шкловский-Корди, В. П. Карп, А. И. Воробьев // Врач и информационные технологии. – 2015. – № 1. – С. 24–27.
5. Костков, Д. Защита облачных вычислений: общие международные подходы / Д. Костков. // Первая миля. – 2015. – № 8. – URL: http://www.lastmile.su/journal/2015/8
6. О персональных данных : федер. закон № 152 от 27.07.2006.
7. Dodis, Y. Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy / Y. Dodis, L. Reyzin, A. Smith // Proc. EUROCRYPT, 2004. – April 13. – P. 523–540,
8. ГОСТ Р 52633.0–2006. Защита информации. Техника защиты информации. Требования к средствам высоконадежной биометрической аутентификации.
9. Об утверждении требований и методов по обезличиванию персональных данных : метод. рекомендации по применению приказа Роскомнадзора № 996 от 05.08.2013.
10. ГОСТ Р 52633.3–2011. Защита информации. Техника защиты информации. Тестирование стойкости средств высоконадежной биометрической защиты к атакам подбора.
11. Разработка аппаратно-программного комплекса обезличивания биометрических данных больных социально-значимыми заболеваниями, содержащихся в цифровых медицинских документах, обрабатываемых в медицинских информационных системах России и Беларуси [Разработчик – ФАУ «ГНИИИ ПТЗИ ФСТЭК России» (2014–2015) : государственный контракт по

Программе союзного государства «Совершенствование системы защиты общих информационных ресурсов Беларуси и России на основе высоких технологий» на 2011–2015 годы.

12. Болл, Руд. Руководство по биометрии : пер. с англ. / Руд Болл, Коннел Джонатан Х., Панканти Шарат, Ратха Налини К., Сеньор Эндрю У. – Москва : Техносфера, 2007. – 368 с.

13. Язов, Ю. К. Нейросетевая защита персональных биометрических данных / Ю. К. Язов, В. И. Волчихин, А. И. Иванов, В. А. Фунтиков, И. Г. Назаров. – Москва : Радиотехника, 2012. – 157 с.

14. ГОСТ Р 52633.2–2010. Защита информации. Техника защиты информации. Требования к формированию синтетических биометрических образов, предназначенных для тестирования средств высоконадежной биометрической аутентификации.

15. ГОСТ Р 52633.5–2011. Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа.

Корнеев, О. В. Обезличивание медицинских электронных документов при их хранении с привлечением облачных сервисов (технология SafeNet, разработка национальной биометрической платформы) / О. В. Корнеев // Безопасность информационных технологий : тр. I Всерос. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2019. – С. 66–80.

УДК 519.2, 612.087, 621.319.7

А. И. Иванов, Ю. С. Семерич

МАЙНИНГ КРИПТОВАЛЮТЫ КАК СПОСОБ НАКАПЛИВАНИЯ ИЗБЫТОЧНЫХ ВЫЧИСЛИТЕЛЬНЫХ РЕСУРСОВ ДЛЯ ОБЫЧНЫХ ПОЛЬЗОВАТЕЛЕЙ

Аннотация. Рассматривается вопрос анализа возможности накапливания избыточных вычислительных ресурсов обычных пользователей за счет майнинга криптовалюты, выполняемого параллельно решению обычных задач на компьютере пользователя. Предложено отказаться от использования обратных криптографических задач для организации «бесполезного майнинга» криптовалюты. Предложено организовывать «полезный майнинг», опираясь на дробление полезных для общества сложных вычислительных задач. Инициатор «полезного майнинга» должен иметь пакет сложных полезных задач, дробить их на фрагменты и раздавать «полезным майнерам». Контроль добросовестности майнеров выполняется дублированием их работы до полного совпадения отчетов о выполнении фрагмента задачи двумя или более майнерами. Политика учетности затраченных вычислительных ресурсов «полезными майнерами» поддерживается выпуском промежуточной криптовалюты, хождение которой является редким событием. Требования к блокчейн реестрам многократно снижаются.

A. I. Ivanov, Y. S. Semerich

MINING OF CRYPTOCURRENCY AS A METHOD OF ACCUMULATING EXCESS COMPUTATIONAL RESOURCES FOR ORDINARY USERS

Abstract. The article discusses the issue of analyzing the possibility of accumulating redundant computing resources of ordinary users by mining cryptocurrency, performed in parallel with solving common tasks on the user's computer. It is proposed to abandon the use of inverse cryptographic tasks for the organization of "useless mining" cryptocurrency. It was proposed to organize "useful mining", based on the fragmentation of complex computational problems useful to society. The initiator of "useful mining" should have a package of complex useful tasks, split them into fragments and distribute them to "useful miners." The control of the miners' conscientiousness is performed by duplicating their work until the reports on the fulfillment of a task fragment by two or more miners completely coincide. The policy of accounting, computational resources expended by “useful miners” is supported by the release of an intermediate cryptocurrency, the circulation of which is a rare event. Requirements for blockchain registries are repeatedly reduced.

Пирамида нерегулируемого майнинга криптовалют, технические издержки на блокчейн поддержку «старых криптовалют»

Одним из способов, задания условия синтеза криптовалюты является подбор значения случайной цифровой последовательности, хэширование, которой дает читаемый (понимаемый) человеком фрагмент. Эта ситуация отображена на рис. 1.

Рис. 1. «Бесполезный майнинг» криптовалюты, непрерывно потребляющий электроэнергию и требующий значительных затрат на аппаратную поддержку

Как только в выходном значении хэш-функции, появится понятное человеку словосочетание на одном из известных языков, обнаруженную комбинацию можно считать «криптомонетой». Стоимость «криптомонеты» может быть определена пропорционально ее редкости. В свою очередь вероятность обнаружения монеты может быть легко вычислена в пространстве расстояний Хэмминга [1–3].

К сожалению, майнинг этого типа «криптомонет» является высокозатратным для «поздних майнеров». «Первые майнеры» быстро находят «первые криптомонеты», так как их много. Однако по мере майнинга «криптомонет» и занесения их в блокчейн реестры их добыча замедляется. Чем больше обнаружено «криптомонет», тем труднее их добывать и тем труднее их поддерживать блокчейн реестрами. Добыча и поддержка криптомонет требует значительных затрат электроэнергии и значительных аппаратных затрат.

Ситуация меняется, если «криптомонеты» выпускает один или несколько эмитентов в ограниченном количестве под некоторый конкретный технический проект. Эта ситуация отображена на рис. 2.

Рис. 2. Эмиссия криптовалюты без затрат электроэнергии и без покупки «железа»

Как показано на рис. 2 эмиссия криптомонет может быть выполнена созданием осмысленного текстового файла, к которому добавлена случайная «соль» и этот текстовый файл должен быть охвачен цифровой подписью эмитентов криптомонеты. В этом случае для выпуска криптомонет нет необходимости тратить значительные вычислительные ресурсы. Стоимость выпуска «криптофлопов» снижается так же как снижается стоимость выпуска бумажных денег по сравнению с золотыми или серебряными монетами.

Структура системы обращения (выпуска и использования) криптофлопов приведена на рис. 3.

Следует подчеркнуть, что «бесполезный затратный майнинг» рис. 1 характерен тем, что «первые майнеры» вполне могут использовать свои первые криптомонеты для накапливания своих избыточных вычислительных ресурсов. На руках у множества пользователей находится значительный парк персональных компьютеров, которые в обычных условиях их эксплуатации недогружены. То есть «первые майнеры», кто тратит малые вычислительные ресурсы на добычу первых криптомонет могут использовать избыточный ресурс своих компьютеров на добычу «первых криптомонет». Потом, когда криптовалюта «постареет» майнить ее становится не выгодно. Она перестает играть роль накопителя криптофлопов.

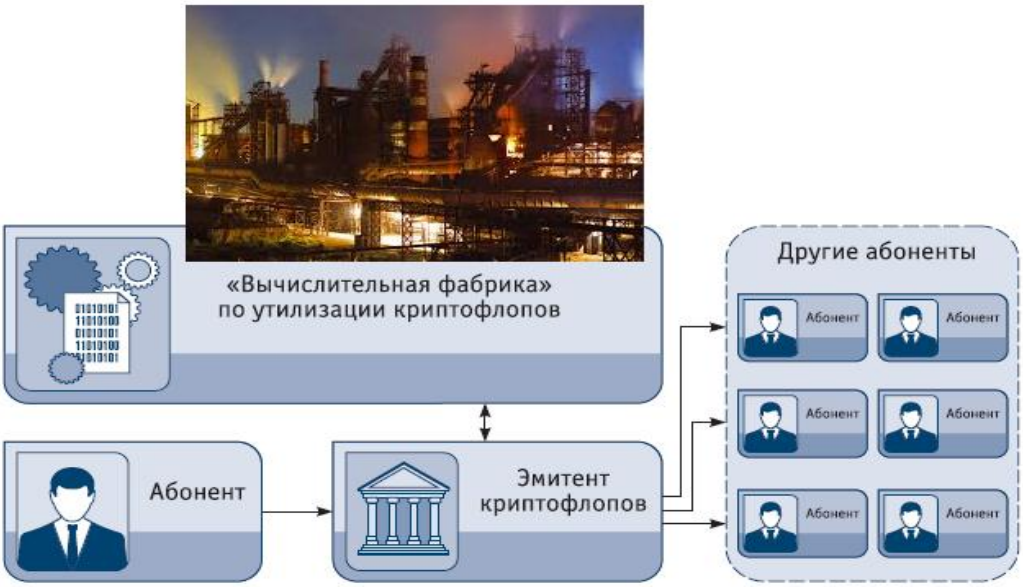


Рис. 3. Абонент, желающий накопить свои вычислительные ресурсы, обращается в банк, эмитирующий криптофлопы, за вычислительной задачей консервации

Положение меняется, если отказаться от майнинга, построенного на решении бесполезных задач. В этом случае организатор «полезного майнинга» должен иметь важную для всех полезную, но сложную в исполнении вычислительную задачу. Он должен уметь дробить полезную задачу на мелкие фрагменты и раздавать «полезным майнерам». Контроль добросовестности майнеров выполняется дублированием их работы до полного совпадения отчетов о выполнении фрагмента задачи двумя или более майнерами. Политика учетности, затраченных вычислительных ресурсов «полезными майнерами» поддерживается выпуском промежуточной криптовалюты, движение которой являются редкими событиями. Требования к блокчейн реестрам многократно снижаются.

В этой схеме взаимодействия эмитента и майнеров принципиальным является то, что задача стороннего ЗАКАЗЧИКА должна легко распараллеливаться на небольшие фрагменты для исполнения их «полезными майнерами». В этом случае система начинает работать. При этом эмитент криптофлопов не должен повторять работу за майнерами, проверяя их. Политика учетности, затраченных вычислительных ресурсов «полезными майнерами» поддерживается выпуском промежуточной криптовалюты, хождение которой является редкими событиями. Требования к блокчейн реестрам многократно снижаются из-за общения «полезных майнеров» только с эмитентом криптофлопов.

Еще одним важным условием является возможность получения коротких однозначных отчетов «полезных майнеров» о проделанной ими работе. Рисунок 4 иллюстрирует эту возможность на примере решения вычислительно сложной задачи факторизации длинных чисел.

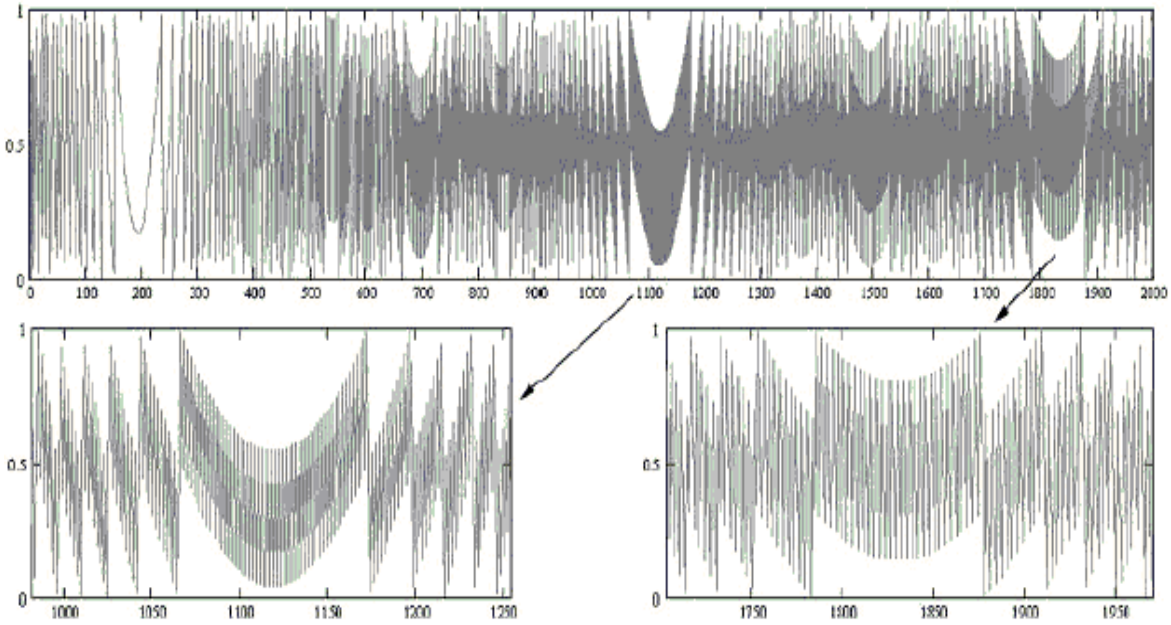


Рис. 4. Портреты остатков от деления числа N на целые числа, перебираемые при поиске произведения простых сомножителей ($pq = N$)

«Полезный майнер», получив свой фрагмент просмотра состояний, будет видеть «бабочки» сильно связанных между собой остатков. Он должен формировать список очень близких к решению примеров простых чисел. Эмитент криптофлопов должен сравнивать списки-отчеты двух и более «полезных майнеров». При совпадении списков, проверяемые майнеры признаются добросовестными.

Следует отметить, что «уплотнение» отчетов «полезных майнеров» может быть значительно увеличено, если перейти от линейных портретов чисел (рис. 4) к двухмерным портретам чисел [5], рис. 5. На этом рисунке видны 16 эллипсов, каждый эллипс описывается двумя параметрами, то есть, представленный портрет есть не что иное, как отображение 32-мерной задачи. Пользуясь этим можно многократно увеличить компактность отчетов о результатах работы «полезных майнеров».

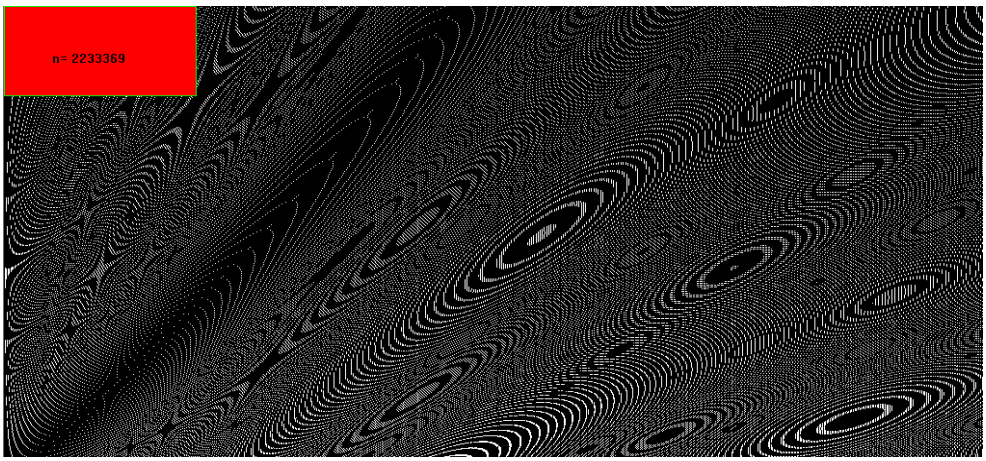


Рис. 5. Двухмерный портрет числа 2233369, полученный из остатков этого числа от деления на проверяемые сомножители, полученные в разных системах счисления

Запасать в прок свои вычислительные ресурсы выгодно в силу того, что «полезный майнер» получает право потратить их на решение его личной сложной задачи. Если личная задача «полезного майнера» легко распараллеливается, то у эмитента криптофлопов проблем не возникает. Он дробит личную задачу одного «полезного майнера» на множество фрагментов и раздает эти фрагменты другим «полезным майнерам». Однако личная задача «полезного майнера» может оказаться сильно связанной, задача ее распараллеливания может оказаться очень сложной сама по себе.

Для того, чтобы гарантированно принять (ликвидировать), выпущенные ранее криптофлопы эмитент должен иметь собственные вычислительные ресурсы, способные вернуть «полезным майнерам» их ранее законсервированные ресурсы при решении их личных задач, не поддающихся распараллеливанию.

На данный момент РЫНКА решения сложных личных задач пока нет. Видимо, он будет сформирован в будущем. В приведенной ниже таблице, даны оценки затрат в рублях на затраты электроэнергии при решении ряда сложных в вычислительном отношении задач. «Полезный майнер» нарабатывал криптофлопы на своем компьютере и, соответственно, он не должен тратить свои ресурсы на «железо». При утилизации криптофлопов для их владельца важен срок утилизации. Например, в первой строке табл. 1 рассматривается задача прогнозирования взаимного совместного поведения 30 пар валют на рынке операций FOREX. Пользователю, заказавшему такой прогноз интересен результат достоверный на время 1 часа, но полученный через 30 минут после запуска ВЫЧИСЛИТЕЛЬНОЙ ФАБРИКИ по утилизации криптофлопов.

Вторя строка табл. 1 отражает тестирование защищенных нейросетевых контейнеров. Эта задача уже не требует высокой скорости утилизации накопленных ранее криптофлопов. Для пользователя не имеет значения сутки или двое будет решаться его задача. Задачи нового рынка будут разнообразны, так в третьей строке таблицы даются оценки стоимости подбора забытых (утерянных) паролей от архиваторов. Актуальность решения таких задач будет расти по мере развития цифровой экономики. Потенциальная стоимость этого сегмента рынка огромна. Достаточно сказать, что 10 % биткоинов утрачены по техническим причинам и в том числе из-за утерянных паролей доступа к криптокошелькам. Видимо, в будущем криптокошельки для хранения криптовалюты будут биометрическими, однако это ослабит проблему, но не снимает ее. Родственники, получившие наследство в форме криптокошелька с биткоинами (возможные реалии цифровой экономики будущего), должны будут потратить часть наследства на открытие криптокошелька, если ЗАВЕЩАТЕЛЬ специально утаил часть цифр кода доступа [4].

Таблица 1

№ п/п	Содержание вычислительно сложной задачи	Необходимый ресурс в терафлопах в час и стоимость энергозатрат
1.	Прогноз соотношения пар валют на рынке FOREX, построенный с учетом совместного поведения 30 наиболее значимых валют в течении последнего года (30-ти мерный прогноз)	1000 Тф/ч, 26 тыс. руб.
2.	Полное тестирование собственного нейросетевого преобразователя биометрия-код, защищенного самошифрованием данных [4] **Распаковка БиоЗащиты архива с биткоинами**	10000 Тф/ч, 260 тыс. руб.
3.	Распаковка архива с утерянным (забытым) паролем:	
	• длинной 8 случайных символов	• 1 Тф/ч, 26 руб.
	• длинной 10 случайных символов	• 1000 Тф/ч, 26 тыс. руб.
	• длинной 12 случайных символов	• 1 000 000 Тф/ч, 26 млн руб.
4.	**Превращение малых выборок в большие для обычного статистического прогнозирования** **20 опытов в 200 опытов — 26 рублей** **30 опытов в 600 опытов — 52 рубля**	

Четвертая строка таблицы отражает сегмент перс перспективного рынка пересчета меленьких выборок в большие [6]. Увеличивать реальный объем тестовой выборки не всегда возможно технически или может оказаться слишком дорогим. Так фармакологическая компания, затратив миллиард на новое лекарство, для возврата вложений должная как можно быстрее выйти на рынок. При этом, необходимо проходить длительный этап тестирования нового лекарства. Если удастся сократить время тестирование в несколько раз, то фармакологическая компания быстрее выйдет на рынок и вернет затраченные средства.

Заключение

Майнинг и полный блокчен открытый реестр являются весьма и весьма энергозатратными и аппаратнозатратными технологиями. Гораздо менее энергозатратными и аппаратнозатратными являются технологии, построенные на выпуске криптомонет одним эмитентом, подкрепленных его капиталовложением в вычислительные ресурсы, ориентированные на решение полезных задач. Банк полезных, вычислительно емких задач пока не сформирован, однако они существуют и отражены в приведенной выше таблице. При этом выпускаемые эмитентом криптофлопы воспринимаются обычными пользователями как способ консервации их избыточных личных вычислительных ресурсов. Накапливание (консервирование) своих избыточных вычислительных ресурсов выгодно пользователям и выгодно эмитенту криптофлопов, так как он привлекает для решения «полезных» задач избыточные вычислительные ресурсы обычных пользователей.

Библиографический список

1. ГОСТ Р 52633.3–2011. Защита информации. Техника защиты информации. Тестирование стойкости средств высоконадежной биометрической защиты к атакам подбора.

2. Ахметов, Б. С. Алгоритмы тестирования биометрико-нейросетевых механизмов защиты информации / Б. С. Ахметов, В. И. Волчихин, А. И. Иванов, А. Ю. Малыгин. – Алматы : КазНТУ им. Сатпаева, 2013. – 152 с. – URL: http://portal.kazntu.kz/files/publicate/2014-01-04-11940.pdf

3. Малыгин, А. Ю. Быстрые алгоритмы тестирования нейросетевых механизмов биометрико-криптографической защиты информации / А. Ю. Малыгин, В. И. Волчихин, А. И. Иванов, В. А. Фунтиков. – Пенза : Изд-во ПГУ. – 161 с.

4. Иванов, А. И. Криптографическая валюта, пригодная для накапливания избыточных вычислительных ресурсов частными лицами / А. И. Иванов // Защита информации. INSAID. – 2014. – № 5. –С. 66–71.

5. Иванов, А. И. Многомерная нейросетевая обработка биометрических данных с программным воспроизведением эффектов квантовой суперпозиции / А. И. Иванов. – Пенза : ПНИЭИ. – 2016. – 133 с. – URL: http://пниэи.рф/activity/science/BOOK16.pdf

6. Иванов, А. И. Простейшие оракулы, обученные корректировать ошибки вычисления младших статистических моментов на малых выборках биометрических данных : учеб. пособие / А. И. Иванов. – Пенза : ПНИЭИ. – URL: http://пниэи.рф/activity/science/noc/BOOK18.pdf

Иванов, А. И. Майнинг криптовалюты как способ накапливания избыточных вычислительных ресурсов для обычных пользователей / А. И. Иванов, Ю. С. Семерич // Безопасность информационных технологий : тр. I Всерос. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2019. – С. 81–89.

УДК 519.7+621

С. В. Туреев, Е. А. Малыгина, А. И. Солопов

МЕТОДИКА ФОРМИРОВАНИЯ ТЕСТОВЫХ БАЗ ДЛЯ ПРОВЕРКИ КАЧЕСТВА ОБУЧЕНИЯ НЕЙРОСЕТЕВЫХ ПРЕОБРАЗОВАТЕЛЕЙ БИОМЕТРИЯ-КОД

Аннотация. Показано, что для тестирования средств высоконадежной биометрии необходимо создавать сбалансированные обезличенные базы биометрических образов с использованием естественных и синтезированных на их основе синтетических биометрических образов.

S. V. Tureev, E. A. Malygina, A. I. Solopov

METHODS OF FORMATION OF TEST BASES FOR TESTING THE QUALITY OF TRAINING OF NEURAL NETWORK CONVERTERS BIOMETRICS-CODE

Abstract. It is shown that to test highly reliable biometrics, it is imperative to create balanced anonymous bases of biometric images using synthetic biometric images that are natural and synthesized on their basis.

В связи с тем, что средства высоконадежной биометрической аутентификации [1, 2] в соответствии с национальным стандартом [3] имеют высокую стойкость к атакам подбора, то для полноты их тестирования возникает необходимость создания баз реальных биометрических образов сопоставимых по своим размерам со стойкостью тестируемых средств [4–6]. При решении этой непростой и трудоемкой задачи возникает целый ряд вопросов без решения, которых невозможно корректное решение самой задачи. Данные вопросы, на наш взгляд, можно выстроить в следующей последовательности:

1. Требования к испытуемым, при сборе биометрических образов.
2. Требования к аппаратуре преобразования физического образа испытуемого в электронный образ и программному обеспечению для обработки реальных биометрических образов.
3. Требования к создаваемым тестовым базам реальных биометрических образов.

Указанные выше вопросы неразрывно связаны между собой и некорректное решение одного из них может значительно исказить конечный результат.

Исследования, проводимые в межотраслевой лаборатории тестирования биометрических устройств и технологий ПГУ показали, что тестовая машина дает стойкость биометрической защиты к атакам подбора кода ключа порядка 10^{16}, что вполне достаточно для ряда практически значимых приложений, которые в настоящее или ближайшее время могут поступить на рынок. Как следствие, для тестирования такого средства прямой подстановкой потребуется база биометрических образов, содержащая на два-три порядка больше биометрических образов, чем заявленная производителем стойкость биометрической защиты [3].

Оценивая затраты времени и людских ресурсов на формирование подобных баз случайных биометрических образов [6] приходится учитывать то, что испытуемый при всей своей лояльности к средствам биометрической аутентификации, формируя данные должен выполнить ряд операций: осознать то, что он должен написать; написать рукописное слово (фразу); проконтролировать корректность рукописного ввода; ввести биометрический образ. Естественно, что все эти операции занимают определенное время. Хронометраж затрат времени показывает, что специально подготовленный человек (время предварительной подготовки занимает порядка 40 минут) вводит рукописные образы длиной по 5–6 символов за время порядка 10 секунд. Затраты времени на похожие операции ввода рукописных символов могут существенно различаться для людей разного темперамента и разного рода профессий. При этом рассматриваются только «идеальные» условия: удобные рабочие места с хорошей освещенностью, отсутствие внешних раздражителей и нормальное психофизиологическое состояние доноров биометрии. При этом, естественно, что люди, сталкивающиеся в своей профессиональной деятельности с необходимостью рукописных записей, вводят в ПЭВМ биометрические рукописные образы быстрее.

Несложные расчеты показывают, что для формирования базы из 10^{19} образов потребуется порядка 10^{14} лет работы одного донора биометрии. Естественно, что такие затраты времени и людских ресурсов не могут быть осуществлены.

Отметим, что при формировании биометрических голосовых образов затраты времени и ресурсов оказываются примерно такими

же. Голосовая информация вводится быстрее, однако ее приходится вводить больше. Голос обладает несколько меньшей информативностью, если сравнивать между собой одинаковые слова-пароли (парольные фразы).

Наряду с вышесказанным отметим, что в результате проведенных исследований при сборе биометрических данных выявлено то, что все реальные биометрические образы обладают некоторой нестабильностью. Именно неоднозначность, нестабильность, размытость биометрических образов является основной проблемой при их преобразовании в однозначный, четкий криптографический ключ [7]. Пример естественных вариаций нестабильности рукописного образа приведен на рис. 1.

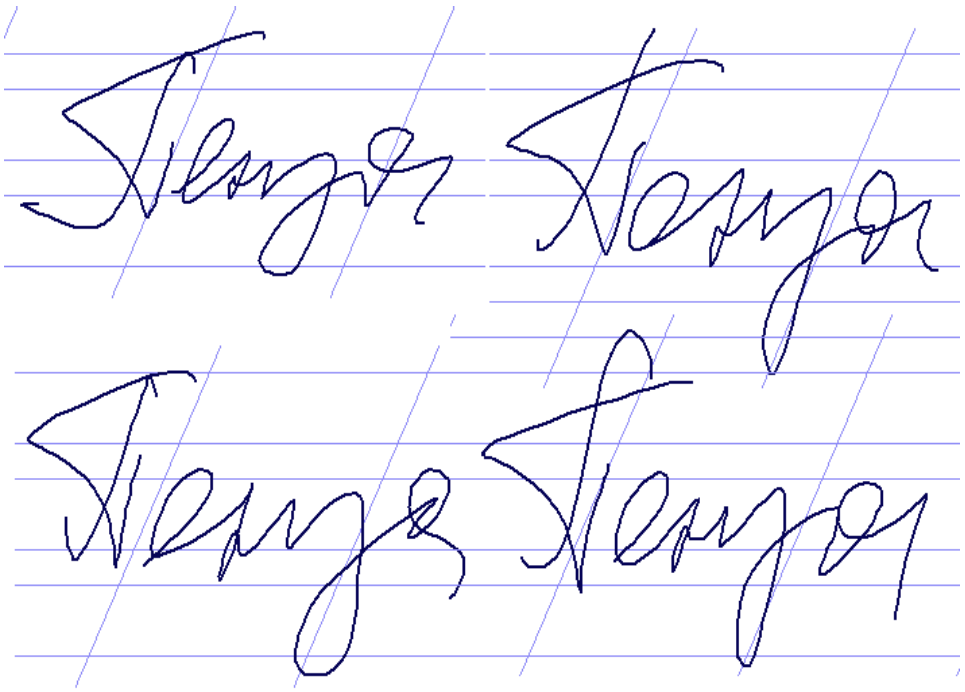


Рис. 1. Примеры естественной нестабильности биометрических образов

Как следствие этого явилась необходимость контролировать нестабильность воспроизведения пользователем его биометрического образа и классифицировать пользователей по их нестабильности.

Для того, чтобы оценить стабильность биометрического образ в целом, необходимо найти математическое ожидание всех дисперсий для каждого из контролируемых биометрических параметров. Сравнивая между собой математические ожидания дисперсий разных биометрических образов $m(б_1)$ и $m(б_2)$ можно сравнить стабильность воспроизведения разных биометрических образов.

Следует подчеркнуть, что стабильность воспроизведения биометрических образов является относительной величиной. Она определяется в первую очередь вектором контролируемых биометрических параметров. Во-вторую очередь она зависит от умения пользователя стабильно воспроизводить свой биометрический образ. Если это касается воспроизведения рукописных образов, то прежде чем тренироваться стабильно воспроизводить слово-пароль желательно попытаться изменить это слово, отыскивая наиболее стабильные для конкретного почерка сочетания рукописных букв.

Все пользователи для каждой конкретной биометрической системы могут быть классифицированы по стабильности воспроизведения ими их биометрических образов. Для формирования классификации необходимо оценить стабильность нескольких сотен пользователей и построить их нормированное распределение показателя стабильности. Эксперименты показали, что реальная гистограмма распределения пользователей по классам смещена в сторону низко стабильных пользователей с повышенными значениями дисперсий. Аппроксимация гистограммы нормальным законом распределения значений дает достаточно хорошую, точность, однако не учитывает естественную асимметрию реальных гистограмм.

Классификация пользователей по стабильности осуществляется делением распределения «Все Свои» пользователи на интервалы равные среднеквадратическому отклонению, так что бы в центре центрального интервала оказывалось математическое ожидание.

Очевидно, что стабильность воспроизведения биометрических образов является крайне важным статистическим показателем. Естественно требовать от формируемых баз биометрических образов того, чтобы они хорошо отражали реальную действительность по процентному содержанию биометрических образов, принадлежащих к разным классам стабильности.

При экспериментальном синтезе классифицирующего распределения пользователей по их стабильности следует обратить внимание на то, что оно должно строиться только для людей способных пользоваться биометрической защитой). Как правило, малые дети, старики, люди с нервными расстройствами имеют очень нестабильный рукописный почерк. Встроенный в средство аутентификации автомат, учитывающий большую нестабильность ввода рукописных образов, должен предупреждать таких пользователей о невозможности обеспечения надежной биометрической защиты.

Еще одним важным фактором применением классификации по стабильности является выявление случаев саботажа (сговора), когда часть пользователей намеренно ослабляет свою биометрическую защиту. Такая ситуация редка при защите пользователями своих интересов, но достаточно часто встречается при защите корпоративных данных. Для прекращения саботажа служащих, как правило, достаточно самого факта его выявления и объяснения служащему негативных для него последствий его же действий (несоответствие занимаемому служебному положению, по квалификации, психофизиологическому состоянию, нелояльность к корпоративной биометрии).

Не менее важна для систем биометрической защиты степень уникальность биометрического образа «Свой». Очевидно, что злоумышленник всегда будет стараться выбирать при атаках подбора наиболее вероятные состояния входов биометрической защиты. То есть, злоумышленник имеет модель среднестатистического «Своего» (или модель «Все Чужие») и будет стараться использовать ее при организации атак. Естественно, что чем больше образ «Свой» будет отличаться от среднестатистического образа, тем выше степень биометрической защиты.

Интуитивно понятно, что уникальность одного биометрического параметра можно оценить через вероятность случайного попадания в интервал «Свой» при эмуляции данных «Все Чужие». Эта вероятность в рамках гипотезы нормальности законов распределения значений будет описываться следующим выражением:

$$P_c = \frac{1}{\sqrt{2\pi}\sigma_B} \int_{m_c-3\sigma_c}^{m_c+3\sigma_c} \exp\left(\frac{(m_B - x)^2}{2\sigma_B^2}\right) dx, \qquad (1)$$

где m_c, $б_c$ – математическое ожидание и дисперсия распределения параметра «Свой»; m_B, $б_B$ – математическое ожидание и дисперсия распределения параметра «ВСЕ Чужие».

Очевидно, что вероятность (1) будет уменьшаться при снижении дисперсии «Своего», а также при вытеснении центра множества «Свой» на периферию распределения «Все Чужие».

Уникальность конкретного биометрического параметра будет описываться обратной величиной вероятности (1). Так как анализируемых биометрических параметров много необходимо оценивать среднюю уникальность этих параметров или меру уникальности всего биометрического образа:

$$U = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{P_{C,i}}.$$

Уникальность, как и любой иной значимый статистический параметр, может быть использована для классификации биометрических образов. Необходимо отметить, что различные биометрические образы обладают разной информативностью (разной сложностью). Очень простые биометрические образы легко подбираются и не могут обладать необходимой стойкостью. Выделяют статические – неизменные биометрические образы, данные человеку от рождения [1]. К ним относятся рисунки отпечатков пальцев, дерево сосудов глазного дна, радужная оболочка глаз, геометрия ладони, геометрия лица. Как правило, человек не может по своему желанию изменять (усложнять) свой статический биометрический образ. Как следствие статический образ легко компрометируется и обладает ограниченной информативностью.

Свои динамические биометрические образы, напротив, человек легко может изменить. Например, рукописный образ слова пароля, легко может быть изменен сменой самого пароля [1]. Динамические биометрические образы могут быть изменены (усложнены) при необходимости. Их можно сохранять в тайне и за счет усложнения повышать их стойкость к атакам подбора. Для динамических биометрических образов (рукописный и голосовой почерк, клавиатурный почерк) в [3] приводятся только ограничения снизу на длину эквивалентного ключа. Ограничений с верху на этот параметр для этих технологий нет, однако сложность самого биометрического образа и эффективная длина ключа связаны. Чем сложнее биометрический образ, тем сложнее его подбор. Это непреложное правило справедливое для любых (статических или динамических) биометрических образов.

В силу того, что стойкость биометрических образов прямо зависим от их сложности и эта сложность может быть достаточно просто оценена, необходимо при формировании больших баз биометрических образов их балансировать по сложности биометрических образов. То есть базы рукописных и голосовых биометрических образов должны содержать число слов из 5 букв (число отпечатков пальцев с 22 особенностями) в процентном отношении столько же, сколько их содержится в естественном языке (в естественном распределении рисунков отпечатков пальцев).

Рассмотрим требования к качеству преобразователей биометрических образов физического уровня в биометрические электронные образы.

Разнотипные высоконадежные распознаватели во многом сходны. Обобщенная схема устройств высоконадежной биометрической аутентификации личности человека приведена на рис. 2.

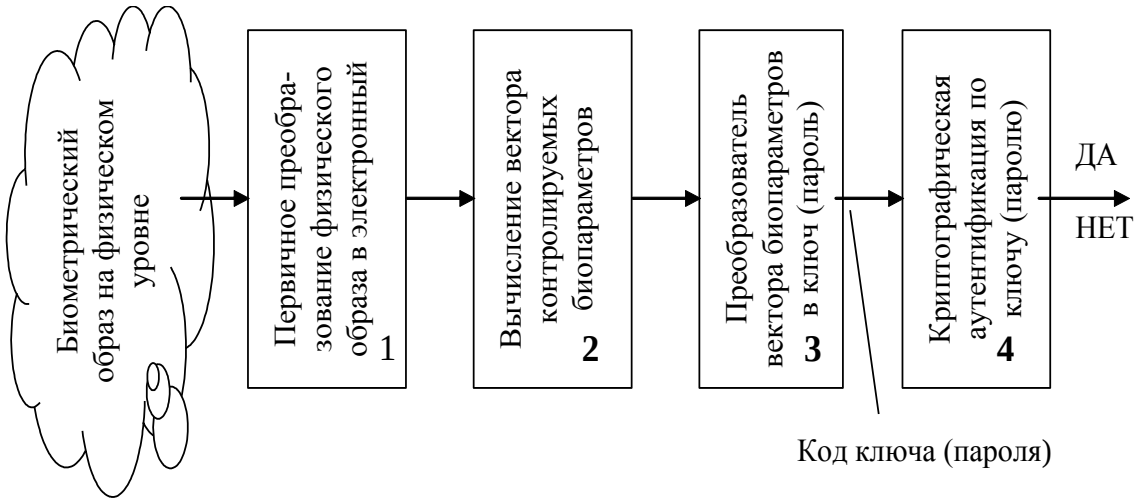


Рис. 2. Структурная схема типового высоконадежного аутентификатора личности человека

В структурной схеме рис. 2, блок-1 осуществляет преобразование физического нечеткого биометрического образа человека в электронный биометрический нечеткий образ через первичные преобразователи физических величин в электронные цифровые данные. Блок-2 осуществляет нормировку электронных образов и вычисление вектора биометрических параметров, например, в виде коэффициентов Фурье, в средствах аутентификации по динамике воспроизведения рукописного пароля. Блок-3 осуществляет преобразование вектора биометрических параметров в код ключа (пароля) для последующей криптографической аутентификации. Блок-4 осуществляет криптографическую аутентификацию пользователя по его ключу или паролю, выдавая на выход решение «ДА/НЕТ».

Очевидно, что при формировании баз биометрических образов рукописных, голосовых, отпечатков пальцев, … необходимо использовать именно те датчики, которые использует конкретная система биометрической идентификации. Однако такой подход может быть использован только при тестировании относительно «слабой» биометрической защиты. Для «слабой» биометрии достаточно формировать тестовые базы, состоящие из малого числа

биометрических образов. Как следствие всегда можно повторить работу по формированию малых баз биометрических образов заново для любого датчика.

При формировании больших баз биометрических образов такой подход неприемлем. Из-за большой трудоемкости этой работы базы должны быть универсальны. То есть при их формировании необходимо использовать наиболее точные на текущий момент средства ввода рукописной графики и звука. Это делается из-за того, что может возникнуть необходимость тестирования очень точных систем аутентификации, тогда собранная информация будет использована в том виде, в каком хранится. Если потребуется снизить точность преобразований физических биометрических образов в электронные биометрические образы, то это всегда можно сделать с помощью специально написанного программного конвертора. Из хороших данных путем ввода «шума» всегда можно сделать плохие данные. Обратная операция просто невозможна.

Таким образом, при формировании больших и универсальных баз биометрических образов требуется использовать как можно более точные и как можно более информативные преобразователи. Для формирования рукописных образов на сегодня подходят для этой цели практически все графические планшеты способные воспринимать 1024 градаций степени нажатия пера и имеющие разрешение 4064 линий на дюйм при рабочем поле 5,5x4 дюйма. Требований к шумам и линейности преобразователей не предъявляется.

При формировании голосовых баз биометрических образов должна использоваться звуковая карта способная оцифровывать звук с частотой 44 кГц, в режиме стереозаписи при дискретизации АЦП по уровню не менее 10 разрядов. Должен использоваться широкополосный микрофон, закрепленный на гарнитуре и второй широкополосный микрофон, закрепленный перед пользователем. При записи голосовых парольных фраз должен быть обеспечен низкий уровень посторонних шумов. Очевидно, что из-за требования к низкому уровню сторонних шумов формировать голосовые базы данных сложнее, так как потребуются помещения подобные студиям звукозаписи.

Немаловажным аспектом при формировании «эталонных» баз биометрических данных являются требования к программному обеспечению автоматизированного формирования базы биометрических тестовых образов.

В связи с высокой трудоемкостью формирования больших баз биометрических образов необходимо стремиться к всемерной экономии затрачиваемых ресурсов. При формировании малых баз возможно привлечение квалифицированного инструктора, под руководством которого каждый пользователь будет выполнять свою работу. Такой подход при формировании больших бах биометрических образов, неприемлем, так как существенно увеличивает их стоимость.

Как показали проведенные исследования при формировании больших баз биометрических образов допустимо привлечение квалифицированного инструктора на этапе обучения тестируемого (инструктаж и пробная работа под контролем). Далее программное обеспечение ввода биометрических образов должно заменить тестируемого инструктора. Оно должно быть максимально доступно для понимания тестируемым, иметь минимум режимов управления и контролировать все действия тестируемого. При отклонениях тестируемого от заданного оптимального режима программное обеспечение должно автоматически сообщать тестируемому о его ошибке. Например, говорите чуть громче или пишите крупнее и быстрее. Сообщения тестируемому должны выдаваться в голосовой и графической формах [7].

При этом особые требования предъявляются также и к программным средствам формирования парольных слов (фраз). Тестируемый должен воспроизводить только заданные ему образы (рукописные и голосовые). Генераторы парольных слов и фраз программного обеспечения должны иметь возможность взаимной синхронизации из центра. Эти генераторы должны быть способны воспроизводить заданную из центра последовательность слов (фраз) или формировать независимую последовательность случайных слов (фраз).

При формировании рукописных и голосовых парольных образов допускается использование словарей, допускается усиление слов словарей их вариациями, принятыми в языке тестируемого, допускается их усиление числами в форме, принятой при написании или голосовом воспроизведении.

Все данные о поведении тестируемого и формировании его биометрического образа обязательно документируются. Программное обеспечение должно вести автоматический журнал регистрации полученной биометрической информации и времени ее получения [4].

При формировании баз биометрических образов должна быть обеспечена анонимность тестируемых с тем, чтобы их биометрическая информация в соответствии с ФЗ № 152 «О персональных данных» [8] не могла быть использована кем-либо против них в настоящем или будущем времени.

Хочется отметить, что ценность естественных баз биометрических образов состоит в том, что они хорошо отражают реальное распределение биометрических признаков среди людей. Обычно предполагают, что чем больше база биометрических образов, тем она точнее отражает действительность. К сожалению, эксперименты, проведенные в Лаборатории при выполнении целого ряда работ, показали, что это не всегда так. Вернее, это может быть так, если представители различных возрастных групп людей, различных профессий и различных темпераментов представлены в тестовых группах в тех же пропорциях, что и в обществе. Добиться подобной представительности крайне сложно. В связи с этим необходимо проводить специальные исследования, позволяющие оценить представительность полученной тестовой выборки. Под каждый тип биометрических образов должны быть сформированы свои критерии представительности тестовой выборки.

Если критерии представительности тестовой выборки построены, то представительность самой выборки может не зависеть от ее размеров. В частности, может быть искусственно сформирована малая по размерам тестовая выборка, хорошо удовлетворяющая по ее представительности вектору критериев представительности. Эта выборка должна строиться таким образом, чтобы с одной стороны она состояла из реальных или правдоподобно синтезированных биометрических образов [9–11], а с другой стороны она должна верно отражать (отображать с заданной погрешностью) вектор выбранных критериев представительности.

В качестве критериев представительности могут выступать:

– статистические характеристики среднестатистического пользователя по некоторому биометрическому параметру (математическое ожидание, среднеквадратическое отклонение, коэффициенты корреляции и другие статистические моменты);

– статистические характеристики среднего по некоторой группе пользователя, выделенной по некоторому биометрическому параметру (математическое ожидание, среднеквадратическое отклонение, коэффициенты корреляции и другие статистические моменты);

– представительность по численности групп пользователей, классифицированных по некоторому биометрическому параметру.

В качестве биометрических параметров могут быть использованы любые параметры, например, стабильность, уникальность, стойкость к атакам подбора.

Заключение

Таким образом, для формирования естественных тестовых баз данных для тестирования средств высоконадежной биометрической аутентификации необходимы большие материальные, людские и временные затраты. Однако без их создания при тестировании нельзя получить полную характеристику стойкости исследуемого устройства. Выходом из создавшегося тупика может стать параллельная работа по созданию реальных и синтезированных на основе реальных искусственных баз биометрических образов, отражающих основные характеристики реальных биометрических образов.

Библиографический список

1. Ахметов, Б. С. Основы биометрической аутентификации личности : учеб. пособие / Б. С. Ахметов, А. И. Иванов, А. Ю. Малыгин, В. А. Фунтиков. – Алматы : Изд-во КазНТУ им. К. И. Сатпаева, 2014. – 151 с.

2. Волчихин, В. И. Быстрые алгоритмы обучения нейросетевых механизмов биометрико-криптографической защиты информации : монография / В. И. Волчихин, А. И. Иванов, В. А. Фунтиков. – Пенза : Изд-во ПГУ, 2005. – 276 с.

3. ГОСТ Р 52633.0–2006. Защита информации. Техника защиты информации. Требования к высоконадежным средствам биометрической аутентификации. – Москва : Стандартинформ, 2007.

4. Малыгин, А. Ю. Формирование больших и сверхбольших баз биометрических образов для тестирования средств высоконадежной биометрической защиты / А. Ю. Малыгин // Вопросы защиты информации. – 2007. – № 4 (79). – С. 17–21.

5. Волчихин, В. И. О проблеме ресурсов при тестировании стойкости высоконадежных биометрических технологий / В. И. Волчихин, А. Ю. Малыгин, М. Ю. Лупанов, А. В. Семенов // Вопросы защиты информации. – Москва : ВИНИ, 2005. – № 4. – С. 15–16.

6. Малыгин, А. Ю. Оценка размеров технически реализуемых баз биометрических образов, необходимых для корректного тестирования высоконадежных нейросетевых преобразователей / А. Ю. Малыгин, В. И. Волчихин, В. В. Федулаев, А. В. Безяев // Нейрокомпьютеры: разработка, применение. – Москва : Радиотехника, 2007. – № 12. – С. 52–54.

7. Язов, Ю. К. Основы теории безопасного преобразования биометрических данных в код личного ключа доступа : учеб. пособие [Электронный ресурс] / Ю. К. Язов, О. А. Остапенко, А. И. Иванов, В. А. Фунтиков. – Воронеж : Воронеж. гос. техн. ун-т, 2014.

8. О персональных данных : федер. закон № 152-ФЗ от 27.07.2006 : [в ред. от 29.07.2017]. – URL:https://fzakon.ru/laws/federalnyy-zakon-ot-27.07.2006-n-152-fz/?yclid=539543931311628356

9. Малыгин, А. Ю. Требования к синтетическим базам биометрических образов и генераторам для их формирования / А. Ю. Малыгин, В. В. Федулаев, Д. Н. Надеев // Нейрокомпьютеры: разработка, применение. – Москва : Радиотехника. – 2007. – № 12 . – С. 60–64.

10. Оценка рисков высоконадежной биометрии : монография / Б. С. Ахметов, Д. Н. Надеев, В. А. Фунтиков, А. И. Иванов, А. Ю. Малыгин. – Алматы : Из-во КазНТУ им. К. И. Сатпаева, 2014. – 108 с.

11. ГОСТ Р 52633.2–2010. Защита информации. Техника защиты информации. Требования к формированию синтетических биометрических образов, предназначенных для тестирования средств высоконадежной биометрической аутентификации. – Москва : Стандартинформ, 2011.

Туреев, С. В. Методика формирования тестовых баз для проверки качества обучения нейросетевых преобразователей биометрия-код / С. В. Туреев, Е. А. Малыгина, А. И. Солопов // Безопасность информационных технологий : тр. I Всерос. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2019. – С. 90–101.

УДК 519.2, 612.087, 621.319.7

А. В. Безяев

НЕЙРОСЕТЕВАЯ МОЛЕКУЛА: МЕХАНИЗМ НАПРАВЛЕННОЙ КВАНТОВОЙ КОРРЕКЦИИ БОЛЬШОГО ЧИСЛА ОШИБОК ДЛИННОГО КОДА ВЫСОКОРАЗМЕРНЫХ БИОМЕТРИЧЕСКИХ ОБРАЗОВ

Аннотация. Целью работы является создание высокоэффективных механизмов корректировки большого числа ошибок в длинных выходных кодах нейросетевых преобразователей. Предложено нейросетевой преобразователь биометрии рассматривать как нейросетевую молекулу и поддерживать для нее режим нейродинамики за счет добавления на ее входы «белого шума». В режиме нейродинамики предложено контролировать показатели стабильности каждого из 256 кубит выходного кода; ускорить подбор верного кодового состояния за счет корректировки наиболее нестабильных разрядов кода.

A. V. Bezyaev

NEURAL NETWORK MOLECULE: THE MECHANISM OF THE DIRECTED QUANTUM CORRECTION OF A LARGE NUMBER OF ERRORS OF THE LONG CODE OF HIGH-DIMENSIONAL BIOMETRIC IMAGES

Abstract. The aim of the work is to create highly efficient mechanisms for correcting a large number of errors in the long output codes of neural network transformers. It is proposed to consider the neural network biometrics converter as a neural network molecule and support the neurodynamics mode for it by adding white noise to its inputs. In the neurodynamics mode, it is proposed to monitor the stability indicators of each of the 256 qubit output code. It was proposed to accelerate the selection of the correct code state by adjusting the most unstable code bits.

Необходимость корректировки ошибок в кодах, получаемых из биометрических данных

Информатизация современного общества приводят к необходимости расширения применения криптографии. Обычные люди не могут запоминать длинные пароли доступа и криптографические ключи. Для решения этой проблемы в США и Евросоюзе развива-

ются технологии «нечетких экстракторов» [1–3], построенных на корректировке ошибок классическими кодами с обнаружением и исправлением ошибок. При этом выходной код «нечетких экстракторов» является коротким из-за того, что классические коды с приемлемой избыточностью в 50 % способны корректировать не более 5 % ошибок [4, 5]. Ошибки исходных кодов «нечетких экстракторов» могут составлять от 20 до 30 % от длины кода, что заставляет использовать самокорректирующиеся коды с 20-ти кратной избыточностью. То есть длина выходного кода «нечеткого экстрактора» оказывается в 20 раз меньше, чем число биометрических параметров, из которых «нечеткий экстрактор» восстанавливает код ключа.

В России развивается технология нейросетевого преобразования биометрии в длинный код доступа или длинный код личного криптографического ключа [5, 6]. Нейросетевые преобразователи биометрия-код, выполненные в соответствии с пакетом стандартов ГОСТ Р 52633.xx [7, 8], обучаются на выборках порядка 20 примеров образа «Свой». При этом вероятность ошибок первого рода (отказ в доступе «Своему») составляет от 0.05 до 0.1. Для снижения вероятности ошибок первого рода необходимо либо увеличивать число примеров в обучающей выборке, либо осуществлять коррекцию ошибок. Для нейросетевых преобразователей биометрия-код нет проблемы коротких выходных кодов так как за один разряд кода отвечает один нейрон, а число нейронов выбирает разработчик биометрического средства защиты информации. Тем не менее, выходные коды нейросетевых преобразователей нуждаются в корректировке.

Корректирующие коды, безопасно хранящие информацию о синдромах ошибок в коротких фрагментах хэш-функций

В биометрических кодах может содержаться до до 10–15 ошибок, если строить коды, способные обнаруживать до 16 ошибок, то весь код следует разбить на 16 фрагментов по 16 бит. Структурная схема хэш-корректора приведена на рис. 1.

Корректирующий код построен на том, что защищаемая последовательность в 256 бит делится на 16 фрагмента по 16 бит. Для верного состояния защищаемого кода вычисляются хеш-функции для 16, 32, 48,…., 256 бит [9, 10]. Три бита каждой хеш-функции запоминают в таблицу хеш-остатков, например, это могут быть последние 3 бита каждой из хеш-функций.

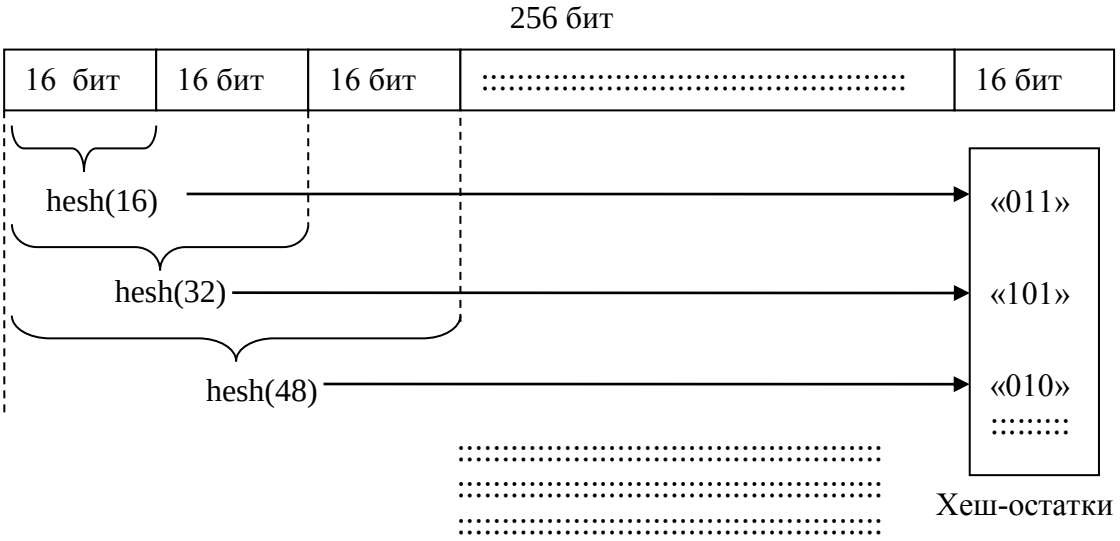


Рис. 1. Безопасная схема рекурсивного формирования эталонных хеш-остатков

Такой код способен корректировать до трех ошибок в каждом из выделенных 16 фрагментах. Корректировка выполняется путем перебора состояний всех возможных положений трех ошибок. Всего приходится проверять: $C_{16}^{3} = \frac{16!}{3!(16-3)!} = 560$ состояний для каждого из 16-битных фрагментов кода. По такой схеме, крайне редко, но все-таки возможно скорректировать 3x16=48 ошибок. Вероятность коррекции столь большого числа ошибок мала и составляет менее 0.00001. Однако, если число ошибок не более 16 и все они попали в разные фрагменты кода, то все они однозначно исправляются.

Так как на каждые 16 бит кода хранится 3 бита информации, в первом приближении можно считать, что эти три бита скомпрометированы. То есть стойкость к атакам подбора кода ключа длинной в 256 бит не может быть выше 208 бит.

Поддержка квантовой суперпозиции для наблюдения показателей стабильности разрядов кода нейронной сети

При выполнении требований пакета национальных стандартов ГОСТ Р 52633.xx энтропия кодов преобразователя приводит к тому, что энтропия кодов примеров образа «Свой» мала:

$$H("c_1, c_2, ..., c_{256}") \approx 0{,}05 \text{ бит.} \qquad (1)$$

Это связано с тем, что данные примеров образа «Свой» находятся внутри многомерного эллипса «Свой». При обучении преобразователя по ГОСТ Р 53633.5 [8] каждый нейрон будет давать разделяющую гиперплоскость, причем все гиперплоскости нейронов пересекаются в центре области все «Чужие». Сечение многомерного пространства по паре любых биометрических параметров дает плоскость, отображенную на рис. 2, где видно, что примеры 1, 2, …., 7 образа «Свой» находятся внутри эллипса, полученного при обучении.

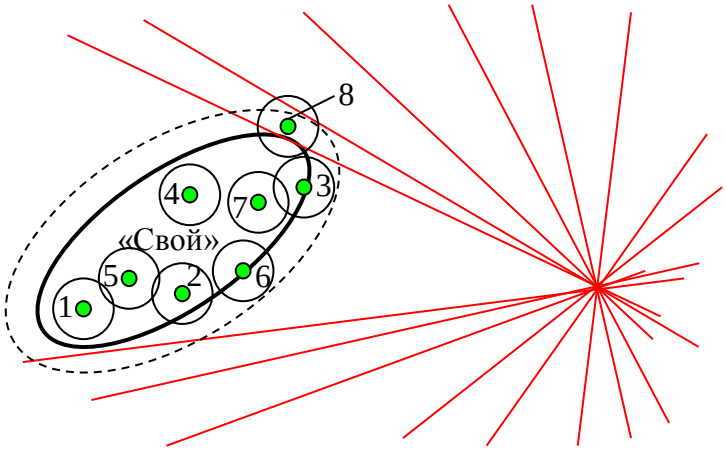


Рис. 2. Обнаружение нестабильных разрядов кода путем добавления тестового «белого шума» во входные данные

Пример-8 находится вне эллипса «Свой» за одной из линий, являющихся проекциями разделяющих гиперплоскостей нейронов. Это означает, что в одном из 256 разрядов кода появится ошибка при предъявлении примера-8 образа «Свой».

Расстояние между примерами и прямыми линиями рис. 2 отражает стабильность разрядов кода. Для того, чтобы оценить стабильность разрядов кода, предъявленного нейронной сети примера необходимо добавить к данным примера «белый шум», как это показано на рис. 3. На рис. 2 добавление «белого шума» приводит к появлению окружностей, накрывающих точку каждого из примеров. Из рис. 2 видно, что окружность шума, накрывшая пример8, пересекает две прямые, что приводит к нестабильности не менее двух бит выходного кода нейронной сети. Другие окружности примеров 1, 2, …, 7 не пересекают разделяющих прямых, то есть эти примеры будут давать стабильные разряды кода [11].

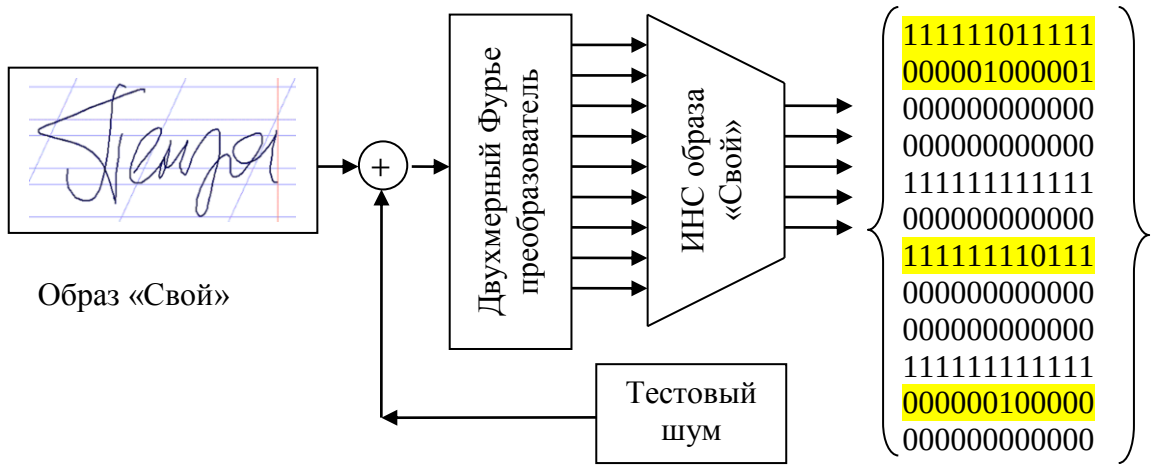


Рис. 3. Схема поддержки квантовой суперпозиции на выходах нейронной сети преобразователей биометрия-код

Для того, чтобы определить уровень нестабильности каждого из разрядов кода необходимо получать множество выходных кодов, как это показано на рис. 3. Поданный на вход нейросети аддитивный белый шум дает множество выходных кодов, при этом полностью стабильные разряды не меняют своего состояния. В нестабильных разрядах состояния меняются.

Для каждого разряда может быть вычислен показатель стабильности:

$$w_i = 2\left|P("0_i") - 0.5\right| = 2\left|P("1_i") - 0.5\right| \qquad (2)$$

где i – номер контролируемого разряда, $P("0_i")$ – вероятность появления состояния «0» в i-том разряде кода, $P("1_i")$ – вероятность появления состояния «1» в i-м разряде кода.

Показатель стабильности может меняться в интервале от 0 до 1, при $w_i = 0$ состояние i-го разряда абсолютно нестабильно (имеем полноценный кубит). При $w_i = 1$ состояние i-го разряда абсолютно стабильно (квантовая суперпозиция полностью отсутствует). На рис. 3 стабильные разряды даны без заливки, а нестабильные разряды помечены заливкой.

Нейросетевая молекула, поддержка ее состояний в нейродинамике

В 1980-х гг. наш соотечественник Юрий Манин занимался струнами и предложил концепцию организации вычислений [12, 13], осуществляемых на квантово-механических вычислитель-

ных элементах нового поколения. За последующие 30 лет развития эта новая концепция активно углублялась, была создана полноценная квантовая математика под квантовую механику уравнения Шредингера. Были найдены эффективные квантовые алгоритмы полиномиальной вычислительной сложности. Например, Питер Шор в 1994 г. создал квантовый алгоритм под решение задачи поиска простых чисел по их произведению (обратная задача для RSA алгоритма шифрования). Возник значительный интерес мировой научно-технической и криптографической общественности к практическим реализациям квантовых вычислений. Однако, исследования квантовых алгоритмов показали, что решение уравнений Шредингера на обычном компьютере становится технически невозможным уже при 30 степенях свободы, так как эти алгоритмы имеют экспоненциальную вычислительную сложность. В рамках квантовой механики сформулирован постулат о невозможности симуляции достаточно большого числа кубит на обычном компьютере.

Ситуация меняется, когда речь идет о моделировании искусственных нейронных сетей, корреляционных молекул, хи-квадрат молекул [14–16]. В этом случае задача моделирования имеет линейную вычислительную сложность.

Применительно к задаче коррекции кодов мы имеем дело с режимом поддержки нейродинамики. Нейронную сеть можно рассматривать как нейросетевую молекулу, откликающуюся спектром выходных дискретных (цифровых) состояний на зашумление входных данных образа «Свой».

На рис. 4 приведена структура нейросетевой молекулы преобразующей континуумы входных состояний в спектр последовательности цифровых кодов.

Нами был рассмотрен только один режим поддержки квантовой суперпозиции через зашумление входов нейросетевой молекулы (рис. 3), однако это не единственный способ. При решении обратной задачи нейросетевой биометрии [14] приходится скрещивать между собой образы «Чужой» и тем самым поддерживать режим нейродинамики преобразователя биометрия-код. Когда речь идет о хи-квадрат молекулах [16, 17] поддержки квантовых эффектов в нейродинамике, приходится поддерживать нейродинамику, случайно выбирая данные серий малых выборок из одной большой выборки.

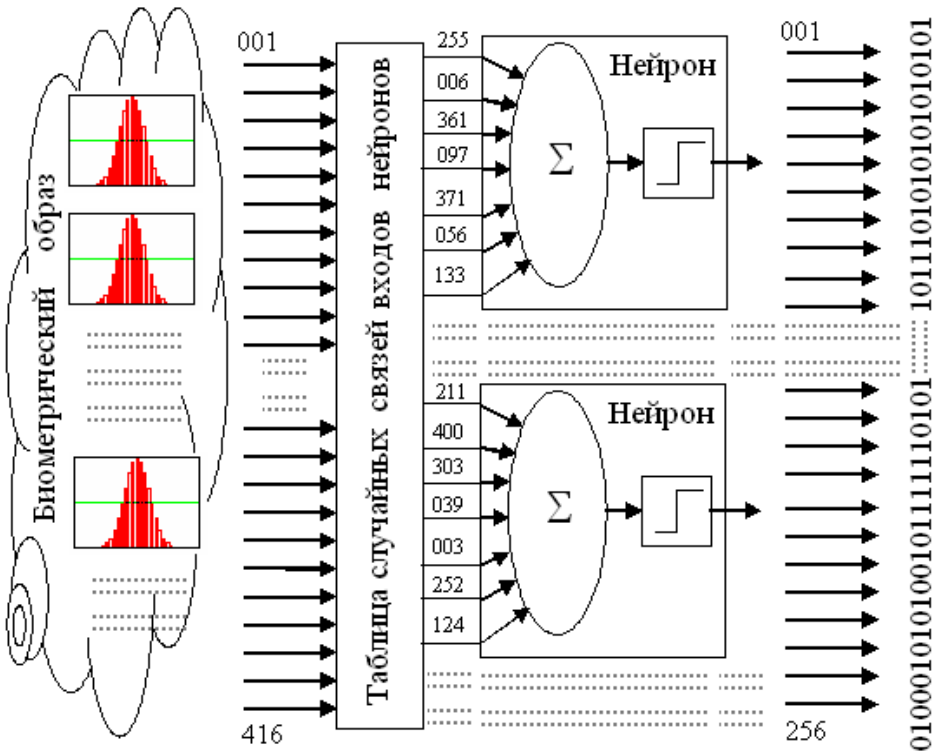


Рис. 4. Нейросетевой преобразователь, находящийся в режиме поддержки нейродинамики

Главное состоит в том, что моделирование нейронов (сумматоров и квантователей) является задачей линейной вычислительной сложности. Если уравнения Шредингера с большим числом степеней свободы нельзя воспроизводить на обычной вычислительной машине, то уравнения нейросетевой молекулы или хи-квадрат молекулы легко воспроизводятся на обычных компьютерах при большом числе степеней свободы. В нашем случае нейросетевой преобразователь биометрия-код с 416 входами и 256 выходами на обычной вычислительной машине удается воспроизводить до 10 000 раз в секунду. За одну секунду мы получаем 10 000 кодов-откликов на входной вектор «Свой», окрашенный «белым шумом».

Преимущества, получаемые за счет поддержки квантовой суперпозиции

Схема самокорректирующихся кодов с хэш-остатками (рис. 1) способна корректировать некоторое число ошибок меньшее, чем число запомненных разрядов остатков. При этом, наблюдая расхождение эталонных хэш-остатков и реально вычисленных хэш-остатков, мы не можем указать положения ошибок при наличии ошибки в первом фрагменте кода.

Однако если мы воспроизведем квантовую суперпозицию и оценим показатели стабильности разрядов кода, то нам удастся определить число и положение «слабых» бит кода. Зная эту информацию, мы можем значительно ускорить поиск и расширить область проверяемых состояний кода.

Например, если мы обнаружили в первых 16 разрядах 5 «слабых» бит, а в следующих трех фрагментах с 17-го бита до 64 бита нет «слабых» разрядов мы можем однозначно восстановить код перебирая $C_{16}^{5} = 4368$ состояний, контролируя при этом 4x3=12 эталонных бит Хэш-остатков.

Ранее использовавшиеся коды, без контроля стабильности разрядов в нейродинамике [9], были не способны корректировать ошибки, группирующиеся в начале схемы корректировки (рис. 1). В нашем же случае даже группировка 8 ошибок в первых 16 битах и 8 ошибок в следующих 16 битах однозначно корректируется при отсутствии ошибок в последующих фрагментах кода.

Выигрыш по времени при корректировке 16 бит при известном положении «слабых» разрядов в сравнении с проверкой по три возможных положения ошибок в каждом из 16 фрагментов кода может составить до 20 000 раз. То есть самокорректирующиеся коды с поддержкой квантовой суперпозиции способны корректировать до 16 ошибок в коде в реальном масштабе времени ввода и обработки биометрических данных.

Библиографический список

1. Juels, A. A Fuzzy Commitment Scheme / A. Juels, M. Wattenberg // Proc. ACM Conf. Computer and Communications Security. – 1999. – November 1–4. – P. 28–36.

2. Ramírez-Ruiz, J. Cryptographic Keys Generation Using FingerCodes / J. Ramírez-Ruiz, C. Pfeiffer, J. Nolazco-Flores // Advances in Artificial Intelligence – IBERAMIA-SBIA, 2006 (LNCS 4140). – 2006. – P. 178–187.

3. Hao, F. Ross Anderson, and John Daugman. Crypto with Biometrics Effectively / F. Hao, R. Anderson, J. Daugman // IEEE TRANSACTIONS ON COMPUTERS. – 2006. – Vol. 55, № 9.

4. Иванов, А. И. Нечеткие экстракторы: проблема использования в биометрии и криптографии / А. И. Иванов // Первая миля. – 2015. – № 1. – С. 40–47.

5. Нейросетевая защита персональных биометрических данных / Ю. К. Язов, В. И. Волчихин, А. И. Иванов, В. А. Фунтиков, И. Г. Назаров. – Москва : Радиотехника, 2012. – 157 с.

6. Технология использования больших нейронных сетей для преобразования нечетких биометрических данных в код ключа доступа : монография / Б. С. Ахметов, А. И. Иванов, В. А. Фунтиков, А. В. Безяев, Е. А. Малыгина. – Алматы : Изд-во LEM, 2014. – 144 с. – URL: http://portal.kaznu.kz/files/publicate/2014-06-27-11940.pdf

7. ГОСТ Р 52633.0–2006. Защита информации. Техника защиты информации. Требования к средствам высоконадежной биометрической аутентификации.

8. ГОСТ Р 52633.5–2011. Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа.

9. Безяев, А. В. Бескомпроматная индикация качества ввода фрагментов тайного составного биометрического образа / А. В. Безяев. // Нейрокомпьютеры: разработка, применение. – 2009. – № 6. – С. 59–62.

10. Безяев, А. В. Оптимизация структуры самокорректирующегося биокода, хранящего синдромы ошибок в виде фрагментов хеш-функций / А. В. Безяев, А. И. Иванов, Ю. В. Фунтикова // Безопасность в информационной сфере : вестник Уральского федерального округа. – 2014. – № 3 (13) . – С. 4–14.

11. Волчихин, В. И. Оценка эффекта ускорения вычислений, обусловленного поддержкой квантовой суперпозиции при корректировке выходных состояний нейросетевого преобразователя биометрии в код / В. И. Волчихин, А. И. Иванов, А. В. Безяев, А. В. Елфимов, А. П. Юнин // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2017. – № 1.– С. 43–55.

12. Нильсон, М. Квантовые вычисления и квантовая информация / М. Нильсон, И. Чанг. – Москва : Мир, 2006. – 821 с.

13. Душкин, Р. В. Квантовые вычисления и функциональное программирование / Р. В. Душкин // ДМК Пресс, 2015. – 234 с.

14. Волчихин, В. И. Нейросетевая молекула: решение обратной задачи биометрии через программную поддержку квантовой суперпозиции на выходах сети искусственных нейронов / В. И. Волчихин, А. И. Иванов // Вестник Мордовского университета. – 2017. – Т. 27, № 4. – С. 518–523.

15. Волчихин, В. И. Квантовая суперпозиция дискретного спектра состояний математической молекулы корреляции для малых выборок биометрических данных / В. И. Волчихин, А. И. Иванов, А. В. Сериков, Ю. И. Серикова // Вестник Мордовского университета. – 2017. – Т. 27, № 2. – С. 230–243.

16. Волчихин, В. И. Перспективы создания циклической континуально-квантовой хи-квадрат машины для проверки статистических гипотез на малых выборках биометрических данных и данных иной природы / В. И. Волчихин, А. И. Иванов, Д. В. Пащенко, Б. Б. Ахметов, С. Е. Вятчанин // Известия высших учебных заведений. Поволжский регион. Технические науки. – Пенза : Изд-во ПГУ, 2017. – № 1. – С. 5–15.

17. Волчихин, В. И. Хаотическая нейродинамика хи-квадрат молекул, квантовая обработка малых выборок биометрических данных на обычном

компьютере / В. И. Волчихин, А. И. Иванов // Информационно-управляющие и телекоммуникационные системы специального назначения : Всерос. науч.-техн. конф., посвящ. 100-летию со дня рождения одного из основоположников советской вычислительной техники Б. И. Рамеева. – Пенза, 2018.

Безяев, А. В. Нейросетевая молекула: механизм направленной квантовой коррекции большого числа ошибок длинного кода высокоразмерных биометрических образов / А. В Безяев // Безопасность информационных технологий : тр. I Всерос. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2019. – С. 102–111.

УДК 519.24; 53; 57.017

К. А. Перфилов, А. И. Газин

ОЦЕНКА СООТНОШЕНИЯ МОЩНОСТЕЙ ХИ-КВАДРАТ НЕЙРОНА И НЕЙРОНА СРЕДНЕГО ГЕОМЕТРИЧЕСКОГО ПРИ ИХ ИСПОЛЬЗОВАНИИ В ПРЕОБРАЗОВАТЕЛЯХ БИОМЕТРИЯ-КОД

Аннотация. Целью статьи является описание искусственных нейронов, построенных как аналоги статистического критерия квадрата среднего геометрического плотностей распределения значений многомерных биометрических данных «Свой» и многомерных плотностей распределения значений, предъявленных биометрических данных. Средствами имитационного моделирования на реальных биометрических данных показано, что мощность созданных квадратичных нейронов намного выше, чем мощность классических квадратичных радиально базисных нейронов. При этом важнейшее свойство линейной вычислительной сложности обучения квадратичных нейронов сохранено, что позволяет быстро обучать как угодно большие искусственные нейронные сети среднего геометрического на малых обучающих выборках.

K. A. Perfilov, A. I. Gazin

ESTIMATION OF THE POWER RATIO OF THE CHI-SQUARE OF A NEURON AND A NEURON OF THE AVERAGE GEOMETRIC WHEN USED IN BIOMETRICS-CODE CONVERTERS

Abstract. The purpose of the presented work is to describe artificial neurons constructed as analogs of the statistical criterion of the square of the geometric mean density of the distribution of the values of multidimensional biometric data “Own” and multidimensional densities of the distribution of values presented by the biometric data. Means of simulation on real biometric data showed that the power of the created quadratic neurons is much higher than the power of the classical quadratic radially basic neurons. At the same time, the most important property of the linear computational complexity of learning quadratic neurons is preserved, which makes it possible to quickly train arbitrarily large artificial neural networks of the geometric mean on small training samples.

Информационное общество предполагает активное использование Интернет ресурсов. Государственные и частные структуры

создают на своих сайтах личные кабинеты пользователей. К сожалению, существующая практика парольной защиты доступа к личным кабинетам обладает существенными уязвимостями. Пользователи не способны запоминать длинные случайные пароли. Владелец информационного ресурса не может быть уверен в том, что к личному электронному кабинету получил доступ именно его хозяин. Пароль может быть перехвачен программной закладкой, так же не составляет проблемы подменить IP адрес Интернет пользователя.

Для усиления защиты доступа к электронным кабинетам в настоящее время разрабатываются технологии биометрической аутентификации личности путем преобразования личных биометрических данных человека в его длинный случайный пароль доступа. Используются такие биометрические образы как: рисунок отпечатка пальца, рисунок радужной оболочки глаза, голосовой пароль, рукописный пароль, рисунок кровеносных сосудов глазного дна или ладони руки. Естественно, что преобразователи биометрия-код не могут быть идеальными и имеют вероятности ошибок первого и второго рода. Возникает необходимость тестирования ошибок первого и второго рода на реальных биометрических данных. Кроме того, при настройке «нечетких экстракторов» и при обучении нейросетевых преобразователей необходимо контролировать отсутствие в биометрических данных грубых ошибок. По сути дела, на небольшом числе примеров биометрического образа необходимо контролировать показатель близости распределения биометрических данных к многомерному нормальному закону. В 1900 году Пирсон предложил хи-квадрат статистический критерий [1], который на сегодняшний день практически стал стандартом [2]. Популярность хи-квадрат критерия Пирсона обусловлена тем, что для больших выборок в 400 и более опытов им была дана аналитическая зависимость плотности распределения значений от числа степеней свободы (от числа столбцов гистограммы экспериментальных данных).

Десятки других статистических критериев [3] на практике куда менее востребованы из-за того, что для них математиками построены таблицы доверительных вероятностей, но нет их точного аналитического описания.

Основная масса таблиц доверительных вероятностей для сотен известных на данный момент статистических критериев перенесены в справочники и стандартизованные рекомендации из первоисточников без независимой серьезной проверки инженерным сообществом. В инженерной среде не принято проверять таблицы доверительных вероятностей, приведенные в справочниках.

В итоге возникает путаница с достоверностью данных, публикуемых в современных статистических справочниках. В этом отношении источник [1] является одним из самых достоверных, так как в нем содержится очень большое число ссылок на первоисточники. По крайней мере, каждый сомневающийся инженер может проследить цепочку ссылок и попытаться найти сведения о независимом подтверждении достоверности таблиц доверительных вероятностей в том или ином первоисточнике.

Тяжесть проблемы состоит в том, что при статистическом анализе биометрических данных приходится настороженно относиться даже к проверенному вдоль и поперек хи-квадрат критерию. Причина состоит в том, что таблицы хи-квадрат критерия для выборки в 16–20 опытов не существует. Еще одной дополнительной проблемой является наличие большого числа предложенных математиками статистических критериев. Часть известных статистических критериев, построенных для интегральных характеристик – сравниваемых функций вероятности приведена в табл. 1. Очевидно, что интегральная функция вероятности – $P(u)$ через дифференциал связана с ее дифференциальным аналогом $p(u)$ – плотностью распределения функции вероятности. В силу линейности операций интегрирования и дифференцирования [4] во всех интегральных статистических критериях табл. 1 функцию вероятности – $P(u)$ можно заменить на ее дифференциал – $p(u)$. В итоге мы получим табл. 2 дифференциальных статистических критериев.

Подобная замена увеличивает число возможных для использования функционалов обогащения данных. Как показано на рис. 1, в ряде случаев дифференциальные функционалы имеют мощность существенно выше интегральных функционалов, если речь идет о разделении биометрических данных с нормальным законом распределения на фоне альтернативного равномерного закона распределения значений [5, 6].

Таблица 1

Статистические критерии проверки гипотезы о соответствии эмпирической функции вероятности $P(\tilde{u})$ некоторому ее аналитическому описанию $\tilde{P}(u)$

№	Название критерия и год создания	Формула вычисления критерия
1	Хи-квадрат критерий Пирсона, 1900 г. [1]	$= N\sum_{i=1}^{m}\left(n_i/N - \tilde{P}_i\right)^2/\tilde{P}_i$, где N – число опытов; m – число интервалов гистограммы; n_i – число отсчетов в i-м интервале; $\tilde{P}_i$ – теоретическая вероятность попадания в i-й интервал
2	Критерий Крамера-фон Мизеса, 1928 г. [1]	$= \int_{-\infty}^{+\infty}\left\{P(\tilde{u}) - \tilde{P}(u)\right\}^2 \cdot du$
3	Критерий Смирнова – Крамера-фон Мизеса, 1936 г. [1]	$= \int_{-\infty}^{+\infty}\left\{P(\tilde{u}) - \tilde{P}(u)\right\}^2 \cdot d\tilde{P}(u)$
4	Критерий Джини, 1941 г. [1]	$= \int_{-\infty}^{+\infty}\left\|P(\tilde{u}) - \tilde{P}(u)\right\| \cdot du)$
5	Критерий Андерсона – Дарлинга, 1952 г. [1]	$= \int_{-\infty}^{+\infty}\frac{\left\{P(\tilde{u}) - \tilde{P}(u)\right\}^2}{\tilde{P}(\tilde{u}) \cdot \left\{1 - \tilde{P}(u)\right\}} \cdot d\tilde{P}(u)$
6	Критерий Ватсона, 1961 г. [1]	$= \int_{-\infty}^{+\infty}\left\{\tilde{P}(u) - P(\tilde{u}) - \int_{-\infty}^{+\infty}\left[\tilde{P}(u) - P(\tilde{u})\right] \cdot d\tilde{P}(u)\right\}^2 \cdot d\tilde{P}(u)$
7	Критерий Фроцини, 1978 г. [1]	$= \int_{-\infty}^{+\infty}\left\|P(\tilde{u}) - \tilde{P}(u)\right\| \cdot d\tilde{P}(u)$
8	Критерий среднего геометрического, сравниваемых функций вероятности, 2014 г. [4]	$= \int_{-\infty}^{+\infty}\sqrt{P(\tilde{u}) \cdot (1 - \tilde{P}(u))} \cdot du$

Как видно из рис. 1, квадрат среднего геометрического сравниваемых функций распределения дает наибольшую мощность (dsg^2), обеспечивая минимальное значение равновероятных ошибок первого и второго рода на малых выборках. Видимо, это самый мощный на текущий момент статистический критерий [4] из известных критериев.

Таблица 2

Статистические критерии проверки гипотезы о соответствии наблюдаемой дифференциальной плотности вероятности $p(u)=\frac{dP(u)}{du}$ некоторому ее аналитическому описанию $\tilde{p}(u)$

№	Название критерия и год создания	Формула вычисления критерия
1	Дифференциальный вариант критерия Крамера-фон Мизеса [4] 2016 г.	$=\int_{-\infty}^{+\infty}\{p(u)-\tilde{p}(u)\}^2\cdot du$
2	Дифференциальный вариант критерия Смирнова-Крамера-фон Мизеса [4] 2016 г.	$=\int_{-\infty}^{+\infty}\{p(u)-\tilde{p}(u)\}^2\cdot\tilde{p}(u)\cdot du$
3	Дифференциальный вариант критерия Джини 2006 г. [5, 7, 8, 9] 2006 г.	$=\int_{-\infty}^{+\infty}\lvert p(u)-\tilde{p}(u)\rvert\cdot du$
4	Интегро-дифференциальный вариант критерия Андерсона-Дарлинга [4] 2016 г.	$=\int_{-\infty}^{+\infty}\frac{\{p(u)-\tilde{p}(u)\}^2}{\tilde{P}(u)\cdot\{1-\tilde{P}(u)\}}\cdot\tilde{p}(u)\cdot du;$
5	Дифференциальный вариант критерия Ватсона [4] 2016 г.	$=\int_{-\infty}^{+\infty}\left\{\tilde{p}(u)-p(u)-\int_{-\infty}^{+\infty}[\tilde{p}(u)-p(u)]\cdot\tilde{p}(u)\cdot du\right\}^2\cdot\tilde{p}(u)\cdot du$
6	Дифференциальный вариант критерия Фроцини [4] 2016 г.	$=\int_{-\infty}^{+\infty}\lvert p(u)-\tilde{p}(u)\rvert\cdot\tilde{p}(u)\cdot du$
7	Среднее геометрическое плотностей сравниваемых вероятностей 2016 г. [5, 6]	$=\int_{-\infty}^{+\infty}\sqrt{p(u)\cdot\tilde{p}(u)}\cdot du$
8	Квадрата среднего геометрического плотностей вероятности 2016 г. [5, 6]	$=\int_{-\infty}^{+\infty}p(u)\cdot\tilde{p}(u)\cdot du$

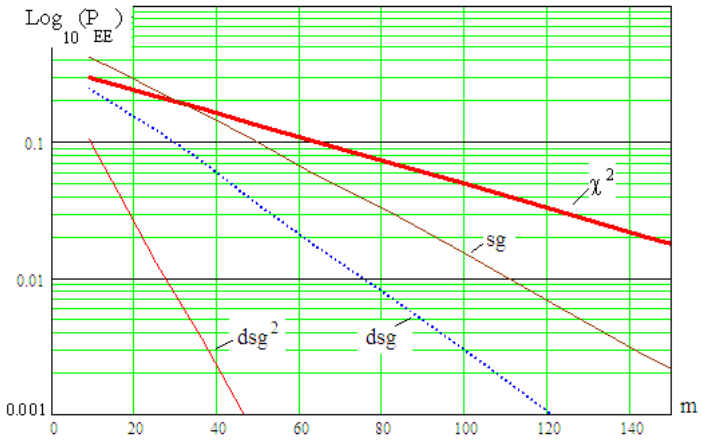


Рис. 1. Эталонная мощность хи-квадрат критерия (толстая линия) в логарифмической шкале равновероятных ошибок, sg – интегральный функционал среднего геометрического (табл. 1 строка 2), dsg –дифференциальный вариант функционала среднего геометрического (табл. 2, строка 8)

Из данных рис. 1 можно наблюдать, что равновероятные ошибки первого и второго рода $P_1 = P_2 = P_{EE}$ для хи-квадрат критерия достигают значения 0.01 при 160 опытах. Та же самая вероятность ошибок P_{EE} = 0.01 для критерия dsg^2 получается на выборке из 27 опытов. Наблюдается 6-ти кратное снижение требований к размеру тестовой выборки, что крайне существенно для биометрических приложений.

При практической реализации многомерного статистического анализа очень удобным оказалось применение искусственных нейронных сетей [10], обучаемых стандартным алгоритмом [11] с линейной вычислительной сложностью и тестируемых после обучения стандартными алгоритмами [12]. Быстрое и абсолютно устойчивое автоматическое обучение может быть организовано не только для сетей из персептронов, но и для иных нейронных сетей, воспроизводящих хорошо исследованные радиально-базисные функции [13] или множество иных, менее изученных, квадратичных функционалов [14–20].

Можно представить, что любому известному статистическому критерию (статистическому функционалу) можно поставить в соответствие некоторый нейрон [4]. Их отличие будет состоять только в том, что нейрон требует обучения (настройки) тогда как статистические критерии, как правило, не настраивают (не регулируют) в части предобработки данных. При использовании статистических критериев необходима настройка только порогового элемента (необходимо выбрать значение требуемого показателя доверительной вероятности).

Так же как все квадратичные функционалы нейрон среднего геометрического со структурой, изображенной на рис. 2, всегда имеет положительный отклик линейной части. Его настройка сводится к нормированию и центрированию m входных биометрических параметров по формуле:

$$u_i = \frac{E(v_i) - v_i}{\sigma(v_i)}, \qquad (1)$$

где i – упорядоченные номера входов нейрона; i = 1, 2,…, m, связанные с 416, контролируемыми биометрическими параметрами БиоОбраза, например, полученного в среде моделирования «БиоНейроАвтограф» [20], таблица связей формируется заранее с использованием генератора псевдослучайных чисел, как это рекомендует стандарт [11].

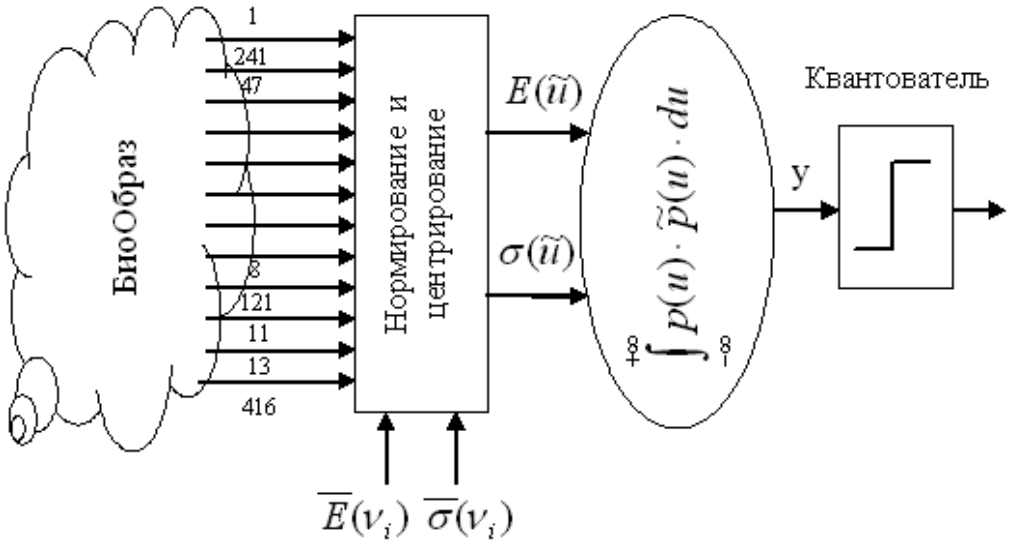


Рис. 2. Структурная схема искусственного нейрона, построенного как эквивалент квадрата среднего геометрического сравниваемых плотностей распределения значений

После нормирования и центрирования (1) для выборки m параметров по n примерам образа «Свой» вычисляют математическое ожидание – $\tilde{E}(u)$ и стандартное отклонение – $\tilde{\sigma}(u)$ для нормального теоретического распределения – $\tilde{p}(u)$.

Если после настройки нейрона dsg^2 подать на его входы тестовые примеры образа «Свой», не участвовавшие в его обучении, то на выходе линейной части получим отклики с малой дисперсией:

$$y = \int_{-\infty}^{+\infty} (\tilde{p}(u))^2 \cdot du = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left\{ \exp\left(\frac{-u^2}{2} \right) \right\}^2 du, \qquad (2)$$

где математическое ожидание $E(u) \approx 0$ – практически не является случайной величиной, стандартное отклонение $\sigma(u) \approx 1$, также практически не является случайной величиной.

Если же на входы обученного нейрона dsg^2 подавать биометрические данные образа «Чужой», то для них нормировка (1) работать не будет:

$$\tilde{u}_i = \frac{E(v_i) - \xi_i}{\sigma(v_i)}, \qquad (3)$$

где ξ_i – биометрический параметр образа «Чужой».

Как следствие, математическое ожидание $E(\tilde{u})$ оказывается случайно величиной, а стандартное отклонение $\sigma(\tilde{u})$ принимает большое значения в интервале от 2 до 5. В итоге отклик –

$\tilde{y}$ нейрона dsg^2 на воздействие вектором биометрических параметров образа «Чужой» – $\bar{\xi}$ будет описываться уравнением совершенно не похожим на уравнение (2):

$$\tilde{y} = \int_{-\infty}^{+\infty} p(u) \cdot \tilde{p}(u) \cdot du = \frac{1}{2\pi \cdot \sigma(\tilde{u})} \int_{-\infty}^{+\infty} \left\{ \exp\left(\frac{-u^2}{2} \right) \right\} \cdot \left\{ \exp\left(-\frac{(E(\tilde{u}) - u)^2}{2 \cdot (\sigma(\tilde{u}))^2} \right) \right\} du, \quad (4)$$

где математическое ожидание $E(\tilde{u})$ – случайная величина с нулевым математическим ожиданием $E(E(\tilde{u})) \approx 0.0$ и значительным стандартным отклонением $\sigma(E(\tilde{u})) \approx 1.41$, стандартное отклонение $\sigma(\tilde{u}) \approx 3.8$ самой переменной не случайно и имеет значительную величину.

Кардинальное отличие уравнений состоит в том, что они дают совершенно разные по своей природе отклики. Уравнение (2) является почти детерминированным, тогда как уравнение (3) дает случайную величину с большим стандартным отклонением – $\sigma(\tilde{y})$. Именно это обстоятельство и давало возможность добиваться высокого уровня подавления шумов квантования, возникающих на малых выборках при применении критерия среднего геометрического от двух сравниваемых плотностей распределения значений [6, 21]. Соотношения математических ожиданий распределения математических ожиданий образов «Свой», «Чужой» и их стандартных отклонений приведено на рис. 3 для нейронов dsg^2 с 8 входами (данные среды моделирования «БиоНейроАвтограф» [20]).

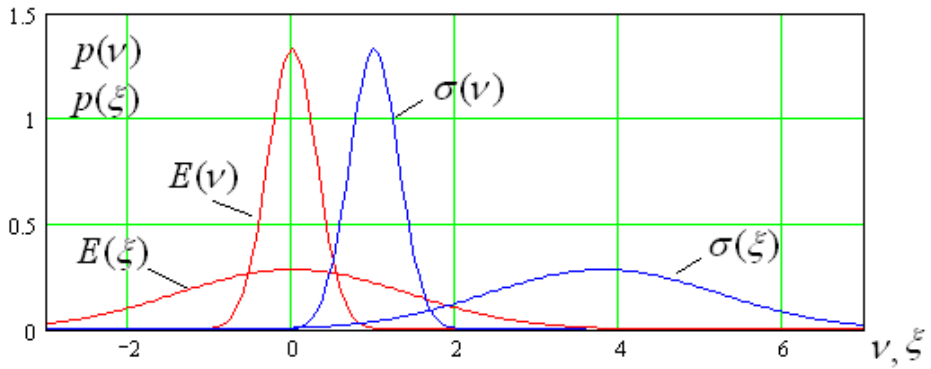


Рис. 3. Распределения математических ожиданий и стандартных отклонений для биометрических параметров образа «Свой» – ν и образа «Чужой» – ξ для нейрона dsg^2 с 8 входами

В силу того, что выражение (2) дает большое и почти детерминированное значение, а выражение (3) дает малое и случайное

значение примеры образа «Свой» и примеры образов «Чужой» оказываются хорошо различимы, если использовать нейрон dsg^2 с 8 входами.

Результат разделения (одинаковая вероятность ошибок первого и второго рода) оказывается намного лучше, чем для линейного нейрона и обычного квадратичного нейрона;

– линейный нейрон с 8 входами $P_1 = P_2 = P_{EE} = 0.45$;

– квадратичный нейрон, имеющий 8 входов, $P_{EE} = 0.26$;

– нейрон dsg^2 с 8 входами $P_1 = P_2 = P_{EE} = 0.21$.

С ростом размерности нейросетевого преобразования выигрыш от замены линейных и квадратичных нейронов на нейроны среднего геометрического квадрата плотностей распределения значений вероятностей усиливается.

Очевидно, что прямые вычисления вида (2), (3) реализовать на низкоразрядных процессорах с малой производительностью трудно. В связи с этим достаточно сложные функции преобразования (2) и (3) следует вычислить заранее и представить в виде двухмерной таблицы:

$$\tilde{Y} = -\log_2(\tilde{y}((E(\tilde{u}), \sigma(\tilde{u}))). \qquad (5)$$

Проблемы вычислений с использованием очень малых значений вероятностей решаются заранее во время вычисления двухмерных таблиц. Эта задача хорошо решается при использовании 64-разрядной сетки вычислительной машины под управлением какой-либо из сред для математических вычислений (например, MathLAB, MathCAD и др.). Сам же нейрон квадрата среднего геометрического может быть реализован программно на любом 8-ми разрядном процессоре низкой производительности.

Заключение

В данной работе впервые сделана попытка показать, что каждому из известных статистических функционалов может быть поставлен в соответствие некоторый нейрон. Особый интерес этот подход представляет при реализации многомерного статистического анализа биометрических данных. При технической реализации нейронов квадрата среднего геометрического проблемы работы с малыми значениями вероятностей легко разрешимы применением заранее вычисленных двухмерных таблиц логарифмов вероятностей. Это позволяет реализовывать нейроны квадрата среднего геометрического не только на обычных ПЭВМ, работающих под

ОС семейства Windows, но и на любом 8-ми битном процессоре малой производительности.

Библиографический список

1. Кобзарь, А. И Прикладная математическая статистика для инженеров и научных работников / А. И. Кобзарь. – Москва : ФИЗМАТЛИТ, 2006.– 816 с.

2. ГОСТ Р 50.1.037–2002. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Критерии типа хи-квадрат. – Москва, 2001. – Ч. 1. – 140 с.

3. ГОСТ Р 50.1.037–2002. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Непараметрические критерии. – Москва, 2002. – Ч. 2. – 123 с.

4. Иванов, А. И. Многомерная нейросетевая обработка биометрических данных с программным воспроизведением эффектов квантовой суперпозиции : монография / А. И. Иванов. – Пенза : Изд-во ПНИЭИ, 2016. –133 с.

5. Волчихин, В. И. Эффект снижения размера тестовой выборки за счет перехода к многомерному статистическому анализу биометрических данных / В. И. Волчихин, А. И. Иванов, Н. И. Серикова и др. // Известия высших учебных заведений. Поволжский регион. Технические науки. – Пенза : Изд-во ПГУ, 2015. – №1.– С. 50–59.

6. Иванов, А. И. Оценка соотношения мощностей семейства статистических критериев «среднего геометрического» на малых выборках биометрических данных / А. И. Иванов, К. А. Перфилов // Современные охранные технологии и средства обеспечения комплексной безопасности объектов : XI Всерос. науч.-практ. конф. – Пенза, Заречный, 2016. – С. 223–229.

7. Малыгин, А. Ю. Быстрые алгоритмы тестирования нейросетевых механизмов биометрико-криптографической защиты информации / А. Ю. Малыгин, В. И. Волчихин, А. И. Иванов и др. – Пенза : Изд-во ПГУ, 2006. –161 с.

8. Серикова, Н. И. Оценка правдоподобия гипотезы о нормальном распределении по критерию Джини для числа степеней свободы, кратного числу опытов / Н. И. Серикова, А. И. Иванов, Ю. И. Серикова // Вопросы радиоэлектроники. – 2015. – № 1 (1). – С. 85–94.

9. Серикова, Н. И. Оценка правдоподобия гипотезы о нормальном распределении по критерию Джини для сглаженных гистограмм, построенных на малых тестовых выборках / Н. И. Серикова // Вопросы радиоэлектроники. – 2015. – № 1. – С. 85–94.

10. ГОСТ Р 52633.0–2006. Защита информации. Техника защиты информации. Требования к средствам высоконадежной биометрической аутентификации. – Москва : Стандартинформ, 2007.– 27 с.

11. ГОСТ Р 52633.5–2011. Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа. – Москва : Стандартинформ, 2012. –16 с.

12. ГОСТ Р 52633.3–2011. Защита информации. Техника защиты информации. Тестирование стойкости средств высоконадежной биометрической защиты к атакам подбора. – Москва : Стандартинформ, 2012. –16 с.

13. Саймон, Х. Нейронные сети: полный курс / Х. Саймон. – Москва : Вильямс, 2006. –1104 с.

14. Ахметов, Б. Б. Многомерный статистический анализ биометрических данных сетью частных критериев Пирсона / Б. Б. Ахметов, А. И. Иванов, А. В. Безяев и др. // Вестник Национальной академии наук. – Алматы, 2015. – № 1. – С. 5– 11.

15. Иванов, А. И. Подавление шумов квантования биометрических данных при использовании многомерного критерия Крамера – фон Мизеса / А. И. Иванов, А. И. Газин, С. Е. Вятчанин и др. // Проблемы информационной безопасности. Компьютерные системы. – Санкт Петербург : Изд-во ПТУ.– 2016. – № 2. – С. 21–28.

16. Ахметов, Б. Многомерные статистики существенно зависимых биометрических данных, порождаемые нейросетевыми эмуляторами квадратичных форм : монография / Б. Ахметов, А. Иванов. – Алматы : LEM, 2016. – 86 с.

17. Иванов, А. И. Снижение требований к размеру тестовой выборки биометрических данных при переходе к использованию многомерных корреляционных функционалов Байеса / А. И. Иванов, П. С. Ложников, А. Е. Сулавко и др. // Инфокоммуникационные технологии. – 2017. – № 15 (2). – С. 186–193.

18. Иванов, А. И. Идентификация подлинности рукописных автографов сетями Байеса – Хэмминга и сетями квадратичных форм / А. И. Иванов, П. С. Ложников, Е. И. Качайкин // Вопросы защиты информации. – 2015. – № 2. – С. 28–34.

19. Иванов, А. И. Биометрическая идентификация рукописных образов с использованием корреляционного аналога правила Байеса / А. И. Иванов, П. С. Ложников, Е. И. Качайкин и др. // Вопросы защиты информации. – 2015. – № 3. – С. 48–54.

20. Иванов, А. И. Среда моделирования «БиоНейроАвтограф» [Программный продукт создан лабораторией биометрических и нейросетевых технологий] / А. И. Иванов, О. С. Захаров. – URL: http://пниэи.рф/activity/science/noc.htm (дата обращения: 10.08.17).

21. Иванов, А. И. Оценка качества малых выборок биометрических данных с использованием дифференциального варианта статистического критерия среднего геометрического / А. И. Иванов, К. А. Перфилов, Е. А. Малыгина // Вестник СИБГАУ. – 2016. – № 4 (17). – С. 864–871.

Перфилов, К. А. Оценка соотношения мощностей хи-квадрат нейрона и нейрона среднего геометрического при их использовании в преобразователях биометрия-код / К. А. Перфилов, А. И. Газин. // Безопасность информационных технологий : тр. I Всерос. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2019. – С. 112–122.

УДК 519.2, 612.087, 621.319.7

А. В. Сериков, С. В. Качалин

КОРРЕЛЯЦИОННАЯ МОЛЕКУЛА С ЭЛЛИПТИЧЕСКИМИ КВАНТОВАТЕЛЯМИ ДЛЯ ВЫЧИСЛЕНИЙ НА МАЛЫХ ОБУЧАЮЩИХ ВЫБОРКАХ

Аннотация. Целью статьи является повышение корректности вычисления коэффициентов корреляции. Показано, что новый метод вычисления корреляции дает погрешность, слабо коррелированную с погрешностью вычисления корреляции классическим методом. То есть новые данные могут быть использованы для корректировки погрешности классических вычислений.

A. V. Serikov, S. V. Kachalin

CORRELATION MOLECULE WITH ELLIPTIC QUANTIZERS FOR COMPUTATIONS ON SMALL TRAINING SAMPLES

Abstract. The aim of the work is to increase the correctness of the calculation of the correlation coefficients. It is shown that the new method for calculating the correlation gives an error that is weakly correlated with the error in calculating the correlation by the classical method. That is, new data can be used to correct the error of classical calculations.

Проблема учета коэффициентов корреляции при обучении нейросетевых преобразователей биометрия-код

В настоящее время Россия создает собственную цифровую экономику. Одной из проблем цифровой экономики является создание эффективных механизмов учета цифровых затрат и цифровых ресурсов. Подобные механизмы учета могут быть реализованы, опираясь на возможность массового применения криптографии. При этом возникает серьезное технологическое ограничение, обусловленное тем, что обычный человек не может запомнить множество своих личных криптографических ключей.

Для того, чтобы снять это ограничение в США и странах НАТО создаются так называемые «нечеткие экстракторы» [1–3], преобразующие биометрию человека в код пароля доступа. Проблемой нечетких экстракторов является то, что они применяют

классические коды с высокой избыточностью для обнаружения и исправления ошибок [4]. При 30-кратной избыточность длина выходного кода «нечеткого экстрактора» становится в 30 раз меньше, чем число контролируемых биометрических параметров человека. По этой причине для рукописных образов, голосовых образов, рисунков отпечатка пальцев, образов лица человека «нечеткие экстракторы» дают коды длинной до 20 бит. Это не позволяет объединять «нечеткие экстракторы» с сильной криптографией длинных ключей.

Россия идет иным путем, создавая и стандартизируя нейросетевые преобразователи биометрия-код [5–7]. Один нейрон в таких преобразователях создает один бит ключа, то есть их выходной код может иметь любую длину [8, 9] из-за того, что в преобразователях можно использовать много искусственных нейронов. Обучение нейросетевых преобразователей биометрия-код должно быть полностью автоматически и иметь низкую вычислительную сложность. В частности, стандартизованный в России алгоритм обучения [7] имеет линейную вычислительную сложность за счет того, что он не является итерационным и вычисляет весовые коэффициенты нейронов как простую функцию младших статистических моментов двух переменных:

$$|\mu_i| = f(E(\nu_i), E(\xi_i), \sigma(\nu_i), \sigma(\xi_i)), \qquad (1)$$

где ν_i – биометрический параметр образа «Свой», ξ_i – биометрический параметр образа «Чужой», $E(.)$ – оператор вычисления математического ожидания, $\sigma(.)$ – оператор вычисления стандартного отклонения.

Ожидается, что следующее поколение стандартизованных нейросетевых преобразователей биометрия-код будет автоматически обучаться алгоритмом, имеющим квадратичную вычислительную сложность за счет дополнительного учета одинаковых корреляционных связей [10–12] биометрических данных одного нейрона:

$$|\mu_i| = f(E(\nu_i), E(\xi_i), \sigma(\nu_i), \sigma(\xi_i), r_i(\nu_i, \nu_j)), \qquad (2)$$

где $r_i(\nu_i, \nu_j) = r_i(\nu_i, \nu_{j+1}) = r_i(\nu_i, \nu_{j+2}) = ...$ одинаково коррелированные биометрические данные одного нейрона, полученные специальной процедурой симметризации корреляционных связей.

Проблема определения корреляционных связей при использовании малых тестовых выборок

К сожалению, классическая процедура вычисления коэффициентов корреляции по формуле Пирсона неустойчива:

$$r(\nu_i, \nu_j) = \frac{1}{n} \sum_{k=1}^{n} \frac{\left(E(\nu_i) - \nu_{i,k}\right) \cdot \left(E(\nu_j) - \nu_{j,k}\right)}{\sigma(\nu_i) \cdot \sigma(\nu_j)}. \qquad (3).$$

На малых выборках ошибки вычисления коэффициентов корреляции оказываются велики. Эта ситуация отображена на рис. 1.

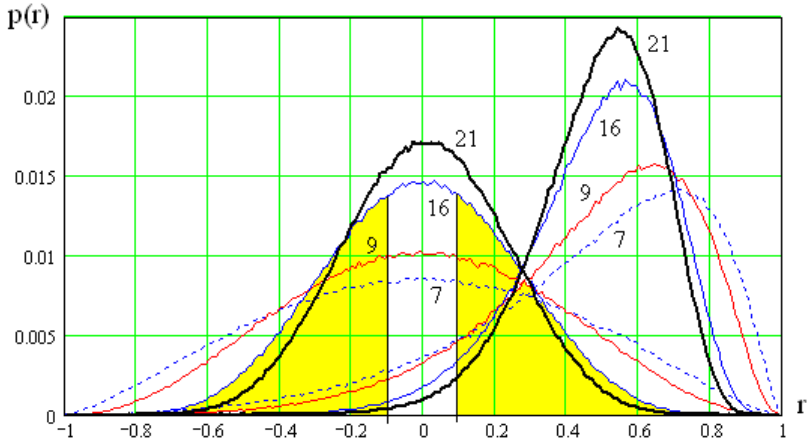


Рис. 1. Распределения значений коэффициентов корреляции при разных тестовых выборках в 7, 9, 16, 21-м примерах

Из рис. 1 видно, что самая большая погрешность возникает при определении коэффициентов корреляции независимых данных. Так для выборок из 16 независимых опытов вычисленный коэффициент корреляции попадает в интервал от –-0.1 до +0.1 с вероятностью только 0.2. Становится актуальной задача повышения точности вычисления коэффициентов корреляции на малых выборках. Причина плохой обусловленности вычисления коэффициентов корреляции состоит в том, что накапливаются четыре ошибки вычисления статистических моментов, входящих в формулу Пирсона (3):

$$\Delta r(\nu_i, \nu_j) = f\left(\Delta E(\nu_i), \Delta E(\nu_j), \Delta\sigma(\nu_i), \Delta\sigma(\nu_j)\right). \qquad (4)$$

Функция (4) обычно монотонна. В связи с этим ошибку вычисления коэффициентов корреляции удается снизить, уменьшая значения ошибок статистических моментов двух контролируемых биометрических параметров [13, 14]. Еще одним способом повы-

шения точности вычислений является синтез новых процедур оценки [15, 16] коэффициентов корреляции и усреднения их данных с данными классической формулы Пирсона (3).

Синтез новой процедуры оценки коэффициента корреляционной связи

При обработке данных реальных стрельб военными [17] практиковалось описание нормальных распределений эллипсами, как это показано на рис. 2.

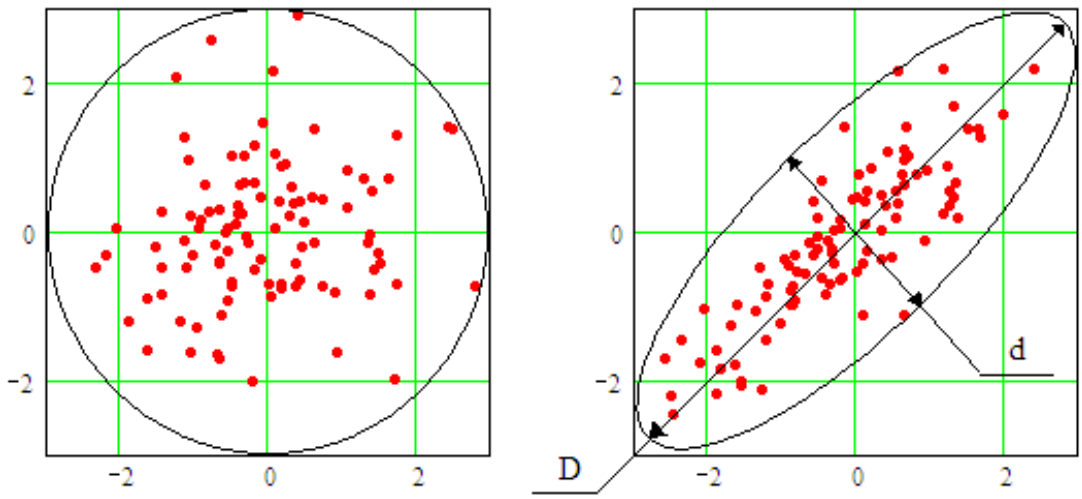


Рис. 2. Приближение двухмерного нормального распределения данных эллипсами

На рис. 2 приведены 2 приближения нормальных распределений, полученные ручным способом. После подбора параметров эллипсов в ручном режиме оценивают соотношение их большого и малого диаметров:

$$r \approx 1 - \frac{d}{D}. \qquad (5)$$

Расчет коэффициентов корреляции по формуле (5) дает значение корреляции близкое к нулю для распределения в левой части рис. 2. Для распределения данных в правой части рисунка корреляция составит r = 1/3.

Следует отметить, что применение эллипсов для описания распределений эквивалентно использованию симметричных эллиптических квантователей:

$$\begin{cases} y_i^{\,2} = \begin{bmatrix} E(\nu_1) - \nu_{1,i} \\ E(\nu_2) - \nu_{2,i} \end{bmatrix}^{T} \cdot \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} E(\nu_1) - \nu_{1,i} \\ E(\nu_2) - \nu_{2,i} \end{bmatrix} \\ z(y_i^2) = "0" \quad \text{если} \quad y_i^2 \le k \\ z(y_i^2) = "1" \quad \text{если} \quad y_i^2 > k \end{cases}, \tag{6}$$

где k – порог квантователя.

Если использовать два квантователя с параметрами r = 1/3 и r = –1/3, мы получим корреляционную молекулу, работа которой иллюстрируется рис. 3.

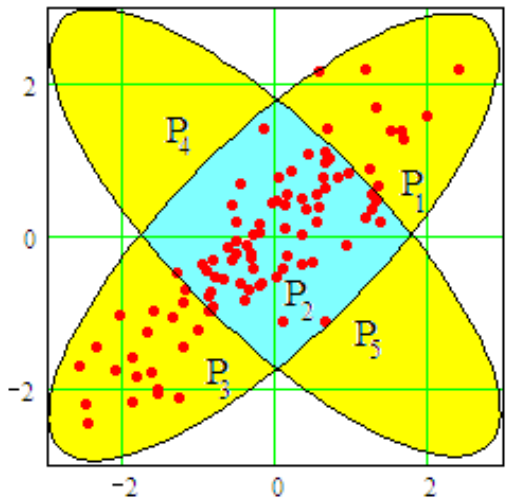


Рис. 3. Работа двух эллиптических квантователей корреляционной молекулы

Работа молекулы построена на том, что число опытов попавших внутрь эллипсов первого и второго квантователя разное:

$$P_1 + P_2 + P_3 > P_2 + P_4 + P_5 . \tag{7}$$

Если исследуемое распределение положительно коррелированно, то число опытов обнаруженное внутри первого квантователя должно быть больше чем число опытов обнаруженное внутри второго квантователя:

$$N_1 > N_2 . \tag{8}$$

Если корреляция отрицательна, то соотношение (8) меняется на противоположное.

Учитывая это, мы получим следующие соотношения для вычисления коэффициентов корреляции:

$$\begin{cases} \tilde{r} \approx 1 - \dfrac{N_1 - N_2}{N_1} \quad \text{при } N_1 > N_2 \\ \tilde{r} \approx -\left(1 - \dfrac{N_2 - N_1}{N_2}\right) \quad \text{при } N_2 > N_1 \end{cases}. \qquad (9)$$

Очевидно, что вычисления коэффициента корреляции по формуле (3) и по формулам (9) являются разными вычислительными процедурами. Как следствие мы получаем разные значения ошибок Δr и $\Delta\tilde{r}$. Значения этих ошибок положительно коррелированы $r(\Delta r, \Delta\tilde{r}) = 0.519$. То, что их корреляция не единична является предпосылкой для использования вычислений (9) совместно с результатами, полученными по классической формуле Пирсона (3).

Еще одним важным результатом является то, что спектры состояний корреляционной молекулы с линейными квантователями [14] и корреляционной молекулы эллиптическими квантователями (6) имеют разную структуру. Так для ситуации, отображенной на рис. 3 квантовые состояния для двух молекул различаются даже по длине векторов дискретных состояний. Для молекулы с двумя линейными квантователями спектр описывается вектором из 4-х параметров $\{P_1 = 31, P_2 = 9, P_3 = 52, P_4 = 8\}$. Для молекулы с двумя эллиптическими квантователями получается вектор из 5 параметров $\{P_1 = 11, P_2 = 67, P_3 = 21, P_4 = 0, P_5 = 0\}$. Если корреляция менее 0.333 два последних компонента вектора становятся не нулевыми.

Библиографический список

1. Monrose, F. Cryptographic key generation from voice / F. Monrose, M. Reiter, Q. Li, S. Wetzel // Proc. IEEE Symp. on Security and Privacy, 2001. – P. 202–213. – URL: https://www.cs.unc.edu/~reiter/papers/2001/SP2.pdf

2. Ramírez-Ruiz, J. Cryptographic Keys Generation Using FingerCodes / J. Ramírez-Ruiz, C. Pfeiffer, J. Nolazco-Flores // Advances in Artificial Intelligence – IBERAMIA-SBIA 2006 (LNCS 4140). – 2006. – P. 178–187. – URL: http://dl.acm.org/citation.cfm?id=2110882

3. Hao, F. Crypto with Biometrics Effectively / F. Hao, R. Anderson, J. Daugman // IEEE TRANSACTIONS ON COMPUTERS, 2006. – Vol. 55, № 9. – P. 1073–1074.

4. Иванов, А. И. Нечеткие экстракторы: проблема использования в биометрии и криптографии / А. И. Иванов // Первая миля. – 2015. – № 1. – С. 40–47. – URL: http://www.lastmile.su/journal/article/4489

5. ГОСТ Р 52633.0–2006. Защита информации. Техника защиты информации. Требования к средствам высоконадежной биометрической аутентификации.

6. ГОСТ Р 52633.3–2011. Защита информации. Техника защиты информации. Тестирование стойкости средств высоконадежной биометрической защиты к атакам подбора.

7. ГОСТ Р 52633.5–2011. Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа.

8. Ахметов, Б. С. Технология использования больших нейронных сетей для преобразования нечетких биометрических данных в код ключа доступа : монография / Б. С. Ахметов, А. И. Иванов, В. А. Фунтиков, А. В. Безяев, Е. А. Малыгина. – Алматы : ТОО «Издательство LEM», 2014. – 144 с. – URL: http://portal.kazntu.kz/files/publicate/2014-06-27-11940.pdf

9. Нейросетевая защита персональных биометрических данных / Ю. К. Язов, В. И. Волчихин, А. И. Иванов, В. А. Фунтиков, И. Г. Назаров. – Москва : Радиотехника, 2012. – 157 с.

10. Ложников, П. С. Биометрическая защита гибридного документооборота / П. С. Ложников. – Новосибирск : Из-во СО РАН, 2017. – 130 с.

11. Иванов, А. И. Идентификация подлинности рукописных автографов сетями Байеса – Хэмминга и сетями квадратичных форм / А. И. Иванов, П. С. Ложников, Е. И. Качайкин // Вопросы защиты информации. – 2015. – № 2. – С. 28–34.

12. Иванов, А. И. Биометрическая идентификация рукописных образов с использованием корреляционного аналога правила Байеса / А. И. Иванов, П. С. Ложников, Е. И. Качайкин, А. Е. Сулавко // Вопросы защиты информации. – 2015. – № 3. – С. 48–54.

13. Кулагин, В. П. Корректировка методических и случайных составляющих погрешностей вычисления коэффициентов корреляции, возникающих на малых выборках биометрических данных / В. П. Кулагин, А. И. Иванов, Ю. И. Серикова // Информационные технологии. – 2016. – Т. 22, № 9. – С. 705–710. – URL: http://novtex.ru/IT/it2016/number09.html

14. Иванов, А. И. Номограммы оценки погрешности, коэффициентов корреляции, вычисленных на малых выборках биометрических данных / А. И. Иванов, Ю. И. Серикова // Вопросы радиоэлектроники. – 2015. – № 2. – С. 123–130.

15. Волчихин, В. И. Фрактально-корреляционный функционал, используемый при поиске пар слабо зависимых биометрических данных в малых выборках / В. И. Волчихин, А. И. Иванов, Б. Б. Ахметов, Ю. И. Серикова // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2016. – № 4. – С. 25–31. – URL: http://izvuz_tn.pnzgu.ru/tn3416

16. Волчихин, В. И. Квантовая суперпозиция дискретного спектра состояний математической молекулы корреляции для малых выборок биометрических данных / В. И. Волчихин, А. И. Иванов, А. В. Сериков, Ю. И. Серикова // Вестник Мордовского университета. – 2017. – Т. 27, № 2. – С. 230–243.

17. Абезгауз, Г. Г. Справочник по вероятностным расчетам / Г. Г. Абезгауз, А. П. Тронь, Ю. Н. Копенкин, И. А. Коровина. – Москва : Воениздат, 1970. – 536 с.

Сериков, А. В. Корреляционная молекула с эллиптическими квантователями для вычислений на малых обучающих выборках / А. В. Сериков, С. В. Качалин // Безопасность информационных технологий : тр. I Всерос. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2019. – С. 123–129.

УДК 519.2, 612.087, 621.319.7

А. В. Майоров, С. А. Сомкин, А. П. Юнин, А. Ж. Акмаев

ОЦЕНКА СТОЙКОСТИ ЗАЩИЩЕННЫХ НЕЙРОСЕТЕВЫХ ПРЕОБРАЗОВАТЕЛЕЙ БИОМЕТРИЯ-КОД С ИСПОЛЬЗОВАНИЕМ БОЛЬШИХ БАЗ СИНТЕТИЧЕСКИХ БИОМЕТРИЧЕСКИХ ОБРАЗОВ

Аннотация. Целью работы является тестирование стойкости к атакам подбора нейросетевых контейнеров, в которых таблицы связей нейронов и таблицы их весовых коэффициентов защищены гаммированием. Используется база естественных биометрических образов «Чужой». Проверяется, сколько первых бит ключа подбирается на базе естественных биометрических образов. После этого размер базы увеличивают путем скрещивания образов-родителей и получения образов-потомков алгоритмом ГОСТ Р 52633.2. Контролируются размеры монотонно увеличивающейся тестовой базы и число подобранных на ней бит ключа. В логарифмической шкале размер тестовой базы связан линейно с длиной ключа, подбираемого при численном эксперименте. Пользуясь этим свойством, удается прогнозировать время, которое займет подбор ключа длиной 256 или 512 бит по данным реального подбора длины ключа от 40 до 64 бит.

A. V. Mayorov, S. A. Somkin, A. P. Junin, A. Zh. Akmaev

EVALUATION OF THE RESISTANCE OF PROTECTED NEURAL NETWORK CONVERTERS BIOMETRICS-CODE USING LARGE BASES OF SYNTHETIC BIOMETRIC IMAGES

Abstract. The aim of the work is to test the resistance to attacks of the selection of neural network containers in which the neuron link tables and the tables of their weights are protected by gamming. Used base of natural biometric images "Alien". Further, it is checked how many first bits of the key are selected on the basis of natural biometric images. After that, the size of the base is increased by crossing the parent images and obtaining the descendant images using the algorithm of GOST R 52633.2. They control the size of the monotonously increasing test base and the number of key bits selected on it. In a logarithmic scale, the size of the test base is linearly related to the key length selected in a numerical experiment. Using this property, it is possible to predict the time that will take the selection of a key length of 256 or 512 bits according to the real selection of the key length from 40 to 64 bits.

Совместное использование биометрии и криптографии

Переход к цифровой экономике приводит к необходимости создания механизмов повышения доверия к электронным документам, находящимся в открытом информационном пространстве (например, хранящимся на облачных сервисах).

Традиционные методы криптографической защиты информации (шифрование, формирование цифровой подписи) обладают низкой эргономичностью. Обычный пользователь не может запомнить длинный пароль из случайных знаков или свой криптографический ключ. В связи с этим в России [1, 2] и за рубежом [3–5] активно развиваются методы преобразования личной биометрии человека в его криптографический ключ.

На рис. 1 приведена схема, так называемых, «нечетких экстракторов» и схема нейросетевых преобразователей биометрия-код. Из рисунка видно, что выходной код криптографического ключа «нечетких экстракторов» короткий. Это происходит из-за того, что «нечеткие экстракторы» используют классические коды с высокой избыточностью, способные обнаруживать и корректировать ошибки.

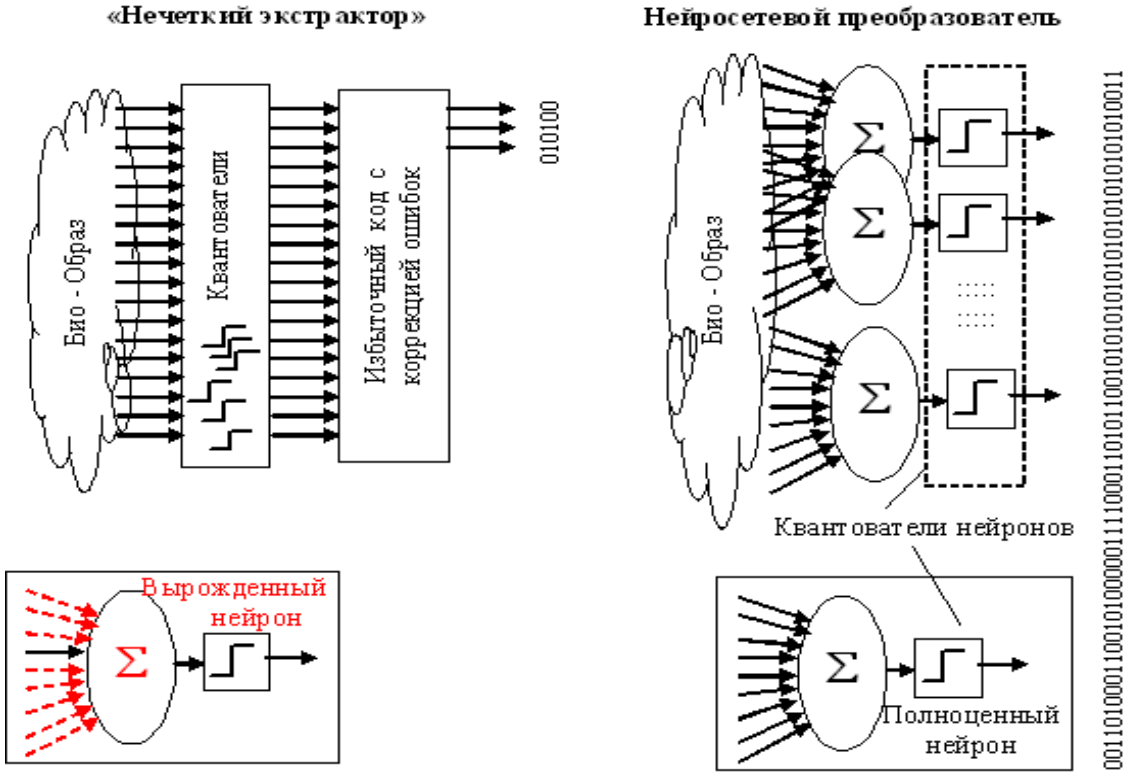


Рис. 1. Схемы организации «нечетких экстракторов» и нейросетевых преобразователей биометрия-код

Так, Даугман [5] при корректировке кода, получаемого из рисунка радужной оболочки глаза, использует код с 20-кратной избыточностью. То есть при расчете длины выходного ключа число входных биометрических параметров следует уменьшить в 20 раз для «нечетких экстракторов».

Как следствие, «нечеткие экстракторы» нельзя применять совместно с сильными криптографическими схемами, ориентированными на использование длинных криптографических ключей.

Ситуация меняется, когда используются нейросетевые преобразователи биометрия-код [1, 2]. В таких конструкциях за каждый бит ключа отвечает один нейрон, число нейронов может быть любым. То есть отечественные нейросетевые преобразователи биометрия-код могут работать с любыми криптографическими схемами защиты информации.

Классические коды с обнаружением и исправлением ошибок хорошо исследованы, по этой причине проектирование «нечетких экстракторов» является простой инженерной задачей. Иначе обстоит дело с нейросетевыми преобразователями биометрия-код. Искусственные нейронные сети плохо учатся, кроме того до конца неизвестно, как их защищать, какова стойкость биометрико-нейросетевой защиты информации.

Для решения задачи обучения искусственных нейронных сетей в России разработан и введен в действие национальный стандарт ГОСТ Р 52633.5 [6]. Для защиты данных обученных нейронных сетей ТК 026 «Криптографическая защита» в настоящее время создает техническую спецификацию [7].

После обучения каждый нейрон преобразователя биометрия-код имеет таблицу связей с биометрическими параметрами и таблицу весовых коэффициентов, как это показано на рис. 2. Известно техническое решение по патенту RU 2346397 [8], по которому осуществляют шифрование таблиц связей и таблиц весовых коэффициентов. Это техническое решение является слишком сложным, так как ориентировано на применения криптографических алгоритмов шифрования, например, выполненных по ГОСТ Р 34.12 [9]. Практическая реализация средств криптографической защиты информации упрощается, если в место полноценного шифрования используется криптографическая хэш-функция, например, выполненная по ГОСТ Р 34.11 [10].

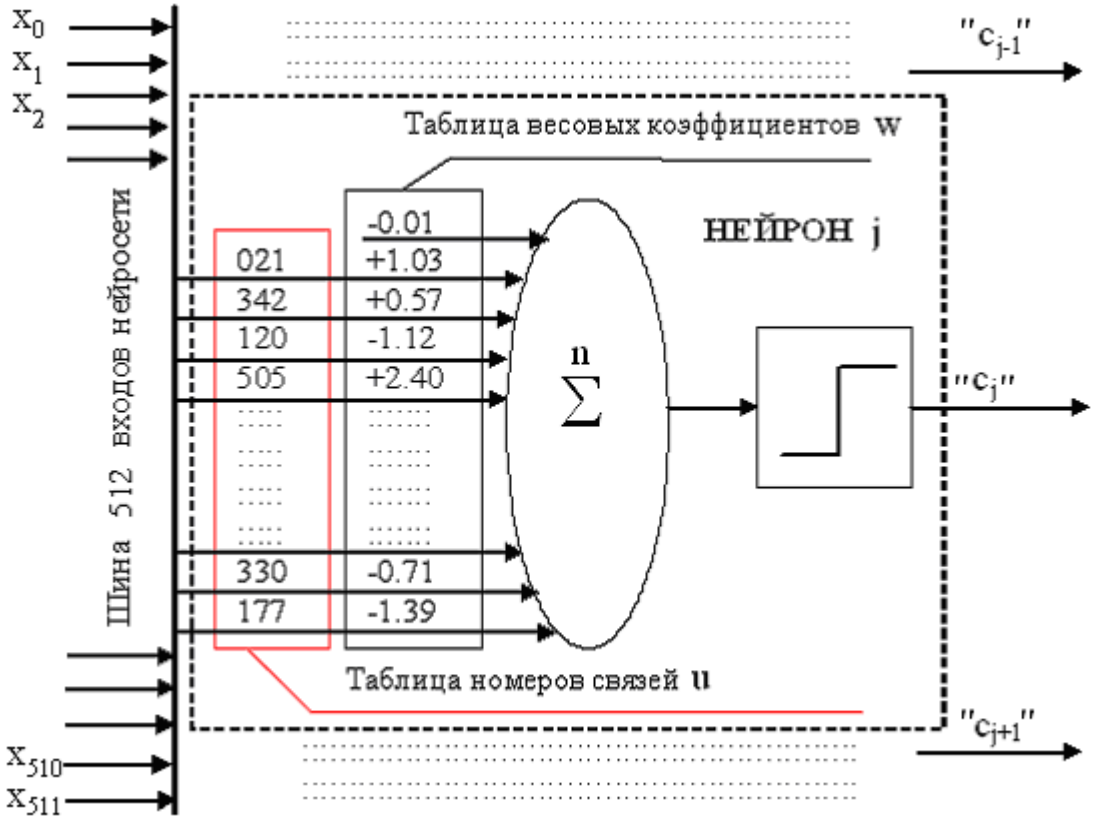


Рис. 2. Каждый нейрон преобразователя биометрия-код имеет собственную таблицу входных связей и собственную таблицу весовых коэффициентов

Техническая спецификация [7] строится на том, что таблицы первого нейрона преобразователя биометрия-код накрыты гаммой – Γ_1, которая получается хэшированием пароля доступа:

$$\Gamma_1 = hesh("пароль"). \quad (1)$$

Таблицы связей второго нейрона накрываются гаммой, полученной хэшированием таблицы связей гаммы первого нейрона в конкатенации с выходным состоянием первого нейрона:

$$\Gamma_2 = hesh("\Gamma_1","c_1"). \quad (2)$$

Получается рекурсивная схема, на каждом шаге которой следующая гамма, зависит от предыдущей гаммы и состояния предшествующего нейрона:

$$\Gamma_k = hesh("\Gamma_{k-1}","c_{k-1}"). \quad (3)$$

Таблицы связей и значения разрядов кода на выходах нейронов известны при обучении нейронной сети. По этой причине после обучения и назначения пароля доступа все гаммы для всех таблиц нейронов могут быть вычислены. После того как таблицы

связей накрыты гаммами, получается защищенный нейросетевой контейнер, который может храниться с привлечением облачного сервиса.

При использовании защищенного контейнера пользователь набирает свой пароль доступа и предъявляет свой биометрический образ. Биометрический образ должен порождать на выходах нейросетевого преобразователя верный выходной код $\{"c_1","c_2","c_3",...,"c_{256}"\}$. Появляется возможность повторно рекуррентно (3) восстановить все гаммы, которыми ранее были накрыты таблицы связей нейронов. Если возникает хотя бы одна ошибка в коде пароля доступа или в коде выходных состояний нейронов, восстановить все гаммы нельзя. Возникает эффект размножения ошибок.

Известно, что стойкость кода биометрико-нейросетевого преобразования к атакам подбора много ниже, чем стойкость криптографического ключа такой же длины [11]. Более того, каждый биометрический преобразователь, обученный на свой биометрический образ, будет обладать своей стойкостью к атакам подбора. То есть, после каждой процедуры обучения и каждой процедуры защиты данных таблиц нейронов необходимо выполнить процедуру тестирования.

Для тестирования следует заранее собрать базу образов «Чужой», например, состоящую из 10 000 образов. Далее пользователь должен ввести свой пароль и начать подставлять на входы защищенного преобразователя данные 10 000 образов «Чужой». При этом, на выходе преобразователя появится 10 000 кодов. О стойкости защищенного нейросетевого преобразователя следует судить, анализируя младшие разряды кодов.

Проведенный численный эксперимент показал, что для рукописного образа «slovo_01» база данных из 10 000 образов дает множество совпадение до 40 младших разрядов кода. Можно утверждать, что число попыток подбора, позволяющих найти часть ключа в 40 бит, составляет 10 000 образов «Чужой». Формальная запись этого утверждения:

$$П_{40} = 10\,000. \qquad (4)$$

Ограниченный размер тестовой базы обусловлен высоким уровнем трудоемкости ее формирования, если придерживаться требований ГОСТ Р 52633.1 [12]. Для того, чтобы увеличить размер тестовой базы необходимо скрещивать пары образов-

родителей по ГОСТ Р 52633.2 [13], получая образы-потомки. Примеры подобного скрещивания иллюстрируются рис. 3.

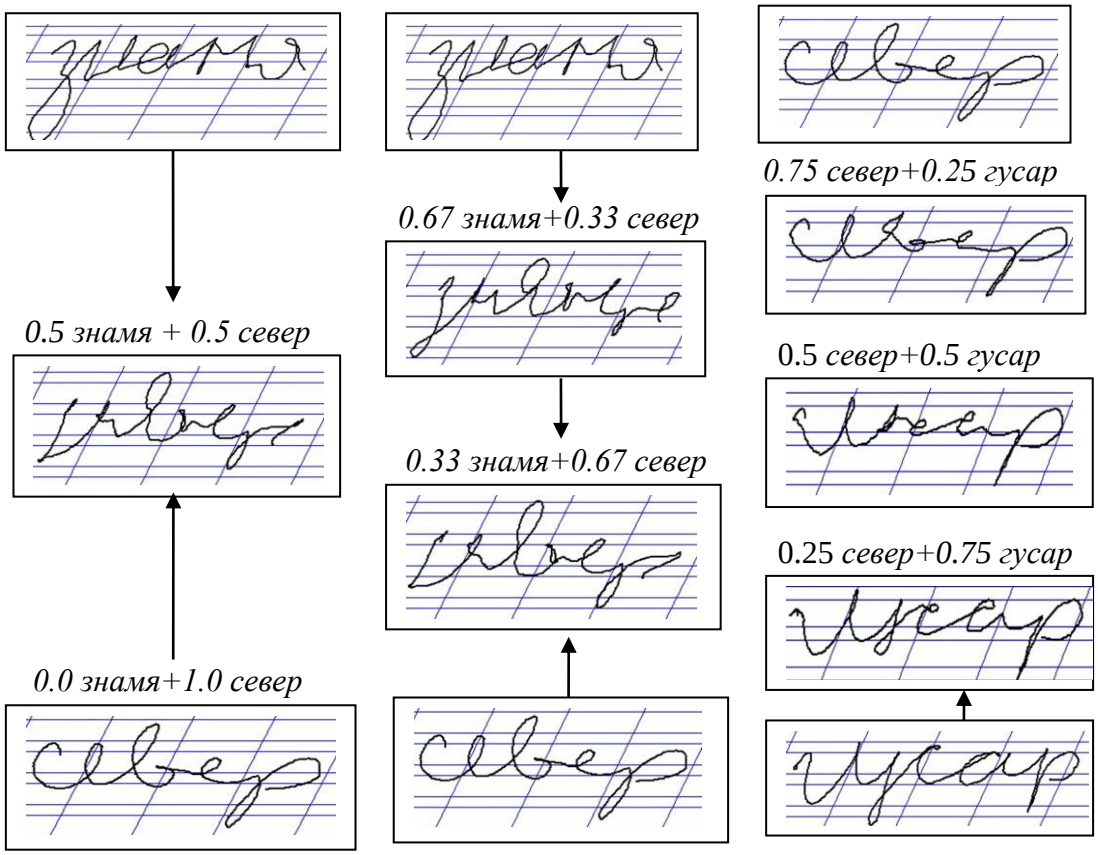


Рис. 3. Получение от образов-родителей 1, 2, 3 образов-потомков линейным морфингом по ГОСТ Р 52633.2

Увеличив размеры тестовой базы, например, в 1000 раз мы вновь можем проверить сколько на этой базе удается подобрать бит ключа. В случае, если будет обнаружен факт совпадения первых 44 бит, можно записать

$$П_{44} = 10\,000\,000. \qquad (5)$$

Очевидно, что постепенно увеличивая размеры тестовой базы, мы будем подбирать все больше и больше бит ключа. На рис. 4 приведена прямая, отражающая результаты подбора.

По результатам численного эксперимента получено время 0.47 секунды перебора первой базы естественных биометрических образов $П_{40}$. После того, как выполнено многократное размножение биометрических данных удалось подобрать первых 64 бита ключа за время 6 минут 51 секунда.

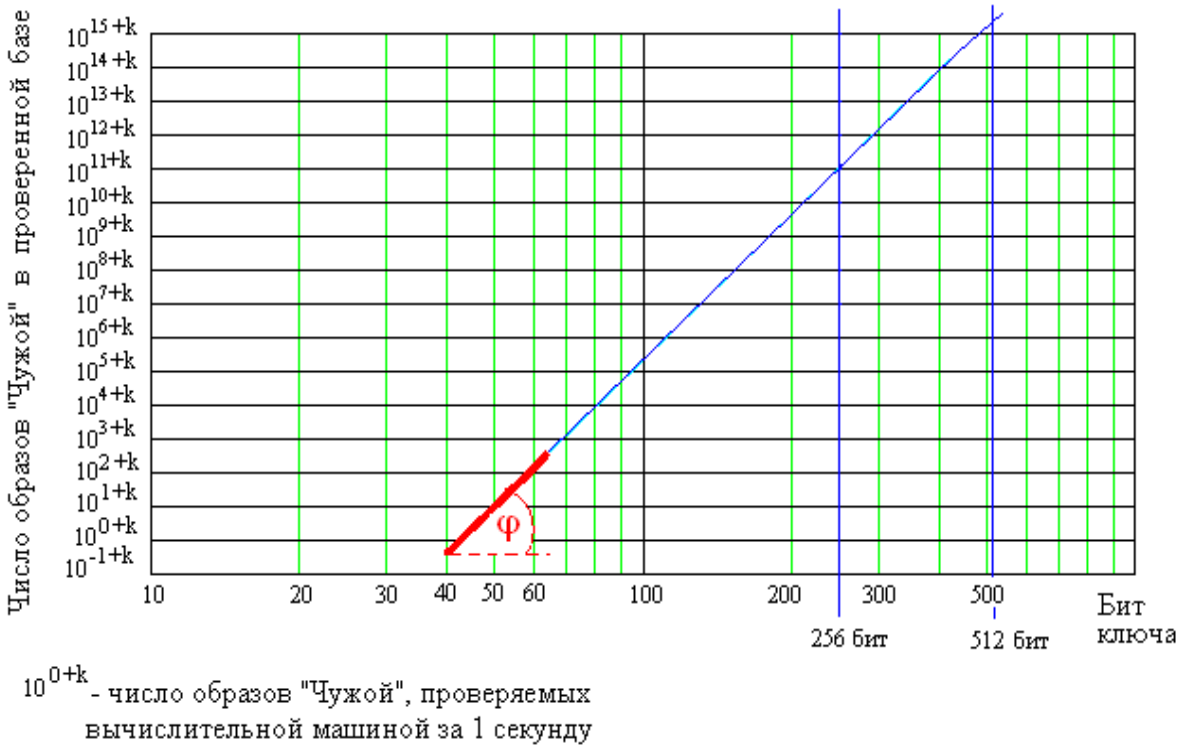


Рис. 4. Прогноз размеров тестовой базы образов «Чужой», необходимой для подбора 256 бит ключа на выходе нейросетевого преобразователя биометрия-код, обученного на биометрическом образе «slovo_01» (данные из приложения)

При реализации численного эксперимента производился последовательный подбор 40, 42, 44, …, 64 бит ключа с контролем размеров базы образов «Чужой» $П_{40}$, $П_{42}$, $П_{44}$, …., $П_{64}$. Как показано на рис. 4 промежуточные данные сливаются в одну прямую (утолщенная линия с углом наклонна – φ). Опираясь на результаты численного эксперимента, удается предсказать размер базы тестовых образов «Чужой», необходимой для подбора 256 бит ключа. Необходимые расчеты описываются, приведенной ниже системой из трех уравнений:

$$\begin{cases} tg(\varphi) = \dfrac{\log_{10}(П_{64}) - \log(П_{40})}{64 - 40}, \\ D = \log_{10}(П_{64}) + tg(\varphi) \cdot (256 - 64), \\ П_{256} \approx 10^{D}. \end{cases} \qquad (6)$$

Из данных рис. 4 следует, что на подбор 256 бит ключа должно уйти 10^{11} секунд, что составит примерно 3 000 лет непрерывной работы вычислительной машины, на которой проводился численный эксперимент. Нами использовалась вычислительная машина

с процессором Intel core i5-2500 (3.3 ГГц) ОЗУ 4Гб. При использовании компьютера большей мощности время на подбор сократится, однако оно остается достаточно большим для обычных настольных компьютеров. Размер базы образов «Чужой» инвариантен к мощности вычислительной машины.

Библиографический список

1. ГОСТ Р 52633.0–2006. Защита информации. Техника защиты информации. Требования к средствам высоконадежной биометрической аутентификации.

2. Нейросетевая защита персональных биометрических данных / Ю. К. Язов, В. И. Волчихин, А. И. Иванов, В. А. Фунтиков, И. Г. Назаров. – Москва : Радиотехника, 2012. – 157 с.

3. Juels, A. A Fuzzy Commitment Scheme / A. Juels, M. Wattenberg // Proc. ACM Conf. Computer and Communications Security. – Singapore, 1999. – P. 28–36.

4. Ramírez-Ruiz, J. Cryptographic Keys Generation Using FingerCodes / J. Ramírez-Ruiz, C. Pfeiffer J. , Nolazco-Flores // Advances in Artificial Intelligence – IBERAMIA-SBIA 2006 (LNCS 4140). – 2006. – P. 178–187.

5. Hao, F. Crypto with Biometrics Effectively / F. Hao, R. Anderson, J. Daugman. // IEEE TRANSACTIONS ON COMPUTERS. – 2006. – Vol. 55, № 9.

6. ГОСТ Р 52633.5–2011. Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа.

7. Техническая спецификация : проект [публичное обсуждение начато с 01.02.2017 членами ТК 26 «Криптографическая защита информации»]. Защита нейросетевых биометрических контейнеров с использованием криптографических алгоритмов.

8. Пат. 2346397 Российская Федерация. Способ защиты персональных данных биометрической идентификации и аутентификации / Иванов А. И., Фунтиков В. А., Ефимов О. В. ; приоритет от 26.06.07 ; опубл.10.02.2009, Бюл. № 4.

9. ГОСТ Р 34.12–2015. Информационная технология. Криптографическая защита информации. Режимы работы блочных шифров.

10. ГОСТ Р 34.11–2012. Информационная технология. Криптографическая защита информации. Функция хэширования.

11. ГОСТ Р 52633.3–2011. Защита информации. Техника защиты информации. Тестирование стойкости средств высоконадежной биометрической защиты к атакам подбора.

12. ГОСТ Р 52633.1–2009. Защита информации. Техника защиты информации. Требования к формированию баз естественных биометрических

образов, предназначенных для тестирования средств высоконадежной биометрической аутентификации.

13. ГОСТ Р 52633.2–2010. Защита информации. Техника защиты информации. Требования к формированию синтетических биометрических образов, предназначенных для тестирования средств высоконадежной биометрической аутентификации.

Майоров, А. В. Оценка стойкости защищенных нейросетевых преобразователей биометрия-код с использованием больших баз синтетических биометрических образов / А. В. Майоров, С. А. Сомкин, А. П. Юнин, А. Ж. Акмаев // Безопасность информационных технологий : тр. I Всерос. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2019. – С. 130–138.

УДК 004.056

А. В. Строков, Е. И. Казанцев

ПРОГРАММНОЕ СРЕДСТВО СОЗДАНИЯ ДЕЙСТВИТЕЛЬНО СЛУЧАЙНЫХ КРИПТОГРАФИЧЕСКИХ КЛЮЧЕЙ ИЗ НЕОДНОЗНАЧНОЙ КОМПОНЕНТЫ БИОМЕТРИЧЕСКИХ ДАННЫХ ДИНАМИКИ РУКОПИСНОГО ПОЧЕРКА ПОЛЬЗОВАТЕЛЯ

Аннотация. Гарантией случайности получаемого криптографического ключа является то, что неоднозначные (неповторяемые) биометрические данные носят случайный характер и число степеней свободы начальных условий достаточно велико. В программное обеспечение по созданию криптографического ключа из биометрических данных должен быть встроен тест, наследующий идеологию ГОСТ Р 52633.0. Этот стандарт рекомендует после каждого обучения искусственной нейронной сети на биометрии пользователя тестировать полученный результат.

A. V. Strokov, E. I. Kazantsev

SOFTWARE TOOL FOR CREATING TRULY RANDOM CRYPTOGRAPHIC KEYS FROM THE CONTROVERSIAL BIOMETRIC HANDWRITING DYNAMICS FEATURES OF HANDWRITING USER.

Abstract. Guarantee the randomness produced a cryptographic key is that ambiguous (not repeatable) random nature of biometric data and the number of degrees of freedom is a large enough initial conditions. Software for creating cryptographic key from biometric data must be embedded test inherits the ideology of GOST r 52633.0. This standard recommends that after each training artificial neural network to the user's biometric test result.

При реализации корректных криптографических преобразований добиться высокой надежности защиты информации удается в только в том случае, когда криптографический ключ имеет высокое качество. При создании криптографического ключа использовать псевдослучайные последовательности нежелательно. Качество ключа криптопреобразования напрямую зависит от алгоритма формирования случайных чисел и последовательностей, точнее от

их степени случайности. Использовать генераторы физического шума для получения ключа не всегда возможно, так как их наличие в аппаратуре является демаскирующим фактором.

Основная сложность генерации последовательности псевдослучайных чисел на компьютере состоит в том, что результат работы сложной программы генерации псевдослучайных чисел может быть подменен или искажен. Желательно, чтобы программа генерации БиоКодов была:

- компактной (не имела внешних вызовов сложных криптографических функций);
- имела электронную цифровую подпись для контроля ее целостности перед запуском;
- имела действительно случайные условия старта.

Особенность биометрии заключается в том, что при реализации одного и того же биометрического образа каждый пример реализации одного и того же рукописного образа будет отличаться от предыдущих [1].

Для создания криптографических ключей из биометрических данных динамического рукописного подчерка пользователя необходимо программное средство, которое считывает координаты движения мыши x и y (рис. 1).

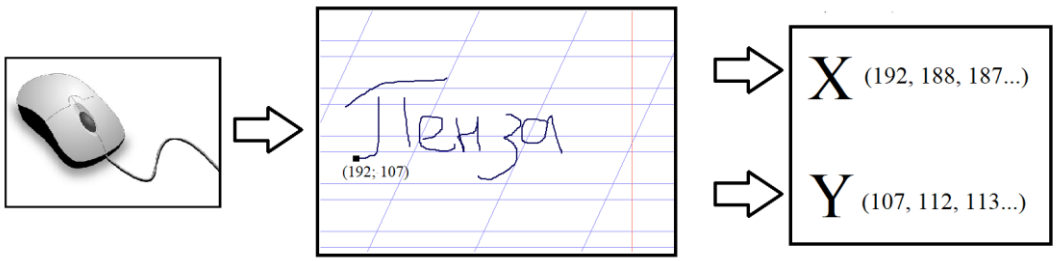


Рис. 1. Получение координат x и y из примера биометрического образа

После получения координат x и y проводится конкатенация и распределение координат в случайном порядке:

$$k_1 = hash(x||y), \quad (1)$$

$$k_2 = hash(y||x). \quad (2)$$

Операции конкатенации и распределения координат в случайном порядке на языке программирования C# представлены на рис. 2.

Случайная последовательность проверяется на случайность по средствам прохождения тестов NIST, их результат подтверждает случайность последовательности.

```
var text = File.ReadAllLines(@"D:\\xy.txt");
var result = string.Join("", text);
var rnd = new Random();
var shuffle = string.Join("", text.OrderBy(x => rnd.Next()));

Console.WriteLine(shuffle);

using (StreamWriter sw = new StreamWriter("D:\\rnd.txt", false))
{
    sw.WriteLine(shuffle);
}
```

Рис. 2. Операции конкатенации и распределения координат в случайном порядке

Список тестов NIST:

- частотный побитовый тест;
- частотный тест блоков;
- тест на длину последовательность одинаковых битов;
- тест на самую длинную последовательность единиц в блоке;
- тест рангов бинарных матриц;
- спектральный тест;
- тест на совпадение неперекрывающихся шаблонов;
- тест на совпадение перекрывающихся шаблонов;
- универсальный статистический тест Муарера;
- тест на линейную сложность;
- тест на периодичность;
- тест приблизительной энтропии;
- тест кумулятивных сумм;
- тест на произвольные отклонения.

Полное тестирование случайной последовательности по всем 14 тестам NIST для биометрических приложений, видимо, избыточно. Предположительно в программном обеспечении предлагается реализовать частотный побитовый тест и частотный тесто блоков (два первых теста комитета по стандартизации США).

От полученной последовательности по алгоритму MD5 вычисляется хэш-функция в виде строки из 32 шестнадцатеричных цифр. Вызов функции MD5, реализованный на языке программирования C# представлен на рис. 3.

Из разных примеров одного и того же биометрического рукописного образа получаются случайные последовательности различной длины из которых вычисляются отличные друг от друга хэш-функции (рис. 4).

```
StringBuilder hash = new StringBuilder();
MD5CryptoServiceProvider md5provider = new MD5CryptoServiceProvider();
byte[] bytes = md5provider.ComputeHash(new UTF8Encoding().GetBytes(input));

for (int i = 0; i < bytes.Length; i++)
{
    hash.Append(bytes[i].ToString("x2"));
}
return hash.ToString();
```

Рис. 3. Вызов функции MD5

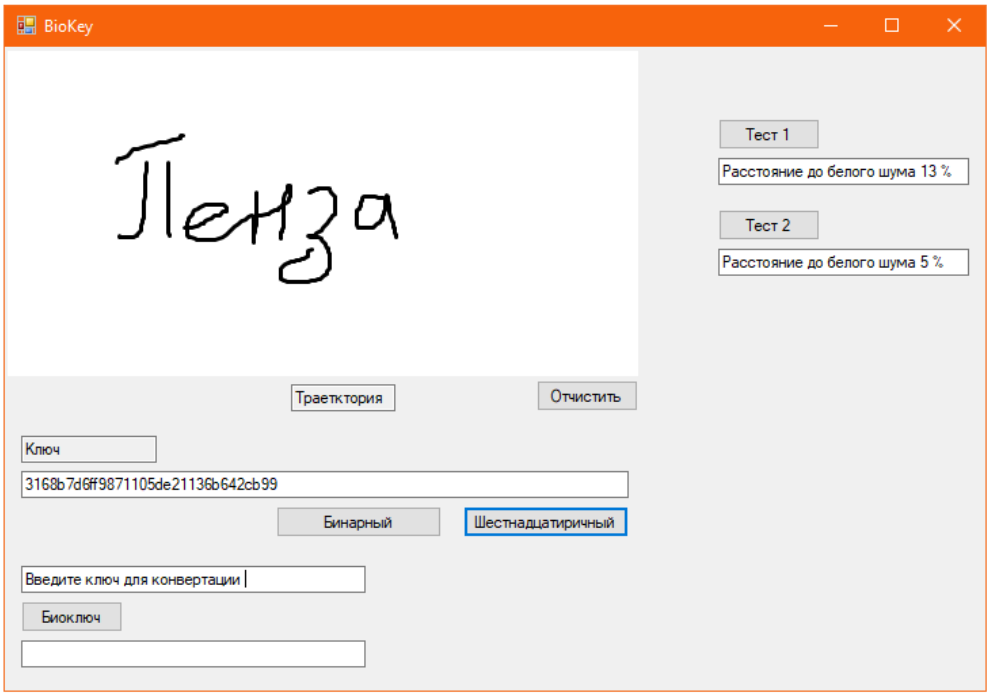


Рис. 4. Экранная форма разрабатываемого программного средства

Гарантией случайности получаемого криптографического ключа является то, что неоднозначные (не повторяемые биометрические данные носят случайный характер и число степеней свободы начальных условий достаточно велико). В рассматриваемом нами случае случайными являются:

- точка старта рукописного образа {x0, y0}, имеющая по каждой координате 512 возможных состояний;
- длина рукописного образа (от 1 до 2048 отсчетов);
- создаваемые последовательности движения пера или манипулятора «мышь» – x(t), y(t) рукописного образа;
- случайный порядок конкатенации последовательности хеширования случайных отсчетов x(t), y(t).

Все выше сказанное позволяет надеяться на то, что криптографическое хэширование данных функцией MD5 даст действительно случайные последовательности. Тем не менее в программное обеспечение по созданию криптографического ключа из биометрических данных должен быть встроен тест наследующий идеологию ГОСТ Р 52633.0 [2]. Этот стандарт рекомендует после каждого обучения искусственной нейронной сети на биометрии пользователя тестировать полученный результат. Так как мы получаем случайную последовательность из биометрии, мы также должны ее тестировать на близость к «белому шуму».

Библиографический список

1. Юнин, А. П. Оценка качества «белого шума»: реализация теста стаи обезьян через множество сверток Хэмминга, построенных для разных систем счисления / А. П. Юнин, А. И. Иванов, К. А. Ратников // Информационно-управляющие и телекоммуникационные системы специального назначения : Всерос. науч.-техн. конф., посвящ. 100-летию со дня рождения одного из основоположников советской вычислительной техники Б. И. Рамеева, г. Пенза, 2018. – URL: Rosoperator.ru/info/docs/Yunin.pdf

2. ГОСТ Р 52633.0–2006. Защита информации. Техника защиты информации. Требования к средствам высоконадежной биометрической аутентификации.

Строков, А. В. Программное средство создания действительно случайных криптографических ключей из неоднозначной компоненты биометрических данных динамики рукописного почерка пользователя / А. В. Строков, Е. И. Казанцев // Безопасность информационных технологий : тр. I Всерос. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2019. – С. 139–143.

УДК 519.2; 519.6

А. Г. Банных, В. И. Семенов

ОСОБЕННОСТИ ПРОГРАММНОЙ РЕАЛИЗАЦИИ СВОБОДНО РАСПРОСТРАНЯЕМОГО КАЛЬКУЛЯТОРА ДЛЯ БЫСТРОГО ВЫЧИСЛЕНИЯ ЭНТРОПИИ ВЫХОДНЫХ СОСТОЯНИЙ НЕЙРОСЕТЕВЫХ ПРЕОБРАЗОВАТЕЛЕЙ БИОМЕТРИИ В КОД ДЛИНОЙ 256 БИТ

Аннотация. Разрабатываемый программный калькулятор будет способен вычислять энтропию данных, имеющих значительное отличие распределений Хэмминга от нормального закона распределения значение. Этого эффекта удается достичь за счет перехода в пространство бета-распределений. Предположительно, удастся снизить методическую погрешность вычислений примерно на порядок. А также позволит заменить коммерческие программные продукты MathCAD, MathLAB и другие, способные выполнять вычисления с высокой разрядностью. Этот эффект достигается применением дополнительных таблиц оптимизации масштабов, вычислений. В каждой ячейке табличных вычислений разрабатываемый калькулятор ведет расчеты в своем масштабе, что в конечном итоге позволяет устранить необходимость в вычислениях с плавающей запятой в разрядной сетке 64 бита.

A. G. Bannich, V. I. Semenov

FEATURES PROGRAMMING FREE CALCULATOR FOR FAST CALCULATION OF THE ENTROPY OF THE OUTPUT STATES OF NEURAL NETWORK IN LONG CODE BIOMETRICS CONVERTERS 256 BIT

Abstract. Developed software calculator will be able to calculate the entropy of the data, with a significant difference from the normal Hamming distributions distribution law value. This effect can be achieved at the expense of moving in space beta distributions. Supposedly will reduce methodological error calculations about the order. And also will replace commercial software products and other MathLAB, MathCAD, capable of performing calculations with high bit depth. This effect is achieved using additional tables optimization magnitude calculations. In each cell of the spreadsheet, calculator developed in its scale calculations leads that eventually helps eliminate the need for floating-point computations in bit 64 bit grid.

В настоящее время активно идут процессы информатизации современного общества, и, как следствие, появляются проблемные аутентификации в информационных системах. Так как зачастую пользователи не способны генерировать «сложные» пароли, и даже та часть пользователей, кто на это способен в конечном счете их не запоминает, следовательно, появляется необходимость перехода от «сложных» сгенерированных паролей к паролям, в основе которых лежат биометрические данные пользователя [1].

Задача оценки энтропии для биометрических образов не может быть решена с помощью классического подхода. Классическое определение информационной энтропии и формулы для ее расчета были введены Шенноном. Основной недостаток классического многомерного вычисления энтропии состоит в том, что требуются огромные размеры исходных данных. Как правило, требуется массив исходных данных, размером превышающий число возможных состояний исследуемого кода. Кроме того, классический метод вычисления многомерной энтропии требует огромных вычислительных затрат. Выходом является вычисление многомерной энтропии в пространстве мер Хемминга [2].

В России на данный момент используется стандартизированная технологи нейросетевого преобразования биометрии пользователя в код пароля доступа. Данная технология должна реализоваться в соответствии с пакетом стандартов ГОСТ Р 52633.xx-20xx. Который рекомендует производить вычисления в расстояниях Хэмминга. Например, ГОСТ Р 52633.3 рекомендует использовать для тестирования вероятностей ошибок второго рода до 32 примеров образов «Чужой» не участвующих ранее в обучении нейросетевого преобразователя [3]. Результат изображен на рис. 1.

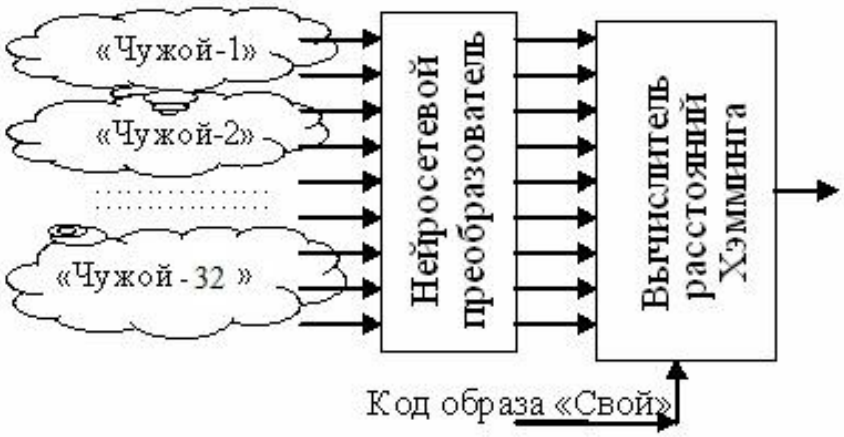


Рис. 1. Тестирование вероятности ошибок второго рода

Однако, данный стандарт предполагает использовать нормальный закон распределения значений, но если использовать бета распределение, то мы можем повысить точность оценок [4]. Для этого воспользуемся формулами математического ожидания E(h) и дисперсии, где α и β произвольные фиксированные параметры:

$$\text{Математическое ожидание} \left| \frac{\alpha}{\alpha+\beta}. \right. \quad (1)$$

$$\text{Дисперсия} \left| \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}. \right. \quad (2)$$

Получив значения, переходим к вычислению стандартного отклонения по формуле

$$\sigma(\beta 1) = \sqrt{\frac{\beta 1\left(\frac{\beta 1-\beta 1\cdot E}{E}\right)}{\left[\beta 1+\left(\frac{\beta 1-\beta 1\cdot E}{E}\right)\right]^2\cdot\left[\beta 1+\left(\frac{\beta 1-\beta 1\cdot E}{E}\right)+1\right]}}, \quad (3)$$

где β1 =[…] – данный параметр задает интервал значений;

β2 = ((β1-β1*E)/E);

E – нормированное значение математического ожидания расстояний Хэмминга.

Если мы возьмем интервал β1 = [1,..., 256], то для пяти значений нормированных математических ожиданий E(h) = [0.15, 0.20, 0.25, 0.35, 045] дают номограмму стандартного отклонения, приведенную на рис. 2.

Как видно на графике, стандартное отклонения на прямую зависит от матожидания.

Если выберем интервал [100…200], то график будет иметь вид, представленный на рис. 3.

После этого было предложено вычислить значения стандартного отклонения на заданном интервале, полученные значения отображены в табл. 1.

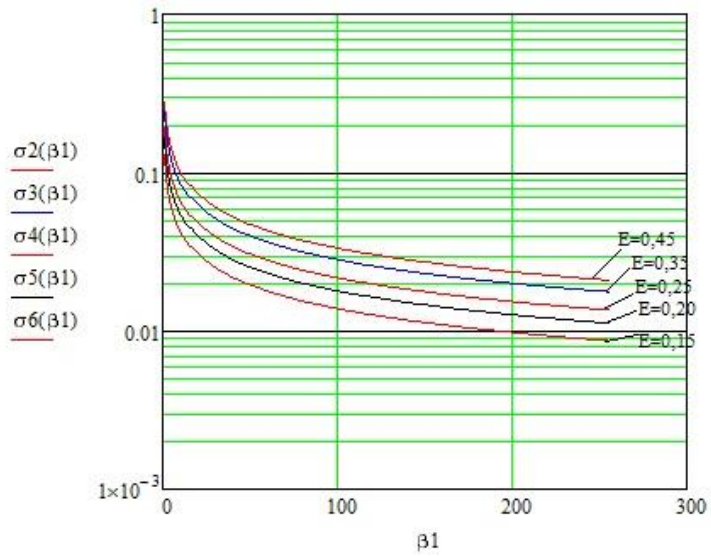


Рис. 2. Номограмма зависимости стандартного отклонения от математического ожидания

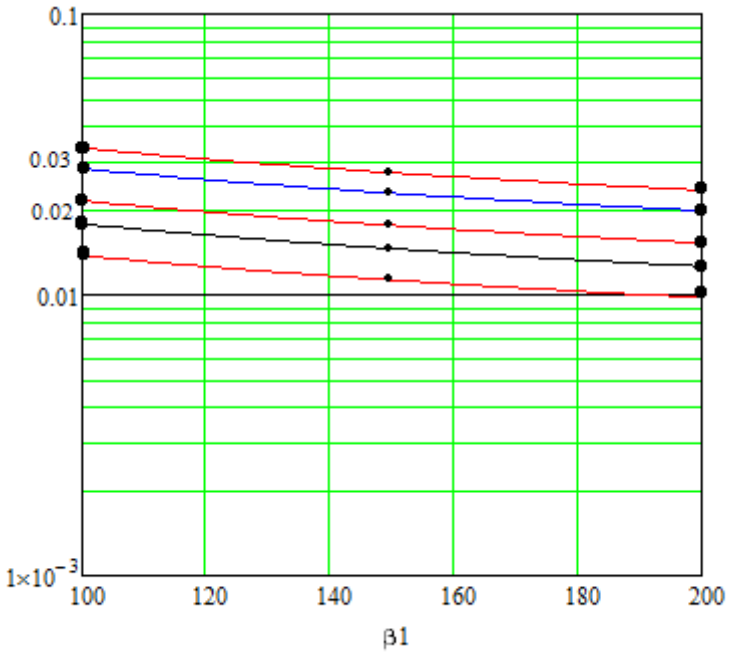


Рис. 3. График в интервале от 100 до 200

Таблица 1

Значения стандартного отклонения

	β1 = 100	β1 = 150	β1 = 200
$\delta 2(\beta 1)$	0.033	0.027	0.024
$\delta 3(\beta 1)$	0.028	0.023	0.02
$\delta 4(\beta 1)$	0.022	0.018	0.015
$\delta 5(\beta 1)$	0.018	0.015	0.013

$\delta 6(\beta 1)$	0.014	0.011	0.009

Из табл. 2 видно, что данные в ней имеют низкую точность. Необходимо увеличить точность вычислений.

Таблица 2

Значения стандартного отклонения с увеличенной точностью

	β1=100	β1=150	β1=200
$\delta 2(\beta 1)$	0.033298	0.027208	0.023572
$\delta 3(\beta 1)$	0.028169	0.023013	0.019936
$\delta 4(\beta 1)$	0.021624	0.017663	0.0153
$\delta 5(\beta 1)$	0.017871	0.014596	0.012642
$\delta 6(\beta 1)$	0.013819	0.011286	0.009775

Проанализировав полученные более точные данные, можно сделать вывод о том, что подобный рост точности представления данных избыточен для создания программного калькулятора с 16-битовым вычислениям на языке C#.

При вычислениях целесообразно ввести смещение и масштаб для значений стандартного отклонения выравнивающих точность вычислений по координатам таблицы с учетом принятой разрядной сетки. Кроме того, для программирования калькулятора было решено перейти в диапазон целых чисел, чтобы сделать программу более универсальной. Итоговые значения с оптимизированной точностью будут иметь вид, представленный в табл. 3.

Таблица 3

Итоговые целочисленные значения стандартного отклонения с увеличенной точностью

	β1=100	β1=150	β1=200	Масштаб	Целочисленные значения при β1=100	Целочисленные значения при β1=150	Целочисленные значения при β1=200
$\delta 2(\beta 1)$	0.0333	0.0272	0.0236	10000	333	272	236
$\delta 3(\beta 1)$	0.0282	0.0230	0.0199	10000	282	230	199
$\delta 4(\beta 1)$	0.0216	0.0177	0.0153	10000	216	177	153
$\delta 5(\beta 1)$	0.0178	0.0145	0.0126	20000	356	290	252
$\delta 6(\beta 1)$	0.0138	0.0113	0.0098	20000	276	226	176

Далее, получив значения математического ожидания, дисперсии и стандартного отклонения можно перейти к вычислению энтропии. В данном случае энтропия вычисляется по формуле:

$$H("x_1, x_2, ..., x_{256}") \approx -\log_2(P_2), \quad (4)$$

где P_2 – это вероятность возникновения ошибки второго рода.

Вероятность возникновения ошибки второго рода в данном случае вычисляется по формуле:

$$P_2 = \frac{(\beta 1 + \beta 2 + 1)!}{\beta 1! \cdot \beta 2!} \cdot \int_0^{\frac{1}{256}} x^{\beta 1} \cdot (1 - x)^{\beta 2} dx, \quad (5)$$

где $x = \frac{h}{\max(h)}$ – нормированное расстояние Хэмминга; β1 и β2 – первый и второй параметры настройки бета-распределения [4, 5].

Имея все необходимые значения, вычислим энтропию на примере значений стандартного отклонения $\delta 2(\beta 1)$ при математическом ожидании E = 0, 45. Полученные результаты представлены в табл. 4.

Таблица 4

Результаты вычислений энтропии

$E(h)$	$\delta(h)$	β1	β2	P_2	$\log_2(P_2)$
0,15	333	100	566	$10^{-13.2}$	42.901
0,15	272	150	850	$10^{-18.87}$	62.865
0,15	236	200	1133	$10^{-24.90}$	82.979

Чтобы вычислить энтропию выходных состояний энтропии нейросетевых преобразователей биометрия – код в пространстве сверток Хэмминга требуются большие временные затраты необходимо разработать программное средство (калькулятор), способное по выше сказанным формулам осуществлять быструю оценку энтропии выходных состояний нейросетевых преобразователей биометрия – код в пространстве сверток Хэмминга.

Результат работы программного средства представлен на рис. 4.

Новый программный калькулятор будет иметь существенно лучшие характеристики, чем сегодняшняя версия.

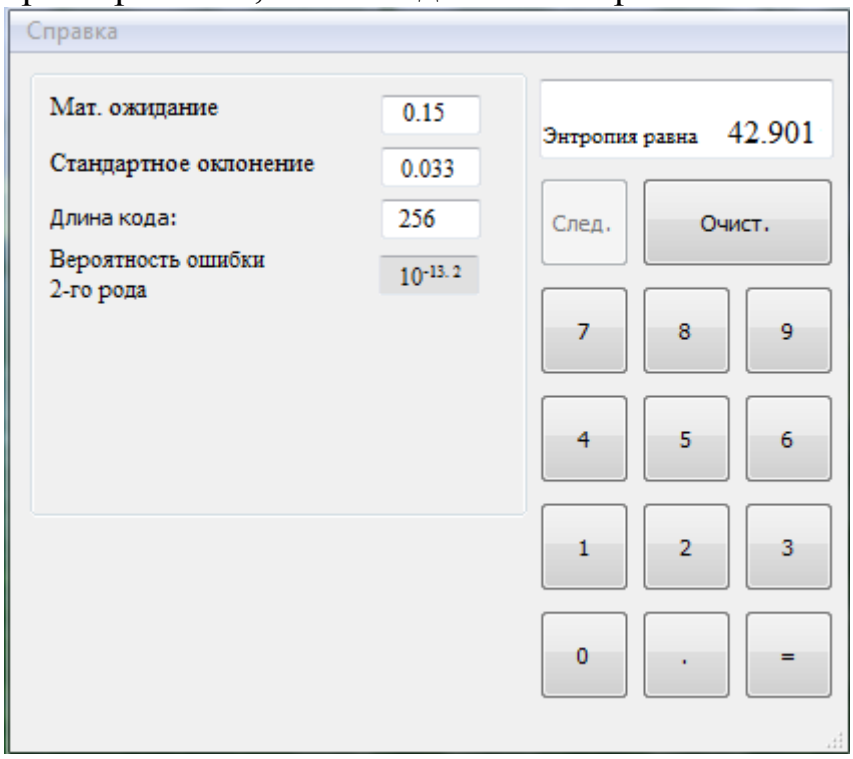


Рис. 4. Экранная форма разрабатываемого программного средства (калькулятора для оценки энтропии и вероятности ошибок второго рода)

Во-первых, он будет способен вычислять энтропию данных, имеющих значительное отличие распределений Хэмминга от нормального закона распределения значение. Этого эффекта удается достичь, за счет перехода в пространство бета-распределений. Предположительно, удастся снизить методическую погрешность вычислений примерно на порядок.

Во-вторых, разрабатываемый программный калькулятор позволит заменить коммерческие программные продукты MathCAD, MathLAB и другие, способные выполнять вычисления с высокой разрядностью. Этот эффект достигается применением дополнительных таблиц оптимизации масштабов, вычислений. В каждой ячейке табличных вычислений, разрабатываемый калькулятор ведет расчеты в своем масштабе, что в конечном итоге позволяет устранить необходимость в вычислениях с плавающей запятой в разрядной сетке 64 бита.

Библиографический список

1. Технология использования больших нейронных сетей для преобразования нечетких биометрических данных в код ключа доступа : монография / Б. С. Ахметов, А. И. Иванов, В. А. Фунтиков, А. В. Безяев, Е. А. Малыгина. – Алматы : Изд-во LEM, 2014. – 144 с. – URL: http://portal.kazntu.kz/files/publicate/2014-06-27-11940.pdf

2. Куликов, С. В. Оценка энтропии биометрических образов через переход к дискретному представлению с асимметричным представлением меры Хэмминга / С. В. Куликов // Безопасность информационных технологий : тр. науч.-техн. конф. кластера пензенских предприятий. – Пенза, 2012. – Т. 8. – С. 18–21. – URL: http://пниэи.рф/activity/science/BIT/T8-p18.pdf

3. ГОСТ Р 52633.3–2011. Защита информации. Техника защиты информации. Тестирование стойкости средств высоконадежной биометрической защиты к атакам подбора.

4. Елфимов, А. В. Сравнение гипотезы нормального распределения и гипотезы бета-распределения расстояний Хэмминга для выходных кодов нейросетевых преобразователей / А. В. Елфимов, А. В. Безяев // Безопасность информационных технологий : тр. науч.-техн. конф. кластера пензенских предприятий. – Пенза, 2012. – Т. 10. – С. 10–14. – URL: http://пниэи.рф/activity/science/BIT/T10-p10.pdf

5. Юнин, А. П. Оценка энтропии легко запоминаемых, длинных паролей со смыслом в ASCII кодировке для русского и английского языков / А. П. Юнин, О. В. Корнеев // Безопасность информационных технологий : тр. науч.-техн. конф. кластера пензенских предприятий. – Пенза, 2016. – Т. 10. – С. 40–42. – URL: http://пниэи.рф/activity/ science/BIT/T10-p10.pdf

Банных, А. Г. Особенности программной реализации свободно распространяемого калькулятора для быстрого вычисления энтропии выходных состояний нейросетевых преобразователей биометрии в код длиной 256 бит / А. Г. Банных, В. И. Семенов // Безопасность информационных технологий : тр. I Всерос. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2019. – С. 144–151.

УДК 519.2; 519.6

С. Ь. Гуреев

СТАТИСТИЧЕСКАЯ ОЦЕНКА ПРОЦЕССА ИЗМЕНЕНИЯ СВОЙСТВ БИОМЕТРИЧЕСКИХ ОБРАЗОВ ТЕСТОВОЙ БАЗЫ «ЧУЖИЕ» ПРИ ИХ ИСКУССТВЕННОМ РАЗМНОЖЕНИИ

Аннотация. Рассматривается процесс вырождения корреляционных матриц биометрических образов при их размножении процедурами ГОСТ Р 52633.2. Дается статистическая оценка скорости утраты реальных значений коэффициентов корреляции, подтвержденная результатами численного эксперимента. Приводятся условия, при которых вырождение корреляционных матриц биометрических образов полностью исчезает, что открывает перспективу разработки новых способов более корректного размножения биометрических данных.

S. V. Tureev

STATISTICAL EVALUATION OF PROCESS CHANGES THE PROPERTIES OF A BIOMETRIC TEST IMAGE DATABASES "STRANGERS" WHEN THEIR ARTIFICIAL REPRODUCTION

Abstract. The process of degeneration of correlation matrices of biometric images when their reproduction procedures GOST r 52633.2. Provides a statistical estimate of the rate of loss of actual values of correlation coefficients, confirmed the results of numerical experiment. Are the conditions in which degeneration of correlation matrices of biometric images completely disappears, that opens up the prospect of developing new ways to more correct reproduction of biometric data.

В настоящее время распознавание личности по его биометрическим данным получило широкое распространение в мире. В России в 2006 г. разработан и введен в действие ГОСТ Р 52633.0 [1], который предполагает использование искусственных нейронных сетей большого размера при биометрической аутентификации личности.

Нейросетевые преобразователи биометрии на выходе имеют длинные выходные коды и, соответственно, требуют использования специальных процедур их тестирования, которые описаны в

ГОСТ Р 52633.3 [2]. Для их реализации необходимо создавать тестовые базы объемом до 10 000 биометрических образов, собранных по требованиям ГОСТ Р 52633.1 [3]. При этом возникает проблема недостаточного размера тестовых баз естественных биометрических образов. Это проблему решает ГОСТ Р 52633.2 [4], рекомендующий получать из пар естественных биометрических образов-родителей синтетические образы-потомки путем морфинга биометрических параметров образов-родителей.

Однако существует негативный эффект морфинг-размножения, приводящий к вырождению корреляционных связей образов-потомков. Устранение этого негативного эффекта или его частичная компенсация должны упростить процедуру обращения матриц нейросетевых функционалов [5, 6], снизив число поколений образов-потомков, при направленном извлечении знаний из обученной нейронной сети.

Для оценки скорости вырождения корреляционных связей [7] был проведен численный эксперимент, по которому были построены корреляционные матрицы для 16-рукописных образов, образующих поколение-1.

Далее биометрические образа попарно скрещивались и их них были получены образы-потомки одинаково похожие на образы-родители поколения-2.

Попарное скрещивание образов-родителей выполнялось в плоть до поколения-5, при этом для каждого нового образа вычислялась его корреляционная матрица и соответствующее ей стандартное отклонение коэффициентов корреляции.

Из полученных данных, описанных в работе [8], видно, что в следующем поколении математическое ожидание стандартных отклонений коэффициентов корреляции монотонно падает:

$$E_1(\sigma(r)) \approx 0.313 > E_2(\sigma(r)) \approx 0.265 > E_3(\sigma(r)) \approx$$
$$\approx 0.196 > E_4(\sigma(r)) \approx 0.169 > E_5(\sigma(r)) \approx 0.144.$$

С ростом числа поколений у синтезированных биометрических образов-потомков наблюдается сжатие плотностей распределения значений коэффициентов корреляции. В первом приближении можно считать, что значение стандартного отклонения коэффициентов корреляции убывает пропорционально квадратному корню номера поколения:

$$E_n(\sigma(r)) \approx \frac{E_1(\sigma(r))}{\sqrt{n}}.$$

Очевидным является так же то, что корреляционная матрица биометрического данных образа никак не зависит от значений математических ожиданий ее параметров. Это означает, что для полного сохранения корреляционной матрицы достаточно синтезировать новые биометрические образы путем нескольких сотен перестановок математических ожиданий биометрических параметров. Формальная запись одной из перестановок двух математических ожиданий иллюстрируется рис. 1.

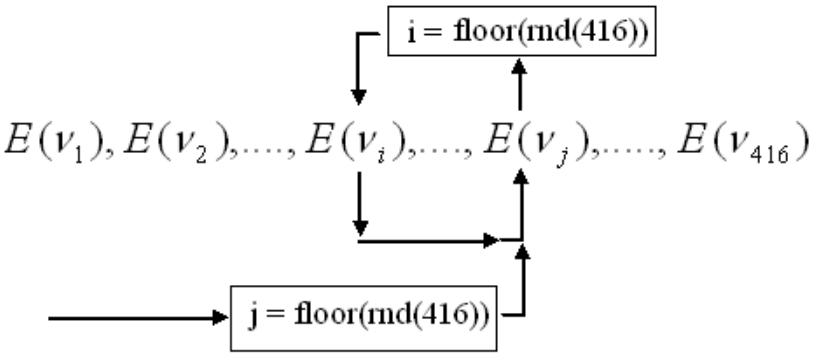


Рис. 1. Механизм перестановки двух математических ожиданий при синтезе нового биометрического образа, сохраняющего первоначальную корреляционную матрицу

Заключение

Предположительно, что сочетание еще одной операции синтеза биометрических образов с уже имеющимися операциями, будет достаточно для эффективного тестирования нейросетевых преобразователей биометрия-код. В новой версии стандарта должны появиться рекомендации о том, в каком сочетании следует использовать циклические случайные перестановки математических ожиданий биометрических данных и размножение образов путем морфинг-скрещивания образов-родителей.

Библиографический список

1. ГОСТ Р 52633.0–2006. Защита информации. Техника защиты информации. Требования к средствам высоконадежной биометрической аутентификации.

2. ГОСТ Р 52633.3–2011. Защита информации. Техника защиты информации. Тестирование стойкости средств высоконадежной биометрической защиты к атакам подбора.

3. ГОСТ Р 52633.1–2009. Защита информации. Техника защиты информации. Требования к формированию баз естественных биометрических образов, предназначенных для тестирования средств высоконадежной биометрической аутентификации.

4. ГОСТ Р 52633.2–2010. Защита информации. Техника защиты информации. Требования к формированию синтетических биометрических образов, предназначенных для тестирования средств высоконадежной биометрической аутентификации.

5. Волчихин, В. И. Нейросетевая молекула: решение обратной задачи биометрии через программную поддержку квантовой суперпозиции на выходах сети искусственных нейронов / В. И. Волчихин, А. И. Иванов // Вестник Мордовского университета. – 2017 . – Т. 27, № 4. – С. 518–523.

6. Ахметов, Б. С. Дополнение нечетких биометрических данных морфинг-размножением примеров родителей в нескольких поколениях примеров потомков / Б. С. Ахметов, С. В. Качалин, А. И. Иванов // Вестник КазНТУ им. К. И. Сатпаева. – Алматы, 2014. – № 4 (104). – С. 194–199.

7. Качалин, С. В. Направленное морфинг-размножение биометрических образов, исключающее эффект вырождения их популяции / С. В. Качалин, А. И. Иванов // Вопросы радиоэлектроники. – 2015. – № 1 (1). – С. 69–76.

8. Туреев, С. В. Численный эксперимент по оценке скорости вырождения корреляционных матриц биометрических образов при размножении данных морфинг-скрещиванием образов-родителей в нескольких поколениях / С. В. Туреев, А. И. Иванов, А. И. Солопов // Охрана, безопасность, связь – 2018 : сб. материалов Междунар. НПК (22 ноября 2018 г.). – Воронеж : Институт МВД, 2018. – С. 67–73.

Туреев, С. В. Статистическая оценка процесса изменения свойств биометрических образов тестовой базы «Чужие» при их искусственном размножении / С. В. Туреев // Безопасность информационных технологий : тр. I Всерос. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2019. – С. 152–155.

УДК 004.93; 57.087.1; 681.139.3

В. Л. Назаров, Д. А. Данилин, М. Д. Суворов, А. А. Устинов

СОВРЕМЕННЫЕ СИСТЕМЫ КОНТРОЛЯ И УПРАВЛЕНИЯ ДОСТУПОМ НА ОСНОВЕ ТЕХНОЛОГИЙ БИОМЕТРИЧЕСКОЙ И ЭЛЕКТРОННОЙ ИДЕНТИФИКАЦИИ ЛИЧНОСТИ

Аннотация. Рассмотрен ряд наиболее популярных способов идентификации личности по биометрическим параметрам, выделены их достоинства и недостатки. Описаны существующие и перспективные технологии идентификации с помощью бесконтактных карт. Проведен обзор современных контроллеров систем контроля и управления доступом (СКУД), включающих в качестве факторов идентификации биометрические параметры человека: ZKTeco MA300, BioSmart-WTC-2-EM, Suprema FS2-D, EyeLock NANO NXT. Проведено сравнение данных моделей по критериям объема памяти, наличия аппаратных интерфейсов взаимодействия с другими компонентами СКУД, их стоимости, факторов идентификации, дополнительных возможностей и др. Выделены общие и частные (для отдельных технологий) особенности эксплуатации биометрико-электронных СКУД. Сделан вывод о перспективности работы в области развития данных технологий.

V. L. Nazarov, D. A. Danilin, M. D. Suvorov, A. A. Ustinov

MODERN ACCESS CONTROL SYSTEMS BASED ON BIOMETRIC AND ELECTRONIC PERSON IDENTIFICATION TECHNOLOGIES

Abstract. The advantages and disadvantages of the most popular methods of biometric identification are considered. Existing and future technologies of contactless magnetic cards are described. A review of modern biometric controllers of access control systems (ACS): ZKTeco MA300, BioSmart-WTC-2-EM, Suprema FS2-D, EyeLock NANO NXT. A comparison was made of their memory size, cost, interfaces, ways of identifying a person, additional capabilities, etc. The general and particular features of the work of biometric ACS are considered. It was concluded about the prospects of work in the field of development of these technologies.

Характерной чертой развития современных систем контроля и управления доступом (а также иных систем, одной из функций

которых выступает идентификации личности) является все большее распространение технологий идентификации личности по ряду биометрических параметров. Об этом свидетельствует рост объема их применения в различных секторах экономики. Примерами тому служат эксплуатация системы распознавания людей в потоке по изображению лица в ГУП «Московский метрополитен», идентификация клиентов ПАО «Сбербанк» по изображению лица и голосу при обращении в банк, допуск людей на территорию Генерального штаба ВС РФ с контролем рисунка вен ладони и т.д. [1]. Преимуществом таких СКУД является их повышенная надежность, достигаемая сложностью воспроизведения контролируемых биометрических параметров злоумышленником, которая позволяет снизить число случаев несанкционированного доступа (НСД) к защищаемым объектам и информационным ресурсам. Наиболее популярными технологиями биометрической идентификации личности в настоящий момент являются технологии распознавание человека по папиллярному узору пальцев (55 % от общего числа), 2D/3D изображению лица человека (17 %), радужной оболочке и сетчатке глаза (7 %), рисунку вен ладони (7 %) [2]. СКУД на основе технологий контроля отпечатка пальца с высокой степенью защиты позволяют идентифицировать пользователя с точностью до 0,001 %, при невысокой стоимости оборудования и малом времени обработки данных [3]. Однако качество их работы может сильно снижаться для отдельных групп пользователей (пожилые люди, строители и др., чьи пальцы подвержены механическим микроповреждениям) и в отдельных условиях эксплуатации (влажная холодная среда, деятельность пользователей, связанная с загрязнением рук и др.). Еще одним недостатком систем на основе данных технологий является их низкая гигиеничность из-за постоянного контакта пальцев рук всех пользователей со сканерами. Кроме того, значительное развитие получили методы подделки отпечатков пальцев. Например, разработанная исследователями Политехниче-ского института Нью-Йоркского университета (New York University Tandon) DeepMasterPrints – модель нейронной сети способна достраивать отдельные части отпечатков пальцев людей на основе статистических данных, что позволяет в ряде случаев успешно обмануть систему идентификации [4]. Технологии идентификации личности, связанные с анализом радужной оболочки и сетчатки глаза, отличаются высокой надежность – FAR (False Acceptance Rate, ошибка распознавания 2 рода) для них составляет 0,00001 %, а FRR (False Rejection Rate, ошибка распознавания 2 рода) – 0,13 %

[3]. Однако значительная их стоимость, а также большое время обработки каждого пользователя не позволяют говорить об их пригодности к установке в местах с высокой интенсивностью потока людей. Все большую популярность обретают системы распознавания людей по изображению лица. В результате использования в их составе нейронных сетей с большим числом внутренних слоев нейронов, возможно достижение показателя ошибки FAR порядка 0,0001 % [3]. Данные решения идеально подходят для КПП, входов в помещения и корпуса предприятий с высокой интенсивностью потока людей. Основным их недостатком является низкая степень масштабируемости, связанная с неизбежным ростом ошибок идентификации, времени обработки и требований к вычислительным ресурсам ЭВМ при увеличении числа пользователей и/или объема анализируемой информации. Протекание процесс внедрения вышеописанных технологий в современные СКУД позволяет говорить не о вытеснении, а о дополнении уже существующих «классических» решений. Наиболее популярным симбиозом в данной области является использование бесконтактных карт семейств EM-Marin или Mifare в совокупности со сканерами биометрических параметров человека. Карты доступа EM-Marin представляют собой бесконтактные карты, работающие на частоте 125 КГц. Данные карты содержат идентификатор – неизменяемый серийный номер встроенного чипа, хранимый без защиты от копирования. Карты Mifare Classic передают данные на частоте 13,56 МГц и обладают энергонезависимой перезаписываемой памятью до 4 КБайт. В отличие от EM-Marin, данные карты позволяют использовать в качестве идентификатора не только код чипа, но и криптографический ключ, а модификации Mifare Plus имеют еще и усиленную защиту от копирования. Но несмотря на это, в ходе эксплуатации данных карт, были неоднократно выявлены уязвимости, способные серьезно снизить надежность СКУД на их основе [5]. В настоящее время рынок систем безопасности насыщен моделями различной стоимости и назначения (рис. 1), (табл. 1). Представителем эконом-класса является модель ZKTeco MA300, выполненная в виде небольшого терминала, в котором объединены считыватель бесконтактных карт EM-Marine, сканер отпечатка пальца и контроллер доступа. Емкость памяти устройства составляет 1500 геометрических моделей пальцев, 1000 номеров бесконтактных карт и до 100 000 событий. Из дополнительных возможностей модели можно отметить ее защищенность от воздействия влажного, горячего и холодного воздуха, а также наличие интерфейса Ethernet. Данный контроллер

широко зарекомендовал себя как недорогое и надежное решение с широкими возможностями использования. Данная модель появилась на рынке в 2013 году и на сегодняшний день ее цена не превышает 10 тыс. рублей, что во многом является основой его популярности [6].

Представителем среднего класса обозреваемых устройств является BioSmart-WTC-2-EM. Это терминал для учета рабочего времени сотрудников посредством идентификации их по отпечаткам пальцев и/или по картам доступа. Существенным отличием терминала от предыдущей модели является наличие LCD дисплея, благодаря которому регистрацию отпечатков и пластиковых карт можно производить непосредственно на терминале, а также больший объем памяти (до 4500 отпечатков пальцев, до 3000 бесконтактных карт). Модель имеет возможность взаимодействовать с элементами СКУД с помощью 2 интерфейсов: Ethernet и USB. Данный терминал был представлен в 2014 году, его актуальная на сегодняшний день цена составляет порядка 35 тыс. рублей [7]. Suprema FS2-D – высокопроизводительный терминал распознавания лиц, со встроенным мультичастотным считывателем карт премиум-класса. Данный терминал очень хорошо подходит для крупных объектов – автономная память на 30 тыс. пользователей, 5 млн. событий 50 тыс. фото. Благодаря расширенной области распознавания, установленное согласно инструкции устройство «видит» лица людей ростом от 145 см. до 210 см. Данный терминал поддерживает как проводные интерфейсы Ethernet, RS485 и USB, так и Wi-Fi, а также предусматривает возможность использования смартфона вместо карты доступа. На рынке Suprema FS2-D появился в 2017 году, на настоящий момент цена на него находится в диапазоне 80–85 тыс. рублей [8]. EyeLock NANO NXT – представитель следующего поколения биометрических считывателей в основе которого лежит принцип сканирования радужной оболочки глаза. Данный терминал позволяет хранить до 20 тыс. образов в собственной памяти и более 1 млн. – во внешней базе данных. Кроме того, представлена возможность портирования образов на мобильное устройство (смартфон/планшет) или смарт-карту типа DESFire EV1/EV2. Регистрация и идентификация пользователей может проводиться как по одному глазу, так и по двум. Пропускная способность – 20 человек в минуту. Используемые интерфейсы: Wiegand, F/2/F, OSDP, PAC, Ethernet. На рынке данный терминал появился в 2016 году, в настоящее время цена на него составляет порядка 310 тыс. рублей [9].

Рис. 1. Внешний вид контроллеров СКУД

Таблица 1

Сравнительная характеристика контроллеров СКУД

Модель	Факторы идент-ии	Интер-фейсы	Объем встроенной памяти			Г/в	Цена т.р.
			Карты досту-па	Био-метр. образы	Собы-тия		
ZKTeco MA300	Карта Отпе-чаток	Ethernet	1000	1500	100 тыс.	2013	10
BioSmart-WTC-2-EM	Карта Отпе-чаток Пин-код	Ethernet USB	3000	4500	100 тыс.	2014	35
Suprema FS2-D	Карта Рисунок лица	Wi-Fi Ethernet RS485 USB	3000	3000.	5 млн.	2017	85
EyeLock NANO NXT	Сетчатка глаза	ETHERNET	–	20 тыс.	5 тыс.	2016	310

Отдельного внимания заслуживают перспективные биометрико-электронные технологии идентификации личности. Так, британский банк Natwest тестирует систему, которая позволяет отказаться от PIN-кода в дебетовой карте, когда данные об отпечатке пальца заносятся в память карты и при совершении платежа запрашиваются системой. Данное решение не только сочетает в себе удобство бесконтактной оплаты с безопасностью подтверждаемых платежей, но и не позволяет украсть биометрические данные пользователя без доступа к карте, т.к. они не хранятся во внешней базе данных [10]. Как для систем с биометрической идентификацией, так и для комбинированных систем характерны общие недостатки,

выражающиеся в увеличении числа ошибок первого рода при повреждении или искажении биометрических данных человека (заболевание глаз, порезы и потертости пальцев, засвет лучей сканеров рисунков вен и т. д.). Кроме того, для них характерны повышенные требования к безопасности информации, вызванные высокой опасностью компрометации (с последующим воспроизведением) данных, т.к. в данном случае невозможно просто «заменить» идентификатор на безопасный новый, в качестве которого выступает один из биометрических параметров человека (в отличие от классических СКУД с различными вариантами электронных идентификаторов). Также, согласно ФЗ РФ № 152-ФЗ «О персональных данных» от 27.07.2006 г., большинство биометрических параметров человека являются персональными данными, что требует заключения с каждым пользователем согласия на обработку персональных данных, а также их хранения внутри системы и ее частей с применением технологий обезличивания данных [11]. Немаловажной особенностью является и требовательность большинства способов идентификации к объему обучающих выборок, что подразумевает собой необходимость использования высококачественных методов размножения биометрических образов [12]. Однако несмотря на то, что вышеописанные особенности применения и эксплуатации СКУД с многофакторной идентификацией личности, контролирующих один или несколько биометрических параметров человека, зачастую могут увеличивать стоимость разработки, внедрения и работы таких систем, а в отдельных случаях – ограничивать зону возможного применения, применение подобных решений позволяет в разы повысить надежность систем безопасности различного уровня. А разнообразие рынка таких систем и все большее их распространение говорит о перспективности данного направления развития СКУД и смежных интеллектуальных систем. Данная статья подготовлена в рамках Договора №23/19 НИР на выполнение научно-исследовательской работы от 09.04.19 г. по программе «Ректорские гранты», реализуемой в ФГБОУ ВО «Пензенский государственный университет».

Библиографический список

1. «Сбербанк» узнает тебя в лицо // habr.ru: Портал с техническими статьями. – 2016. – URL: https://habr.com/ru/news/t/369181/

2. Бекмурзин, М. С. Биометрические технологии в антитеррористической деятельности правоохранительных органов: перспективы и проблемы

использования / М. С. Бекмурзин, В. П. Захаров, О. И. Зачек // Вестник Московского университета МВД России. – 2014. – № 10. – С. 44–49.

3. Современные биометрические методы идентификации // habr.ru: Портал с техническими статьями. 2016. – URL: https://habr.com/ru/post/126144/

4. Roy, A. MasterPrint: Exploring the Vulnerability of Partial Fingerprint-Based Authentication Systems / Aditi Roy, Student Member, Nasir Memon // IEEE transactions on information forensics and security. – 2017. – № 9. – С. 2013–2025.

5. Быков, О. Слабое звено в банковских системах контроля доступа / О. Быков // Алгоритм безопасности. – 2015. – № 3. – С. 57–58.

6. ZKTECO. Биометрия и безопасность. Производитель оборудования и поставщик готовых решений в сфере обеспечения безопасности, контроля доступа и контроля рабочего времени. – 2019. – URL: https://www.zkteco.ru/product_detail/MA300.html

7. Биометрические системы контроля доступа BioSmart. – 2019. – URL: https://www.tinko.ru/catalog/product/240534/

8. Биометрическое оборудование для систем контроля доступа и учета рабочего времени. 2019. – URL: https://www.supremainc.ru/products/biometricheskoe-oborudovanie/facestation-2/

9. ВЗОР. Технолоджи-услуги по удаленной идентификации личности, а также системы решения в области управления доступом, мониторинга рабочего времени и безопасности, 2019. – URL: http://vzortechnology.com/produkt-nano-nxt/

10. Porter, J. Debit card with built-in fingerprint reader begins trial in the UK / J. Porter // The Verge. – 2019. – 11 марта.

11. О персональных данных : федер. закон № 152-ФЗ от 27.07.2006 : [в ред. от 31.12.2017]. – URL: http://www.consultant.ru/cons/cgi/online.cgi?req=doc&base=LAW&n=286959&fld=134&dst=1000000001,0&rnd=0.40222648846555065#07030617617285386

12. Быстрый алгоритм обучения больших сетей искусственных нейронов квадрата среднего геометрического плотностей распределения значений многомерных биометрических данных / В. И. Волчихин, А. И. Иванов, К. А. Перфилов, Е. А. Малыгина, Ю. И. Серикова // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2018. – № 3 (47). – С. 23–35.

Назаров, В. Л. Современные системы контроля и управления доступом на основе технологий биометрической и электронной идентификации личности / В. Л. Назаров, Д. А. Данилин, М. Д. Суворов, А. А. Устинов // Безопасность информационных технологий : тр. I Всерос. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2019. – С. 156–162.

УДК 519.2, 612.087, 621.319.7

А. И. Иванов, А. Г. Банных, Е. Н. Куприянов,
В. С. Лукин, К. А. Перфилов, К. Н. Савинов

КОЛЛЕКЦИЯ ИСКУССТВЕННЫХ НЕЙРОНОВ, ЭКВИВАЛЕНТНЫХ СТАТИСТИЧЕСКИМ КРИТЕРИЯМ, ДЛЯ ИХ СОВМЕСТНОГО ПРИМЕНЕНИЯ ПРИ ПРОВЕРКЕ ГИПОТЕЗЫ НОРМАЛЬНОСТИ МАЛЫХ ВЫБОРОК БИОМЕТРИЧЕСКИХ ДАННЫХ

Аннотация. Целью работы является снижение вероятности ошибок при проверке гипотезы нормальности распределения малой выборки биометрических данных за счет нейросетевого обобщения нескольких старых и новых статистических критериев. Используются давно известные статистические критерии и новые статистические критерии: среднего геометрического, среднего гармонического, эксцесса. Показано, что эти критерии дополняют друг друга и могут использоваться совместно. Для всех критериев построены эквивалентные им искусственные нейроны, что позволяет их совместно применять в форме сети из 12 нейронов с 4096 дискретными выходными кодовыми состояниями. Показано, что статистические решения рассмотренных нейронов имеют существенную независимую компоненту, что позволяет эффективно их совместно использовать. Оценены ожидаемые вероятности нейросетевого обобщения рассмотренных статистических критериев.

A. I. Ivanov, A. G. Bannich, E. N. Kupriyanov,
V. S. Lukin, K. A. Perfilov, K.N. Savinov

COLLECTION OF ARTIFICIAL NEURONS EQUIVALENT STATISTICAL CRITERIA FOR THEIR USE WHEN TESTING THE HYPOTHESIS OF NORMALITY OF SMALL SAMPLES OF BIOMETRIC DATA

Abstract. The aim of the research is to reduce the likelihood of errors when checking the hypothesis of normality distribution of small sample biometric data by neural network synthesis of several old and new statistical criteria.

Used for a long time known statistical criteria and new statistical criteria: the geometric mean, Harmonic, kurtosis. It is shown that these criteria are complementary and can be used together. For all benchmarks built their equivalent artificial neurons, allowing them to jointly apply in the form of a network of 12 neurons with 4096 discrete outputs code States.

It is shown that the statistical decisions have significant independent neurons reviewed component that allows you to share them effectively. Estimated expected probability neural network generalization reviewed statistical criteria.

Введение

Экспериментально полученных данных всегда мало. Особенно остро эта проблема стоит в биометрии. Пользователи биометрии рассматривают средство как дружественное, если оно обучается на 20 примерах БиоОбраза «Свой». На такой маленькой выборке нельзя стандартными методами оценить нормальность обучающей выборки.

Если пользоваться стандартными рекомендациями проверки гипотезы по критерию хи-квадрат требуется выборка в 200 и более опытов [1]. Заставлять пользователе предъявлять 200 примеров свого биометрического образа нерационально. Более рациональным является повысить мощность статистических оценок за счет нейросетевого обобщения нескольких статистических критериев.

Для примера построим нейрон, эквивалентный хи-квадрат критерию, для выборки в 21 опыт. Для этой цели будем воспроизводить хи-квадрат критерий для нормальных и равномерно распределенных данных. Результат численного эксперимента отображен на рис. 1.

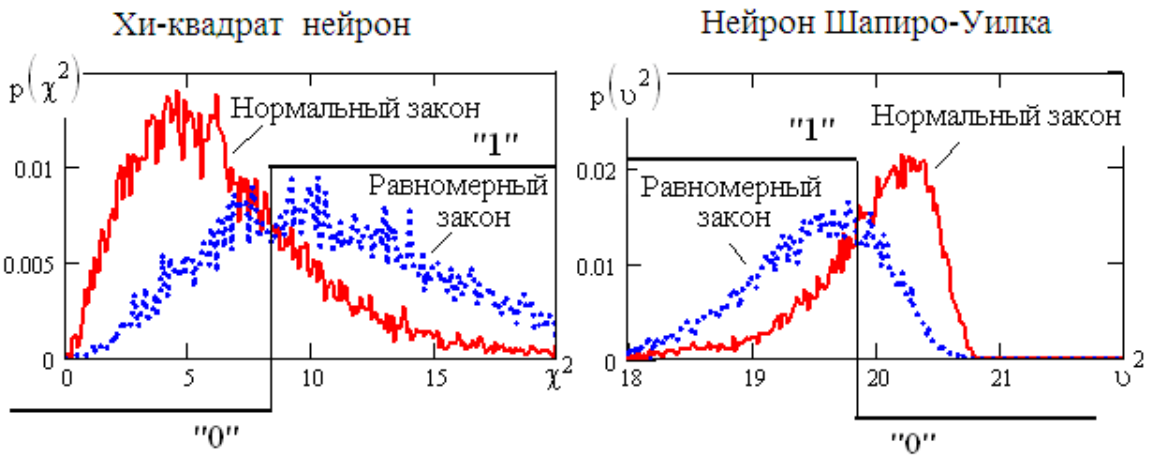


Рис. 1. Результаты численного моделирования хи-квадрат критерия и критерия Шапиро–Уилка

Из левой части рис. 1 видно, что статистики нормальных данных и равномерных данных достаточно хорошо различимы при

применении хи-квадрат критерия. Это означает, что мы можем построить нейрон, воспроизводящий работу хи-квадрат критерия. Функциональные связи такого нейрона описываются системой в табл. 1.

Таблица 1

Хи-квадрат нейрон

$\begin{cases} x \leftarrow \text{sort}(x) \\ \Delta \leftarrow \dfrac{x_{20} - x_0}{5} \\ \tilde{x}_i \leftarrow x_0 + \Delta \cdot i, \ i = 0, \ 1, ..., 5 \\ \chi^2 \leftarrow 21 \cdot \sum_{i=0}^{4} \dfrac{\left(\dfrac{n_i}{21} - (P(\tilde{x}_{i+1}) - P(\tilde{x}_i))\right)^2}{P(\tilde{x}_{i+1}) - P(\tilde{x}_i)}, \\ z(\chi^2) \leftarrow "0" \ \text{if} \ \chi^2 \le 7.72 \\ z(\chi^2) \leftarrow "1" \ \text{if} \ \chi^2 > 7.72 \\ P_1 \approx P_2 \approx P_{EE} \approx 0.328 \end{cases}$	где Δ – ширина интервала гистограммы; n_i – число отсчетов, попавших i-й интервал гистограммы; $P(\tilde{x}_i)$ – теоретическая вероятность для нормального распределения; P_1 – вероятность ошибки первого рода; P_2 – вероятность ошибки второго рода

В правой части рис. 1 даются данные численного эксперимента моделирования критерия Шапиро–Уилка для малой выборки из 21 примера. Поведение критерия Шапиро–Уилка может быть воспроизведено нейроном, чьи функциональные связи описываются системой, размещенной в табл. 2.

Таблица 2

Нейрон Шапиро – Уилка

$\begin{cases} x \leftarrow \text{sort}(x) \\ \upsilon^2 \leftarrow \dfrac{1}{\sigma(x)} \cdot \left\{ \sum_{i=0}^{9} a_i \cdot (x_{20-i} - x_i) \right\}^2 \\ z(\upsilon^2) \leftarrow "0" \ \text{if} \ \upsilon^2 \ge 19.2 \\ z(\upsilon^2) \leftarrow "1" \ \text{if} \ \upsilon^2 < 19.2 \\ P_1 \approx P_2 \approx P_{EE} \approx 0.301 \end{cases}$	где a_i – коэффициенты Шапиро-Уилка ($a_0 = 0.4834$, $a_1 = 0.3185$, $a_2 = 0.2575$, $a_3 = 0.2119$, $a_4 = 0.1736$, $a_5 = 0.1399$, $a_6 = 0.1039$, $a_7 = 0.0804$, $a_8 = 0.0530$, $a_9 = 0.0263$), $\sigma(x)$ – стандартное отклонение малой выборки в 21 опыт

Повторим численный эксперимент для критериев Крамера – фон Мизеса, его результаты отображены на рис. 2–6, и в табл. 3–13.

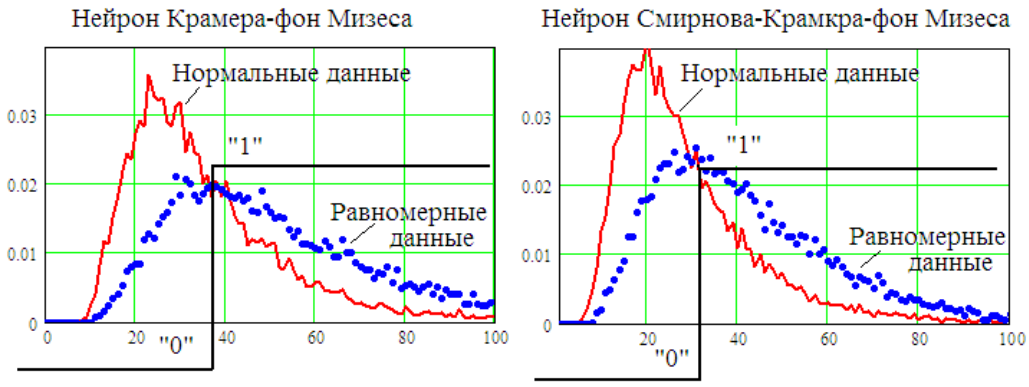


Рис. 2. Численное моделирование критериев Крамера – фон Мизеса

Таблица 3

Нейрон Крамера – фон Мизеса

$x \leftarrow \mathrm{sort}(x)$ $\omega^2 \leftarrow \sum_{i=0}^{19} (i+1-21\cdot P(x_i))^2 \cdot \frac{21\cdot(x_{i+1}-x_i)}{x_{20}-x_0}$ $z(\omega^2) \leftarrow$ "0" if $\omega^2 \le 38.52$ $z(\omega^2) \leftarrow$ "1" if $\omega^2 > 38.52$ $P_1 \approx P_2 \approx P_{EE} \approx 0.359$	Примечание: $\mathrm{corr}(\omega^2, \chi^2) = 0.556$ $\mathrm{corr}(\omega^2, \upsilon^2) = -0.667$

Таблица 4

Нейрон Смирнова – Крамера – фон Мизеса

$x \leftarrow \mathrm{sort}(x)$ $\omega_c^{\,2} \leftarrow \sum_{i=0}^{20} (i+1-21\cdot P(x_i))^2$ $z(\omega_c^{\,2}) \leftarrow$ "0" if $\omega_c^{\,2} \le 31.6$ $z(\omega_c^{\,2}) \leftarrow$ "1" if $\omega_c^{\,2} > 31.6$ $P_1 \approx P_2 \approx P_{EE} \approx 0.337$	Примечание: $\mathrm{corr}(\omega^2, \upsilon^2) = 0.885$, нет полной коррелированности

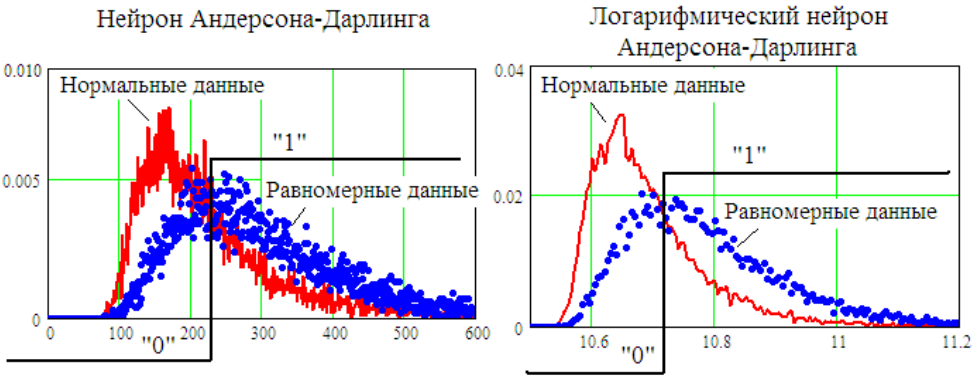


Рис. 3. Численное моделирование двух вариантов критерия Андерсона – Дарлинга

Таблица 5

Нейрон Андерсона – Дарлинга

$\begin{cases} x \leftarrow sort(x) \\ a^2 \leftarrow \sum_{i=0}^{19} \cdot \frac{(i+1-21\cdot P(x_i))^2}{P(x_i)\cdot(1-P(x_i))} \\ z(a^2) \leftarrow "0" \text{ if } a^2 \le 241 \\ z(a^2) \leftarrow "1" \text{ if } a^2 > 241 \\ P_1 \approx P_2 \approx P_{EE} \approx 0.336 \end{cases}$	Примечание: $corr(a^2, \omega_c{}^2) = 0.393$ $corr(a^2, \chi^2) = 0.424$

Таблица 6

Логарифмический нейрон Андерсона – Дарлинга

$\begin{cases} x \leftarrow sort(x) \\ al^2 \leftarrow -\sum_{i=0}^{20} \left[\frac{i+1}{21} \cdot \log(P(x_i)) - \left(1 - \frac{i+1}{21}\right) \cdot \log(1-P(x_i)) \right] \\ z(al^2) \leftarrow "0" \text{ if } al^2 \le 31.6 \\ z(al^2) \leftarrow "1" \text{ if } al^2 > 31.6 \\ P_1 \approx P_2 \approx P_{EE} \approx 0.321 \end{cases}$	Примечание: $corr(a^2, al^2) = 0.644$

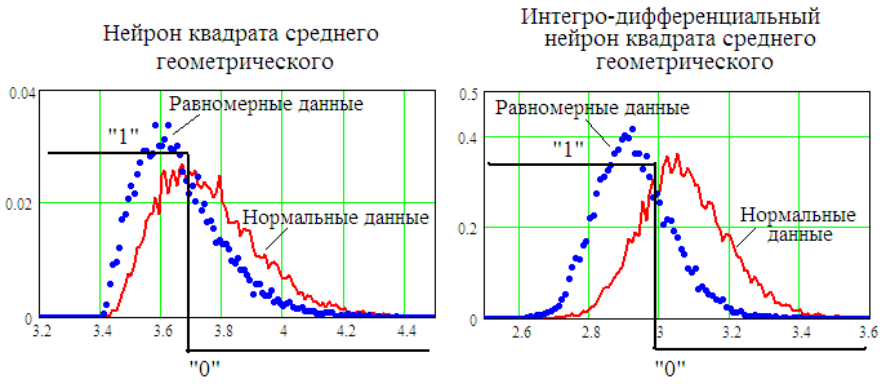


Рис. 4. Моделирование нового семейства статистических критериев квадрата среднего геометрического

Таблица 7

Нейрон квадрата среднего геометрического

$\begin{cases} x \leftarrow sort(x) \\ s^2 \leftarrow \sum_{i=0}^{20} \left(\sqrt{\frac{(i+1)\cdot(1-P(x_i))}{21}} \right)^2 \\ z(s^2) \leftarrow "0" \text{ if } a^2 \ge 3.67 \\ z(s^2) \leftarrow "1" \text{ if } a^2 < 3.67 \\ P_1 \approx P_2 \approx P_{EE} \approx 0.382 \end{cases}$	Примечание: $corr(s^2, ds^2) = 0.132$, очень низкий уровень коррелированности

Таблица 8

Нейрон интегродифференциального квадрата среднего геометрического

$\begin{cases} x \leftarrow \text{sort}(x) \\ ds^2 \leftarrow \sum_{i=0}^{20} \left(\sqrt{\frac{(i+1)\cdot p(x_i)}{21}} \right)^2 \\ z(ds^2) \leftarrow "0" \text{ if } ds^2 \geq 2.99 \\ z(ds^2) \leftarrow "1" \text{ if } ds^2 < 2.99 \\ P_1 \approx P_2 \approx P_{EE} \approx 0.265 \end{cases}$,	где $p(x) = \frac{\partial P(x)}{\partial x}$ – плотность распределения значений нормального закона [4]

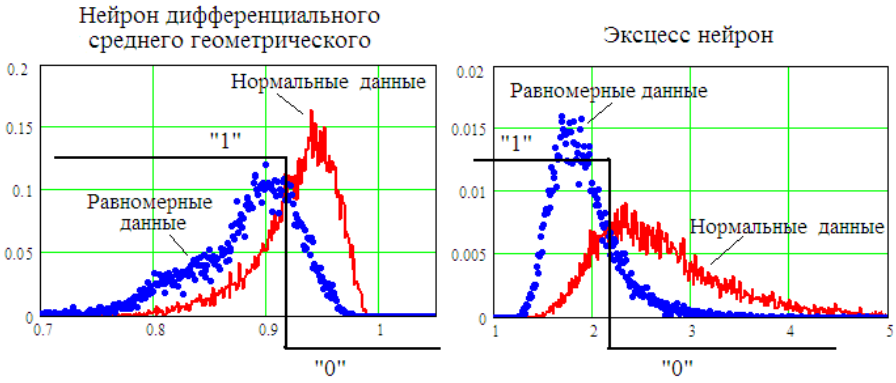


Рис. 5. Моделирование нового критерия дифференциального среднего геометрического и нового эксцесс-критерия

Таблица 9

Нейрон дифференциального среднего геометрического

$\begin{cases} x \leftarrow \text{sort}(x) \\ \Delta \leftarrow \frac{x_{20} - x_0}{7} \\ \tilde{x}_i \leftarrow x_0 + \Delta \cdot i, \ i = 0, 1, ..., 7 \\ \partial s \leftarrow \sum_{i=0}^{6} \sqrt{\frac{n_i}{21} \cdot (P(\tilde{x}_{i+1}) - P(\tilde{x}_i))} \\ z(\partial s) \leftarrow "0" \text{ if } \partial s \geq 0.912 \\ z(\partial s) \leftarrow "1" \text{ if } \partial s < 0.912 \\ P_1 \approx P_2 \approx P_{EE} \approx 0.336 \end{cases}$	Примечание: $\text{corr}(s^2, ds^2) = 0.132$ $\text{corr}(s^2, \omega^2) = 0.390$ $\text{corr}(\partial s, s^2) = 0.324$

Таблица 10

Эксцесс нейрон

$\begin{cases} x \leftarrow \text{sort}(x) \\ \mu^4 \leftarrow \sum_{i=0}^{20} \frac{(x_i - E(x))^4}{21 \cdot (\sigma(x))^4} \\ z(\mu^4) \leftarrow "0" \text{ if } \mu^4 \geq 2.14 \\ z(\mu^4) \leftarrow "1" \text{ if } \mu^4 < 2.14 \\ P_1 \approx P_2 \approx P_{EE} \approx 0.197 \end{cases}$	Примечание: $\text{corr}(\partial s, \mu^4) = 0.142$

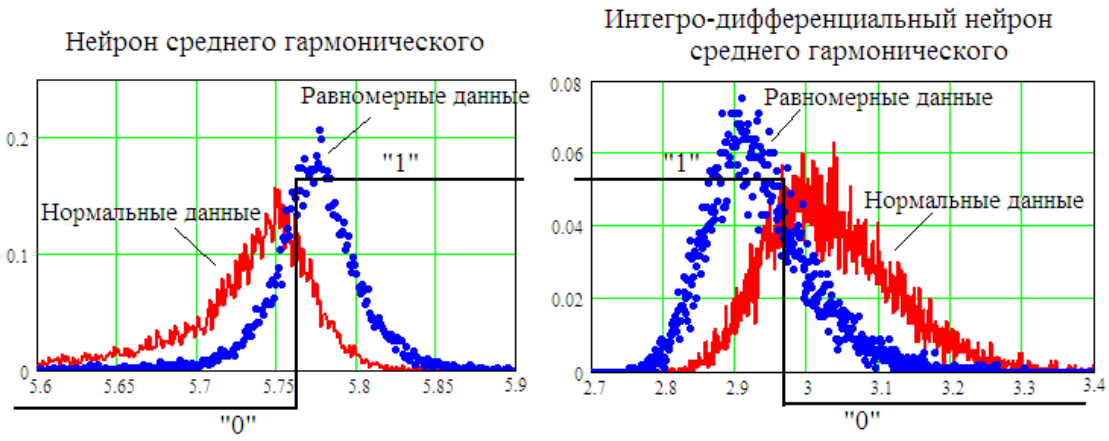


Рис. 6. Моделирование нового семейства критериев среднего гармонического

Таблица 11

Нейрон среднего гармонического

$\begin{cases} x \leftarrow \text{sort}(x) \\ g^2 \leftarrow \sum_{i=0}^{20} \frac{(i+1)\cdot(1-P(x_i))}{i+1+21\cdot(1-P(x_i))} \\ z(g^2) \leftarrow \text{"0" if } g^2 \le 5.69 \\ z(g^2) \leftarrow \text{"1" if } g^2 > 5.69 \\ P_1 \approx P_2 \approx P_{EE} \approx 0.249 \end{cases}$	Примечание: $\text{corr}(g^2, dg^2) = -0.222$

Таблица 12

Интегро-дифференциальный нейрон среднего гармонического

$\begin{cases} x \leftarrow \text{sort}(x) \\ dg^2 \leftarrow \sum_{i=0}^{20} \frac{(i+1)\cdot p(x_i)}{i+1+21\cdot(1-P(x_i))} \\ z(dg^2) \leftarrow \text{"0" if } dg^2 \ge 2.96 \\ z(dg^2) \leftarrow \text{"1" if } dg^2 < 2.96 \\ P_1 \approx P_2 \approx P_{EE} \approx 0.262 \end{cases}$	Примечание: $\text{corr}(dg^2, s^2) = 0.903$ $\text{corr}(dg^2, ds^2) = 0.414$

Нейросетевое обобщение нескольких критериев выполняется объединением нескольких нейронов в сеть, как это показано на рис. 7.

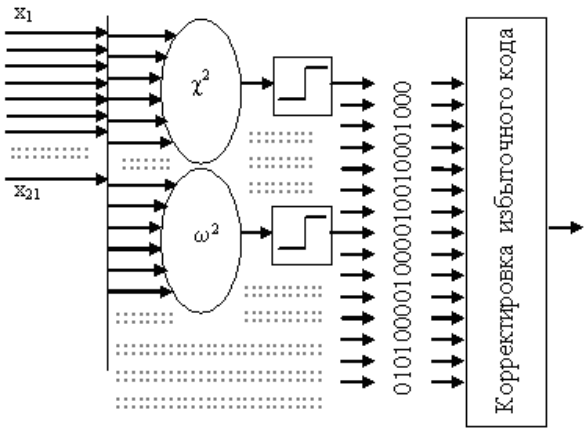


Рис. 7. Нейросетевое обобщение разнородных статистических критериев при анализе малых выборок биометрических данных

Таблица 13

Матрица корреляционной связанности 8, рассмотренных выше статистических критериев

	χ^2	a^2	al	s^2	ds^2	ω^2	ω^2_c	υ^2
χ^2	**1**	0.423	0.672	0.037	–0.042	0.559	0.401	–0.726
a^2	0.423	**1**	0.644	0.018	–0.145	0.226	0.393	–0.113
al	0.672	0.644	**1**	0.056	0.209	0.827	0.832	–0.917
s^2	0.037	0.018	0.056	**1**	0.132	0.414	0.402	–0.212
ds^2	–0.042	–0.145	0.209	0.132	**1**	–0.242	–0.142	–0.041
ω^2	0.559	0.226	0.827	0.414	–0.242	**1**	0.885	–0.667
ω^2_c	0.401	0.393	0.832	0.402	–0.142	0.885	**1**	–0.764
υ^2	–0.726	–0.113	–0.917	–0.212	–0.041	–0.667	–0.764	**1**

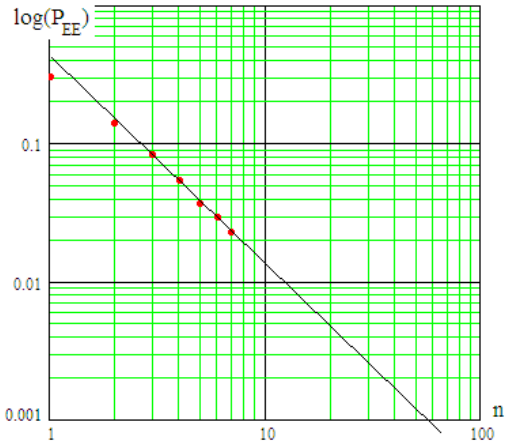


Рис. 8. Моделирование вероятностей ошибок первого и второго рода для 1, 2, …, 7 обобщаемых критериев (7 нейронов) при среднем значении вероятностей ошибок $E(P_{EE}) = 0.306$, среднем значении модулей корреляции $E(|corr(.,.)|) = 0.398$ (усреднение выполнено по матрице из табл. 13)

Выводы

Имитационное моделирование показало, что большинству старых статистических критериев (хи-кадрат, Крамера – фон Мизеса [2], Шапиро – Уилка, Андерсона – Дарлинга,…) удается построить эквивалентные им нейроны. Это позволяет их легко обобщать в виде нейронной сети (рис. 7). Эффективность нейросетевого обобщения зависит не столько от мощности нейрона, но и от его корреляционных связей с другими нейронами.

В этом контексте в Пензенском государственном университете предприняты попытки синтеза новых статистических критериев, которые будут слабо коррелированны между собой и слабо коррелированны со старыми статистическими критериями. В частности, было создано семейство критериев среднего геометрического [3, 4], среднего гармонического и эксцесс-критерий.

Симметризация задачи нейросетевого обобщения [5] позволяет ожидать для 12 нейронов вероятностей ошибок менее 0.01. При обобщении 60 статистических критериев вероятность ошибок должна снизиться до величины 0.001.

Библиографический список

1. Р 50.1.037–2002. Рекомендации по стандартизации. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. – Москва : Госстандарт России, 2001. – Ч. I. Критерии типа $\chi 2$. – 140 с.

2. Иванов, А. И. Закон распределения значений критерия Крамера – фон Мизеса для проверки гипотезы нормальности малых выборок / А. И. Иванов, Е. А. Малыгина, С. Е. Вятчанин, С. В. Туреев // Электронные информационные системы. – 2019. – № 1 (20). – С. 95-105.

3. Иванов, А. И. Сравнение мощности критерия среднего геометрического и Крамера – фон Мизеса на малых выборках биометрических данных / А. И. Иванов, Е. А. Малыгина, П. А. Перфилов, С. Е. Вятчанин // Модели, системы, сети в экономике, технике, природе и обществе. – 2016. – № 2. – С. 155–158.

4. Иванов, А. И. Оценка качества малых выборок биометрических данных с использованием дифференциального варианта статистического критерия среднего геометрического / А. И. Иванов, П. А. Перфилов, Е. А. Малыгина // Вестник СИБГАУ. – 2016. – № 4 (17). – С. 864–871.

5. Ivanov, A. I. Simplification of Statistical Description of Quantum Entanglement of Multidimensional Biometric Data Using Simmetrization of Paired Cor-

relation Matrices / A. I. Ivanov, A. V. Bezyaev, A. I. Gazin // Journal of Computational and Engineering Mathematics. – 2017. – Vol 4, № 2. – С. 3–13. – URL: http://jcem.susu.ru/jcem/issue/view/14

Иванов, А. И. Коллекция искусственных нейронов, эквивалентных статистическим критериям, для их совместного применения при проверке гипотезы нормальности малых выборок биометрических данных / А. И. Иванов, А. Г. Банных, Е. Н. Куприянов, В. С. Лукин, К. А. Перфилов, К. Н. Савинов // Безопасность информационных технологий : тр. I Всерос. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2019. – С. 163–172.

УДК 519.2, 612.087, 621.319.7

**А. И. Иванов, А. В. Безяев, Е. А. Малыгина,
Ю. И. Серикова**

ВТОРОЙ НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИИ ПО БЫСТРОМУ АВТОМАТИЧЕСКОМУ ОБУЧЕНИЮ БОЛЬШИХ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ НА МАЛЫХ ВЫБОРКАХ БИОМЕТРИЧЕСКИХ ДАННЫХ

Аннотация. Целью работы является пояснение причин необходимости создания второго национального стандарта России по автоматическому обучению больших искусственных нейронных сетей. Ранее был создан и введен в действие первый национальный стандарт ГОСТ Р 52633.5–2011 по автоматическому обучению персептронов. Кроме того, была разработана техническая спецификация по применению криптографических механизмов для защиты таблиц связей и весовых коэффициентов сетей персептронов. Криптоанализ показал ряд уязвимостей сетей персептронов, которые устраняются переходом к использованию квадратичных нейронов с многоуровневыми квантователями. Мощность квадратичных нейронов в метрике равновероятного значения ошибок первого и второго рода много выше, чем мощность нейронов с накоплением (обогащением) входных данных в линейном пространстве. Однако обычные бинарные квадратичные нейроны не пригодны для применения в биометрии, так как не способны хэшировать (перемешивать) данные образов «Чужой».

**A. I. Ivanov, A. V. Bezyaev, E. A. Malygina,
Y. I. Serikova**

THE SECOND NATIONAL STANDARD OF RUSSIA'S RAPID AUTOMATIC LEARNING OF LARGE ARTIFICIAL NEURAL NETWORKS ON SMALL SAMPLES OF BIOMETRIC DATA

Abstract. The aim of the work is the explanation of the reasons, the need to create a second Russian national standard for automatic learning of large artificial neural networks.

Previously created and launched the first national standard GOST r 52633.5-2011 automatic learning perseptronov. In addition, technical specification has been developed on the use of cryptographic mechanisms to protect links and tables of weights perseptronov networks. Cryptanalysis showed a series of vulnerabilities networks perseptronov that the transition to the use of quadratic neurons with multilevel kvantovateljami.

Power of quadratic neurons in metric linear error values of the first and second kind a lot higher than power stacked neurons (enrichment) of input data in linear space. However, the usual binary quadratic neurons are unsuitable for use in biometrics, as not capable of hashing (stir) image data "Alien".

Введение

Цифровая экономика оказывается устойчивой, если действия людей в Интернет «облаках» криптографически защищены. К сожалению, обычные люди не могут запоминать длинные пароли доступа из случайных символов. Как следствие доверие к хранению личной информации в Интернет облаках низкое.

В США, Канаде и странах Евросоюза эту проблему пытаются решить с использованием, так называемых, «нечетких экстракторов» [1–3], преобразующих длинный неоднозначный БиоКод в короткий, но стабильный код криптографического ключа. На рис. 1 иллюстрируется идея, положенная в основу работы «нечетких экстракторов».

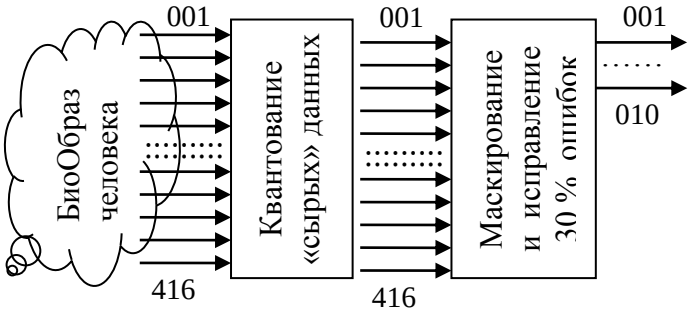


Рис. 1. Блок-схема «нечетких экстракторов», использующих классические избыточные коды с обнаружением и исправлением ошибок

Нечеткий биометрический образ преобразуется в контролируемые «сырые» биометрические параметры и сравниваются с порогом среднего значения. Параметр примера БиоОбраза больше среднего дает состояние «1», параметр меньше среднего дает состоянии «0». Получившийся БиоКод имеет примерно 30 % ошибок. Самые не стабильные разряды кода накрываются маской и становятся стабильными. Остальные разряды накрывают исправляющим ошибки кодом с 20-кратной избыточностью. Длина выходного кода криптографического ключа более чем в 20 раз меньше, чем длина корректируемого БиоКода.

Использование нейросетевых преобразователей биометрия-код

Использовать классические коды с обнаружением и исправлением ошибок не рационально из-за того, что исправленный код криптографического ключа оказывается коротким. Гораздо более рациональным для обогащения данных является использование искусственных нейронов. Эта идея положена в основу серии национальных стандартов России с номерами ГОСТ Р 52633.xx-20xx.

Это техническое решение позволяет получать коды криптографического ключа любой длины. Все зависит от числа нейронов в нейронной сети, каждый нейрон отвечает за один бит криптографического ключа. Обучение нейронов ведется только в автоматическом режиме алгоритмом ГОСТ Р 52633.5–2011 [4].

На рис. 1 указана длина БиоКода – 416 и длина вектора биометрических параметров БиоОбраза-416. Это обусловлено тем, что достоверных биометрических данных достаточного объема в Интернете нет. Единственным исключением являются данные среды моделирования «БиоНейроАвтограф» [5], скачав из Интернет, эту среду любой пользователь может получить как угодно большую базу примеров БиоОбразов. Для этого необходимо манипулятором «мышь» или через чувствительный экран компьютера вводить рукописные образы. Среда моделирования выполняет двухмерное преобразование Фурье и формирует вектор из 416 контролируемых БиоПараметров. Далее сеть из 256 нейронов практически полностью устраняет естественные вариации БиоОбраза «Свой», получая с вероятностью 0.97 код заданного при обучении криптографического ключа.

Симметрия, присутствующая у нейронных сетей, обученных по ГОСТ Р 52633.5

Особенностью алгоритма обучения ГОСТ Р 52633.5–2011 является то, что для образов «Свой» устраняется неопределенность выходного кода. Для образов «Чужой» естественная неопределенность усиливается, происходит хэширование (перемешивание) данных образов «Чужой». Небольшие изменения данных «Чужой» (полученные, например, добавлением малого шума) приводят к очень большим изменениям выходного кода. Эта ситуация иллюстрируется рис. 2.

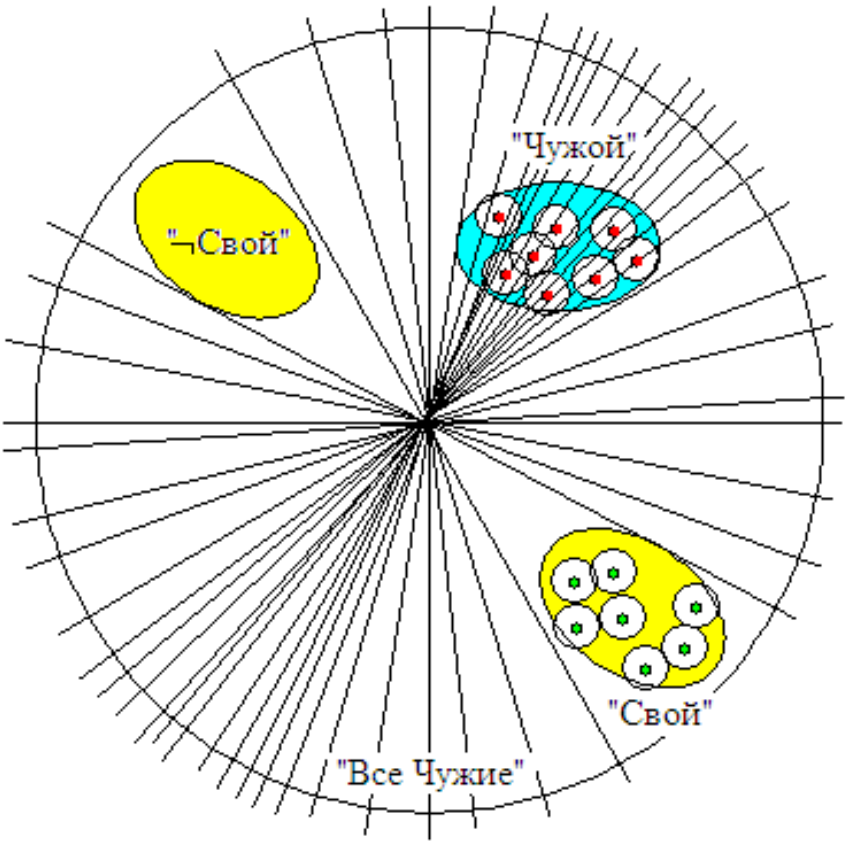


Рис. 2. Особенности топологии нейронных сетей,
обученных алгоритмом ГОСТ Р 52633.5

Все разделяющие гиперплоскости обученных нейронов проходят через цент множества «Все Чужие». Это приводит к тому, что «Чужой», пытающийся подобрать образ «Свой» видит равновероятные состояния «0» и «1» для каждого из разрядов выходного кода. К сожалению, выполнение этого требования приводит к появлению двух многомерных пирамид стабильности. Одна пирамида своим объемом, охватывает распределение образов «Свой», а вторая многомерная пирамида своим объемом охватывает инверсию данных образа «Свой». Данные инверсного образа «Свой» дают на выходе обученной сети нейронов инверсный очень стабильный выходной код образа «Свой».

Из-за того, что ни одна из линий разделяющих функций нейронов не пересекает эллипс «Свой» добавление небольшого шума не меняет выходного кода. Энтропия кодов «Свой» близка к нулю.

Совершенно иная ситуация возникает при добавлении малого шума к образу «Чужой». В этом случае разряды кода оказываются нестабильны с высокой вероятностью, так как эллипс «Чужой»

многократно пересекается разделительными линиями нейронов сети преобразователя.

Машина по извлечению знаний из однослойной сети персептронов

Сети с искусственными нейронами, осуществляющими накопление данных в линейном пространстве, много лучше чем «нечеткие экстракторы». Длина выходного кода в место 10 бит для «нечетких экстракторов» имеет длину криптографического ключа в 256 бит. Однако разряды выходного 256 битного ключа не являются независимыми.

Пользуясь зависимостью выходных разрядов нейросети и ее симметрий (см. рис. 2.) в 2009 году была создана машина по извлечению знаний из таблиц связей нейронов и таблиц весовых коэффициентов обученной нейронной сети. Для противодействия созданной в 2009 г. машине извлечения знаний необходимо шифровать таблицы связей и таблицы весовых коэффициентов. Шифрование таблиц регламентируется технической спецификацией [6]. При применении спецификации (расчетная) длина ключа с независимыми разрядами снижается с 256 бит до 30 бит, что в 3 раза больше чем длина ключа «нечеткого экстрактора». Однако при такой защите машина извлечения знаний [7] уже не способна выполнять свою работу.

Проблемы квадратичных нейронов с бинарным квантованием

Оценка в 30 бит эквивалентного выходного ключа нейросетевого преобразователя биометрических данных построена на том, что по требованиям спецификации [6] необходимо использовать нейроны без общих входных связей. То есть играет роль число входов у нейронов. Чем меньше входов у каждого нейрона, тем больше независимых выходных бит получится при применении спецификации [6].

В этом отношении перспективным оказывается переход от использования нейронов с линейным накоплением данных, к нейронам с накопления данных в квадратичном пространстве [8].

Проблема состоит в том, что все квадратичные нейроны с бинарным квантователем не способны хэшировать (перемешивать) данные образов «Чужой». Эта ситуация иллюстрируется рис. 3.

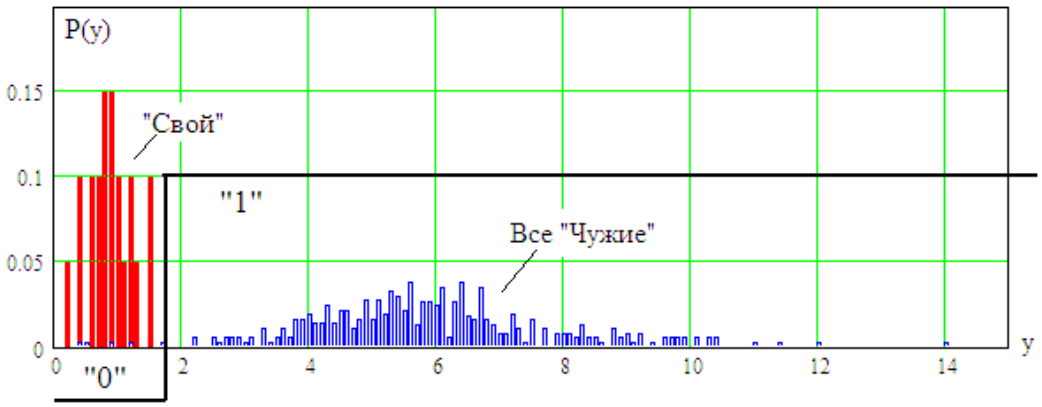


Рис. 3. Работа квадратичного нейрона с бинарным квантователем

Если мы возьмем 256 квадратичных нейронов и обучим сеть выдавать код «Свой», то любой образ «Чужой» будет давать инверсный код «Свой» с очень высокой вероятностью. Если мы добавим к образу «Чужой» небольшой шум, то выходной код меняться не будет. Замена линейных нейронов с бинарным квантованием на квадратичные нейроны с бинарными квантователями лишает нейросетевой преобразователь биометрии всякой защиты от попыток извлечения знаний. Чтобы узнать код «Свой» достаточно предъявить любой образ «Чужой» и инвертировать полученный выходной код.

Устранение проблемы квадратичных нейронов

Исследования, проведенные в «Пензенском государственном университете» [9, 10] показали, что эта проблема может быть решена, если перейти к нейронам с выходными многоуровневыми квантователями. Как итоговый результат был создан второй национальный стандарт по быстрому автоматическому обучению больших искусственных нейронных сетей [11]. Работа квадратичного нейрона с 8-уровневым квантователем иллюстрируется рис. 4.

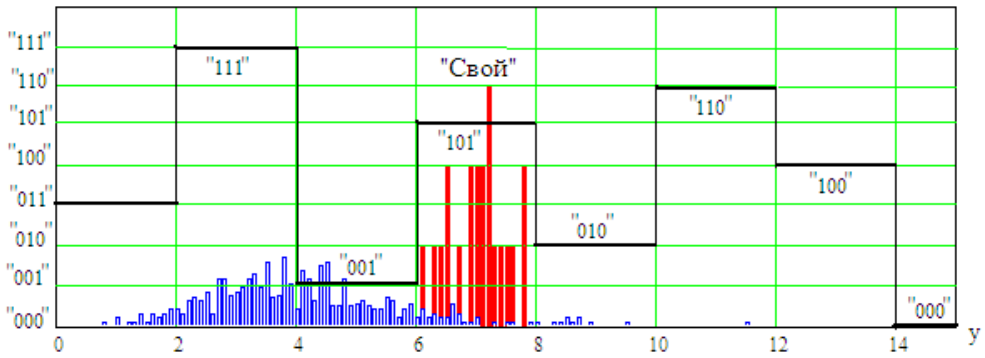


Рис. 4. Работа квадратичного нейрона с многоуровневым квантователем

Если для такой конструкции к образу «Чужой» добавить небольшой шум, то на выходе нейросетевого преобразователя возникают значительные флюктуации выходных кодовых состояний. Наблюдается эффект хэширования (перемешивания) данных «Чужой», причем этот эффект выше, чем у сетей персептронов, обученных по ГОСТ Р 52633.5.

Выводы

Имитационное моделирование показало, что мощность квадратичных нейронов в метрике равной вероятности ошибок первого и второго рода от трех до четырех раз выше мощности линейных нейронов [8]. То есть для биометрических данных одинакового качества для квадратичных нейронов потребуется в три-четыре раза меньше входов. Это означает, что при защите таблиц связей и весов шифрованием по спецификации [6] число независимых нейронов (не имеющих общих входных связей) увеличивается в три-четыре раза. Предположительно длина эквивалентного криптографического ключа по отношению к сетям персептронов вырастает с 30 бит до 90 бит или даже до 120 бит.

Очевидно, что для нового класса нейросетевых преобразователей биометрии в код с обогащением данных в квадратичном пространстве необходимо запрограммировать новую машину для извлечения из не защищенных (открытых) таблиц нейросети знаний. Только после создания и эксплуатации второй машины для извлечения знаний можно провести полноценный криптографический анализ нового класса нейросетевых преобразователей.

Следует также отметить, что стандартизованный алгоритм обучения требует маленьких обучающих выборок объемом от 12 до 20 примеров. Новый алгоритм по автоматическому обучению квадратичных нейронов сохраняет это положительное свойство своего аналога.

Библиографический список

1. Monrose, F. Cryptographic key generation from voice. In Proc. / F. Monrose, M. Reiter, Q. Li, S. Wetzel // IEEE Symp. on Security and Privacy. – Oakland, CA, USA. – 2001. – № 14–16 May. – P. 202–213.

2. Dodis, Y. Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy / Y. Dodis, L. Reyzin, A. Smith // EUROCRYPT. – 2004. – April 13. – P. 523–540.

3. Ramírez-Ruiz, J. Cryptographic Keys Generation Using FingerCodes. / J. Ramírez-Ruiz, C. Pfeiffer, J. Nolazco-Flores //Advances in Artificial Intelligence – IBERAMIA-SBIA 2006 (LNCS 4140). – 2006. – P. 178–187.

4. ГОСТ Р 52633.5–2011. Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа.

5. Иванов, А. И. Среда моделирования «БиоНейроАвтограф» / А. И. Иванов, О. С. Захаров [Программный продукт создан лабораторией биометрических и нейросетевых технологий, размещен с 2009 г. на сайте АО «ПНИЭИ»]. – URL: http://пниэи.рф/activity/science/noc/bioneuroautograph.zip

6. Техническая спецификация «Защита нейросетевых биометрических контейнеров с использованием криптографических алгоритмов» : проект [Публичное обсуждение начато с 01.02.2017 членами ТК 26 «Криптографическая защита информации», голосование по второй редакции оканчивается 01.08.19].

7. Волчихин, В. И. Нейросетевая молекула: решение обратной задачи биометрии через программную поддержку квантовой суперпозиции на выходах сети искусственных нейронов / В. И. Волчихин, А. И. Иванов // Вестник Мордовского университета. – 2017. – Т. 27, № 4. – С. 518–523.

8. Волчихин, В. И. Соотношение мощности нейронов с линейным и квадратичным обогатителями биометрических данных / В. И. Волчихин, А. И. Иванов, Е. А. Малыгина, А. П. Юнин // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2018. – № 1 (45). – С. 17–25.

9. Волчихин, В. И. Перспективы использования искусственных нейронных сетей с многоуровневыми квантователями в технологии биометрико-нейросетевой аутентификации / В. И. Волчихин, А. И. Иванов, В. А. Фунтиков, Е. А. Малыгина // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2013. – № 4 (28). – С. 88–99

10. Волчихин, В. И. Абсолютно устойчивый алгоритм автоматического обучения сетей вероятностных нейронов «Крамера–фон Мизеса» на малых выборках биометрических данных / В. И. Волчихин, А. И. Иванов, С. Е. Вятчанин, Е. А. Малыгина // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2017. – № 2 (42). – С. 55–56

11. ГОСТ Р 52633.xx-20xx. Защита информации. Техника защиты информации. Автоматическое обучение сетей квадратичных нейронов с многоуровневым квантованием биометрических данных : проект [головной разработчик ФГБОУ ВО «Пензенский государственный университет»].

Иванов, А. И. Второй национальный стандарт России по быстрому автоматическому обучению больших искусственных нейронных сетей на малых выборках биометрических данных / А. И. Иванов, А. В. Безяев, Е. А. Малыгина, Ю. И. Серикова // Безопасность информационных технологий : тр. I Всерос. науч.-техн. конф. – Пенза : Изд-во ПГУ, 2019. – С. 173–180.

СОДЕРЖАНИЕ

АО «Пензенский научно-исследовательский электротехнический институт» (АО «ПНИЭИ»)

В АО»ПНИЭИ» в настоящее время разрабатываются и серийно выпускаются комплексы и технические средства криптографической защиты информации, средства специальной связи, обеспечивающие конфиденциальность, достоверность, целостность информации при передаче ее по различным каналам связи. Активно развиваются такие направления как

– создание средств управления защищенными информационно-телекоммуникационными сетями;

– создание специальных систем передачи данных;

– создание средств электронного документооборота;

– развитие и внедрение биометрико-нейросетевых технологий.

Учеными и специалистами ПНИЭИ создаются и внедряются новые поколения аппаратуры и комплексов технических средств для обработки и защиты мультимедийной информации, передаваемой по разнородным каналам и информационно-телекоммуникационным системам связи, созданным на базе современных международных протоколов.

Акционерное общество
ПЕНЗЕНСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ ИНСТИТУТ

440000, г. Пенза, ул. Советская, 9
info@pniei.penza.ru
сайт: пниэи.рф

(841-2) 59-33-50
приемная (841-2) 59-33-35
служба маркетинга (841-2) 59-33-43

Комплекс технических средств ПОРТАЛ

КТС ПОРТАЛ предназначен для организации защищенного корпоративного (ведомственного) портала с мультимедийными сервисами, имеющего собственную логическую инфраструктуру управления, не зависящую от зарубежных ресурсов логического управления публичной сетью Интернет, и комплексно реализующего современные технологии безопасности доверенных вычислений на основе отечественной элементной базы.

В основе системных решений лежит разработка собственной облачной криптографически защищенной среды, реализующей идеологию «интернет в интернете», и предоставляющей набор как стандартных, так и узкоспециализированных веб-сервисов.

Комплекс разворачивается на базе существующих ведомственных локальных сетей и не требует установки и настройки программного обеспечения на рабочих станциях. Работа пользователей осуществляется также, как если бы они работали через Интернет, но реальный выход в глобальную сеть пользователям будет недоступен, и наоборот, доступ в ведомственную сеть со стороны открытой сети Интернет также невозможен.

КТС включает в себя серверную составляющую, аппаратные средства криптографической защиты информации, мобильное приложение для доступа с Android-устройств и программное обеспечение взаимосвязанных и объединенных между собой прикладных сервисов.

Пользователь получает доступ к защищенной электронной почте, защищенному мессенджеру мгновенных сообщений с различный устройств (смартфон, планшет, ноутбук и т.д.), организует защищенные аудио и видеоконференции между разнородными техническими средствами, ведет защищенные переговоры с помощью сервиса виртуальной АТС, получает доступ к защищенному облачному хранилищу файлов и защищенному сервису справочной информации. При этом действия пользователя мало отличаются от привычных ему действий при работе в интернете через стандартный браузер.

Акционерное общество
ПЕНЗЕНСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ ИНСТИТУТ

440000, г. Пенза, ул. Советская, 9
info@pniei.penza.ru
сайт: пниэи.рф

(841-2) 59-33-50
приемная (841-2) 59-33-35
служба маркетинга (841-2) 59-33-43

ПОРТАЛ-СЕРВЕР

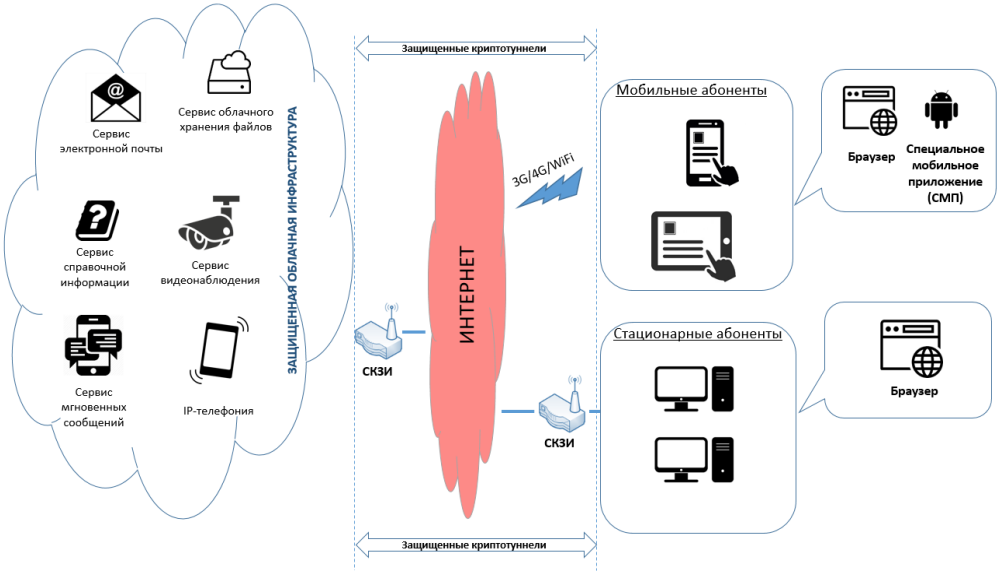
Продукт Портал-сервер представляет собой облачную криптографически замкнутую среду информационного взаимодействия и входит в линейку изделий комплекса "Портал".

Состав системных служб

- DNS-служба
- служба синхронизации времени (NTP)
- сетевая служба
- служба маршрутизации

Состав защищенных прикладных сервисов

- сервис электронной почты
- сервис обмена мгновенными сообщениями
- сервис видеоконференций
- сервис справочной информации (Вики-страницы)
- сервис облачного хранения файлов (Диск)
- видеонаблюдение
- голосовая и ведеосвязь



Логическая организация защищенной облачной системы реализована так как это делается в глобальной сети Интернет – т.е. с собственной внутренней структурой доменных имен для доступа к ресурсам. По аналогии с сетью Интернет, пользователи могут использовать облачные сервисы через привычные для них браузеры без каких-либо дополнительных условностей.

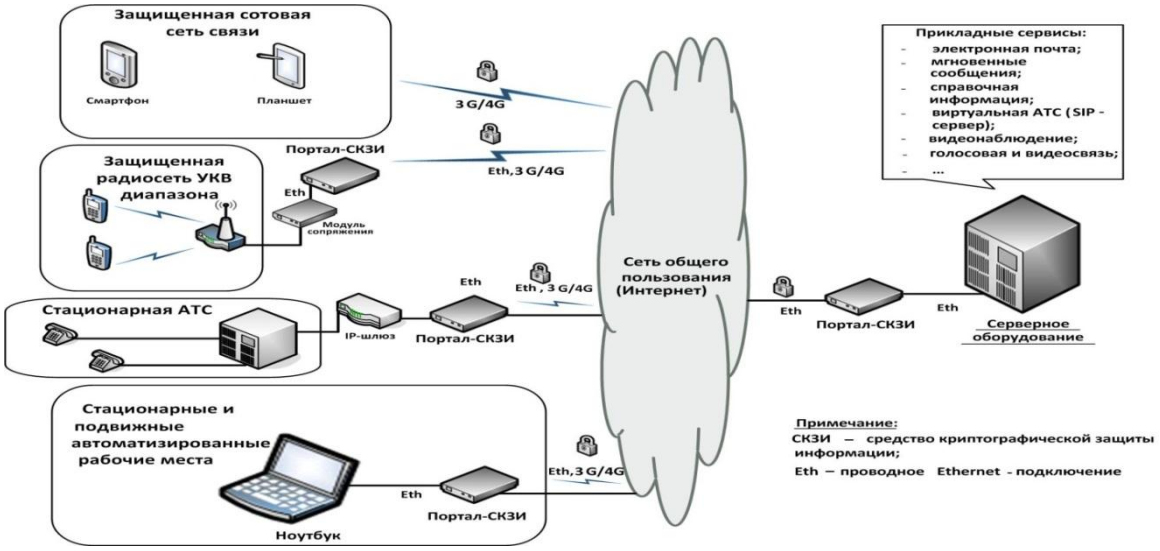
Акционерное общество
ПЕНЗЕНСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ ИНСТИТУТ

440000, г. Пенза, ул. Советская, 9
info@pniei.penza.ru
сайт: пниэи.рф

(841-2) 59-33-50
приемная (841-2) 59-33-35
служба маркетинга (841-2) 59-33-43

Доступ к информационным ресурсам может осуществляться с любого клиентского устройства сети, находящегося за СКЗИ. Разграничение доступа осуществляется с помощью механизмов аутентификации.

Также возможна организация следующей схемы связи разнородных информационных систем и отдельных технических средств в единой криптографически защищенной сервис-ориентированной среде:



Доступна организация системным администратором с помощью виртуальной АТС аудиоконференции между разнородными техническими средствами (смартфон, планшет, радиостанция, стационарный телефон, стационарное и подвижное рабочее место).

Средства криптографической защиты (Портал-СКЗИ)

Криптографическую защиту передаваемых данных в этих средах возможно осуществлять на различных уровнях стека телекоммуникационных протоколов:

- на канальном уровне (в проработке)
- на прикладном уровне (Портал-ПО, Портал-1-SD)
- на сетевом уровне (Портал-10, М-687, Швейцар-М, Портал-1000)

Акционерное общество
ПЕНЗЕНСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ ИНСТИТУТ

440000, г. Пенза, ул. Советская, 9
info@pniei.penza.ru
сайт: пниэи.рф

(841-2) 59-33-50
приемная (841-2) 59-33-35
служба маркетинга (841-2) 59-33-43

Таблица 1 – Возможные типы СКЗИ для использования в предлагаемой архитектуре

Наимено-вание	**Класс защиты**	**Скорость**	**Габариты, мм**	**Примечание**
Портал-ПО	длина ключа до 56 бит	Ограничена производи-тельностью устройства	–	Реализуется на прикладном уровне
Портал-1-SIM	длина ключа до 56 бит, для класса КС1 – в разработке	1 Мбит/с	форм-фактор SIM-карты	Совместная работа с программным СКЗИ
Портал-1-SD	длина ключа до 56 бит, для класса КС1 – в разработке		форм-фактор SD-карты, microSD-карты	Совместная работа с Портал-10, Портал-1000
Портал-10	длина ключа до 56 бит, для класса КС1 – в разработке	10 Мбит/с	165×115×25	Поддержка Wi-Fi, работа в динамических IP- адресах Совместная работа с Портал-1-SD, Портал-1000
Портал-1000	длина ключа до 56 бит, для класса КС1, КА – в разработке	600 Мбит/с	440×380×58	Совместная работа с Портал-1-SD, Портал-10, М-687А, Швейцар-М
Швейцар-М	КА, КВ	35 Мбит/с	230×165×55	Совместная работа с М-687, Портал-1000, Швейцар-М
М-687 (М-687А, М-687В)	КА, КВ, гостайна	95 Мбит/с	392x316x53	Совместная работа с Швейцар-М, Портал-1000

Акционерное общество
ПЕНЗЕНСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ ИНСТИТУТ
440000, г. Пенза, ул. Советская, 9
info@pniei.penza.ru
сайт: пниэи.рф
(841-2) 59-33-50
приемная (841-2) 59-33-35
служба маркетинга (841-2) 59-33-43

Изделие ПОРТАЛ-1-SD

Изделие является составной частью комплекса ПОРТАЛ и обеспечивает шифрование информации с длиной ключа 56 бит.

Обеспечивает

- **обмен с изделиями КТС *ПОРТАЛ***
- чтение ключевой, служебной и пользовательской информации из встроенной FLASH-памяти
- криптографическую обработку информации в соответствии с заданным алгоритмом специального преобразования

Основные технические характеристики

- максимальная тактовая частота – 20 МГц
- ряд напряжений питания – 1,8 В ± 10%; 3,0 В ± 10%; 5,0 В ± 10 %
- интерфейс ввода-вывода информации – последовательный в соответствии со стандартом SD и микроSD (SPI)
- объем встроенной FLASH-памяти – 16 Гбайт
- количество циклов стирания/записи FLASH-памяти – не менее 100 000
- время сохранности информации во FLASH-памяти – не менее 10 лет
- встроенный аппаратный ускоритель операций для криптографических алгоритмов DES
- диапазон рабочих температур: от минус 25 °С до плюс 85 °С.

Производится на базе отечественного микроконтроллера «Курган» с доверенным загрузчиком нулевого уровня.

Акционерное общество
ПЕНЗЕНСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ ИНСТИТУТ

440000, г. Пенза, ул. Советская, 9
info@pniei.penza.ru
сайт: пниэи.рф

(841-2) 59-33-50
приемная (841-2) 59-33-35
служба маркетинга (841-2) 59-33-43

Изделие ПОРТАЛ-1-SIM

Изделие является составной частью комплекса ПОРТАЛ и обеспечивает шифрование информации с длиной ключа 56 бит по интерфейсу стандарта ISO/IEC 7816-3.

Обеспечивает

- обмен с изделиями КТС *ПОРТАЛ*
- запись, хранение ключевой, служебной и пользовательской информации во встроенной FLASH-памяти
- чтение ключевой, служебной и пользовательской информации из встроенной FLASH-памяти
- криптографическую обработку информации в соответствии с заданным алгоритмом специального преобразования

Основные технические характеристики

- максимальная тактовая частота – 20 МГц
- ряд напряжений питания – 1,8 В ± 10 %; 3,0 В ± 10 %; 5,0 В ± 10 %
- интерфейс ввода-вывода информации – последовательный в соответствии со стандартом ISO/IEC 7816-3
- протокол информационно-логического взаимодействия – оригинальный на основе протокола T0 стандарта ISO/IEC 7816-3
- объем встроенной FLASH-памяти – 384 Кбайт
- количество циклов стирания/записи FLASH-памяти не менее 100 000
- время сохранности информации во FLASH-памяти не менее 10 лет
- встроенный аппаратный ускоритель операций для криптографических алгоритмов – ГОСТ 28147–89 и DES
- встроенный сопроцессор модульной арифметики
- форм-фактор – SIM
- диапазон рабочих температур: от минус 25 до плюс 85 °C

Акционерное общество
ПЕНЗЕНСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ ИНСТИТУТ

440000, г. Пенза, ул. Советская, 9
info@pniei.penza.ru
сайт: пниэи.рф

(841-2) 59-33-50
приемная (841-2) 59-33-35
служба маркетинга (841-2) 59-33-43

Изделие ПОРТАЛ-10

Изделие является составной частью комплекса ПОРТАЛ и обеспечивает шифрование и имитозащиту конфиденциальной информации с длиной ключа 56 бит

Обеспечивает

- встречную работу с аналогичным изделием ПОРТАЛ-10, а так же с изделием ПОРТАЛ-1000
- криптографическую защиту IP-пакетов методом полной инкапсуляции
- прозрачное автоматическое шифрование/расшифрование информации с заданной стойкостью по алгоритму шифрования – ГОСТ 28147-89
- контроль целостности пакетов данных – имитозащиту по ГОСТ 28147–89
- аутентификацию источника данных
- ключевую систему – полносвязную ключевую матрицу с индивидуальными ключами на каждом направлении обмена, ключевая структура предусматривает работу с 5000 абонентами
- гибкую полнофункциональную настройку изделия (с ПЭВМ)
- возможность встречной работы через NATP преобразователи (через маршрутизаторы, межсетевые экраны) в сетях с «серой IP адресацией»
- возможность встречной работы через сети операторов сотовой связи 3G (HSDPA, HSUPA,UMTS), EDGE, GPRS, LTE при наличии сервера маршрутизации мобильного трафика;
- возможность подключения USB-модема непосредственно к изделию
- возможность встречной работы по каналам Ethernet (100BASE-T), Wi-Fi
- возможность работы в режиме сервера маршрутизации мобильного трафика
- контроль технического состояния готовности к работе
- контроль наличия действующих и очередных ключей
- контроль целостности программного обеспечения
- контроль меток точного времени
- функцию дистанционного конфигурирования (реконфигурирования)
- дистанционное управление ключами (ввод ключевой информации, переход с действующего ключа на очередной, полное и выборочное стирание ключевой информации)
- круглосуточную необслуживаемую работу

Электропитание изделия ПОРТАЛ-10 осуществляется от:

- сети переменного тока напряжением 220 В (+ 22 В; – 33 В) с частотой 50 Гц
- сети постоянного тока напряжением 5 В (стандартный USB интерфейс)

Акционерное общество
ПЕНЗЕНСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ ИНСТИТУТ

440000, г. Пенза, ул. Советская, 9
info@pniei.penza.ru
сайт: пниэи.рф

(841-2) 59-33-50
приемная (841-2) 59-33-35
служба маркетинга (841-2) 59-33-43

Изделие подключается к локальной сети, а также к оборудованию транспортной сети по интерфейсу Ethernet (100BASE-T на скорости 100 Мбит/с), Wi-Fi или к сети операторов сотовой связи 3G (HSDPA, HSUPA,UMTS), EDGE, GPRS, LTE через модемное оборудование и поддерживает протокол Ethernet 802.3 на портах, не внося ограничений в работу протоколов верхних уровней.

Габаритные размеры изделия: 165x110x30 мм; масса 0,7 кг.

Акционерное общество
ПЕНЗЕНСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ ИНСТИТУТ

440000, г. Пенза, ул. Советская, 9
info@pniei.penza.ru
сайт: пниэи.рф

(841-2) 59-33-50
приемная (841-2) 59-33-35
служба маркетинга (841-2) 59-33-43

Комплекс «Швейцар»

В АО «ПНИЭИ» создан комплекс технических средств, позволяющий решить задачу по защите информации, передаваемой по протоколам IP. Комплекс предназначен для организации защищенной связи и обмена информацией между сегментами телекоммуникационных систем ведомств, а также для решения задач автоматизации управления безопасностью, включая функции дистанционного управления средствами криптографической защиты.

На базе комплекса предусматривается построение подсистем имеющих в составе до 5000 объектов (объект – локальная сеть или отдельный пользователь): архитектура комплекса позволяет строить подсистему криптографической защиты с единым центром управления безопасностью либо с иерархической структурой управления и контроля, содержащей до 200 подсетей.

Разработка велась с учетом потребности средств защиты как в сегменте защиты информации, содержащей сведения, составляющие государственную тайну со степенью секретности «секретно», так и защиты конфиденциальной информации:

– изделия М-687 (гос. тайна), М-687А и М-687В (конфиденциальный контур) с пропускной способность до 100 Мбит/с со стыками Ethernet, обеспечивающие шифрование и имитозащиту информации и режим ввода ключевой информации по каналам связи и удаленного мониторинга, обеспечивают взаимодействие с изделием Швейцар-М (при защите конфиденциальной информации);

– изделие Швейцар-М (конфиденциальный контур) с пропускной способность до 40 Мбит/с, обеспечивающее шифрование и имитозащиту информации и режим ввода ключевой информации по каналам связи и удаленного мониторинга, обеспечивает взаимодействие с изделиями М-684А и М-687В;

– аппаратура М-684 (гос. тайна), М-684А и М-687В (конфиденциальный контур)- станции децентрализованного изготовления ключей и их распределения по каналам связи, с функциями удаленного мониторинга состояния технических средств комплекса, автоматизированного сбора и учета сведений о событиях безопасности в подсистеме криптографической защиты. Обеспечивают возможность организации многоуровневой иерархической подсистемы управления безопасностью в сетях IP, реализованных на базе изделий М-687 (М-687А, М-687В) и Швейцар-М.

Акционерное общество
ПЕНЗЕНСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ ИНСТИТУТ

440000, г. Пенза, ул. Советская, 9
info@pniei.penza.ru
сайт: пниэи.рф

(841-2) 59-33-50
приемная (841-2) 59-33-35
служба маркетинга (841-2) 59-33-43

– Все изделия комплекса удобны в эксплуатации, не требуют длительной специальной подготовки персонала, обеспечивают круглосуточную необслуживаемую работу, имеют относительно низкую стоимость по сравнению с аналогами.

– Комплекс является самостоятельной разработкой в полном объеме схемных решений и программного обеспечения, в нем отсутствует системное программное обеспечение сторонних разработчиков и не предъявляются требования к смежной аппаратуре.

Продукт прошел сертификацию на соответствие требованиям ФСБ по защите информации, содержащей сведения, составляющие государственную тайну, и на соответствие по защите информации, не содержащей сведений, составляющих государственную тайну.

Акционерное общество
ПЕНЗЕНСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ ИНСТИТУТ

440000, г. Пенза, ул. Советская, 9
info@pniei.penza.ru
сайт: пниэи.рф

(841-2) 59-33-50
приемная (841-2) 59-33-35
служба маркетинга (841-2) 59-33-43

Изделие М-687 (М-687А, М-687В)

ИЗДЕЛИЕ КРИПТОГРАФИЧЕСКОЙ ЗАЩИТЫ ДАННЫХ В IP-СЕТЯХ

Изделие обеспечивает работу

- М-687 в режиме шифрования и имитозащиты информации, содержащей сведения, составляющие государственную тайну с грифом не выше «секретно» (встречная работа с аналогичным изделием и изделием М-641)
- М-687А в режиме шифрования и имитозащиты конфиденциальной информации, класс КА (встречная работа с аналогичным изделием и изделиями М-641К
- М-687В в режиме шифрования и имитозащиты конфиденциальной информации, класс КВ (изделие работает встречно с аппаратурой Швейцар-Я)

Примечание – изделия изготавливаются по единой документации, различие – ключевые документы, вводимые на объектах эксплуатации.

Изделие имеет два исполнения

- аппаратура М-687 (М-687А, М-687В) РИВУ.465644.01 – с пультом управления ПБ090 РИВУ.468381.010
- аппаратура М-687 (М-687А, М-687В) РИВУ.465644.031-01 – без пульта управления ПБ090 РИВУ.468381.010

Изделие обеспечивает

- криптографическую защиту IP-пакетов методом полной инкапсуляции
- прозрачное автоматическое шифрование/расшифрование информации с заданной стойкостью по алгоритму шифрования – ГОСТ 28147–89
- контроль целостности пакетов данных – имитозащиту по ГОСТ 28147–89
- аутентификацию источника данных
- поддержку фрагментации пакетов
- возможность генерации «ложного трафика» и выравнивание размеров передаваемых пакетов (нормализацию трафика). Режим аналогичный режиму работы изделия «Сито»
- ключевую систему – полносвязную ключевую матрицу с индивидуальными ключами на каждом направлении обмена, ключевая структура предусматривает работу с 5000 абонентов
- гибкую полнофункциональную настройку изделия (с ПЭВМ)
- межсетевое экранирование информационных потоков с выполнением следующих требований:

Акционерное общество
ПЕНЗЕНСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ ИНСТИТУТ

440000, г. Пенза, ул. Советская, 9
info@pniei.penza.ru
сайт: пниэи.рф

(841-2) 59-33-50
приемная (841-2) 59-33-35
служба маркетинга (841-2) 59-33-43

- возможность задания правил фильтрации IP-пакетов для обоих направлений передачи (LAN-WAN, WAN-LAN), с не менее чем 100 правил для каждого направления передачи
- возможность протоколирования событий межсетевого экранирования
- поддержку классификации трафика на основе IP-адресов, номеров протоколов, номеров портов транспортных протоколов TCP и UDP, полей ToS или DiffServ и поддерживает маркировку и перемаркировку трафика по полям ToS или DiffServ в соответствии с заданными правилами.
- возможность назначения IP-адреса «вручную» (статическая адресация) и динамически по протоколу DHCP. Аппаратура с динамически назначенным IP-адресом WAN обеспечивает возможность встречной работы только с аппаратурой с «вручную» назначенным IP-адресом WAN
- возможность дистанционного мониторинга и управления ключевой информациият аппаратуры децентрализованного изготовления ключей (М-684):
  - контроль технического состояния готовности к работе
  - контроль наличия действующих и очередных ключей
  - контроль целостности программного обеспечения
  - контроль меток точного времени
  - функцию дистанционного конфигурирования (реконфигурирования)
  - дистанционное управления ключами (ввод ключевой информации, переход с действующего ключа на очередные, полное и выборочное стирание ключевой информации)
- защиту от НСД при вскрытии корпуса
- круглосуточную необслуживаемую работу
- пропускную способность 94 Мбит/с при длине передаваемых пакетов 1400 байт

М-687 имеет оригинальный, разработанный специалистами АО «ПНИЭИ» конструктив, выполняющий функции экранирования и теплоотвода с возможностью установки в 19” стойку (высота-1U).

Электропитание изделия осуществляется от сети переменного тока напряжением 220 В (+ 22 В; – 33 В) с частотой 50 Гц ± 2,5 Гц. Мощность, потребляемая изделием от сети переменного тока, не превышает 15 В·А.

Изделие подключается к локальной сети (или отдельной станции), а также к оборудованию транспортной сети по интерфейсам Ethernet (10BASE-T, 100BASE-TX, RJ-45 на скоростях 10 и 100 Мбит/с) и поддерживает протокол Ethernet 802.3 на портах, не внося ограничений в работу протоколов верхних уровней.

Габаритные размеры изделия: 392x316x52,5 мм, масса-5,3 кг.

Акционерное общество
ПЕНЗЕНСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ ИНСТИТУТ

440000, г. Пенза, ул. Советская, 9
info@pniei.penza.ru
сайт: пниэи.рф

(841-2) 59-33-50
приемная (841-2) 59-33-35
служба маркетинга (841-2) 59-33-43

Швейцар-М

ИЗДЕЛИЕ КРИПТОГРАФИЧЕСКОЙ ЗАЩИТЫ ИНФОРМАЦИИ В СЕТЯХ ПЕРЕДАЧИ ДАННЫХ

Предназначено для обеспечения безопасности конфиденциальной информации в IP-сетях стандарта IEEE 802.3/802.3u

Обеспечивает

- встречную работу с изделиями Швейцар-Я, Швейцар-М, М-687А, М-687В
- криптографическую аутентификацию изделий встречной работы
- криптографическую защиту IP-пакетов методом полной инкапсуляции
- прозрачное шифрование информации в режиме гаммирования с обратной связью по ГОСТ 28147-89
- контроль целостности пакетов данных – имитозащиту по ГОСТ 28147–89
- создание не менее 50 криптографически защищенных туннелей
- создание 10 новых криптографически защищенных туннелей в секунду
- защиту от кодирования открытой информации (выравнивание трафика, генерация ложного трафика, маркировка поля ToS)
- межсетевое экранирование сетевого трафика на основе пакетной фильтрации
- наличие механизма QoS на сетевом уровне
- наличие ключевой системы – полносвязной ключевой матрицы с индивидуальными ключами на каждом направлении обмена
- возможность встречной работы с 5000 изделий в сети
- ввод ключевой информации с использованием пульта ПБ090
- взаимодействие со станцией генерации и распределения ключей
- функциональную настройку с использованием пульта ПБ090, USB-flash накопителей и ПЭВМ, а также со стороны станции генерации и распределения ключей
- мониторинг работы изделия на ПЭВМ, подключаемой к управляющему порту изделия
- регистрация событий безопасности
- ведение статистики межсетевого экранирования
- контроль целостности программного обеспечения
- защиту от НСД при вскрытии корпуса
- круглосуточную необслуживаемую работу

По условиям эксплуатации изделие удовлетворяет требованиям групп 1.1, 1.3. Диапазон рабочих температур: от – 10 до +50 °С.

Изделие имеет специально разработанный малогабаритный экранированный, теплоотводящий корпус.

Габаритные размеры изделия: 230×165×30 мм.

Вес: ~1кг.

Акционерное общество
ПЕНЗЕНСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ ИНСТИТУТ

440000, г. Пенза, ул. Советская, 9
info@pniei.penza.ru
сайт: пниэи.рф

(841-2) 59-33-50
приемная (841-2) 59-33-35
служба маркетинга (841-2) 59-33-43

М-684

АППАРАТУРА ДЕЦЕНТРАЛИЗОВАННОГО ИЗГОТОВЛЕНИЯ, РАСПРЕДЕЛЕНИЯ КЛЮЧЕЙ И ОРГАНИЗАЦИИ МНОГОУРОВНЕВОЙ СИСТЕМЫ ДИСТАНЦИОННОГО МОНИТОРИНГА В IP-СЕТЯХ

Предназначена для децентрализованного изготовления и распределения шифрключей по каналам связи и организации многоуровневой системы дистанционного мониторинга в сетях передачи данных IP.

Обеспечивает

- децентрализованное изготовление ключевых документов
- распределение и доведение ключей до изделий Швейцар-М, М-687 (М-687А, М-687В) по каналам связи согласно заданной схемы распределения, а также до М-684, находящейся на нижележащих уровнях управления
- дистанционное управление ключами в изделиях Швейцар-М, М-687 (М-687А, М-687В) и М-684, находящихся на нижележащих уровнях управления, по каналам связи, включая управление сменой, стиранием (сбросом) ключей, контроль их наличия и состояния
- запись ключевой и служебной информации, в том числе и контроль правильности осуществленной записи на носители ДК-6 в целях доставки ключевой информации и ее непосредственного ввода в изделия Швейцар-М, М-687 (М-687А, М-687В), Швейцар-Я согласно заданной схемы распределения
- запись больших массивов ключевой и служебной информации, в том числе и контроль правильности осуществления записи на ВНИ в целях доставки до М-684, находящейся на нижележащих уровнях управления, при отсутствии канала связи
- поэкземплярный учет ключей, сроков их действия, стирания, а также отображение сведений о наличии действующих и очередных ключей в обслуживаемых изделиях Швейцар-М, М-687 (М-687А, М-687В), Швейцар-Я, а также в М-684, находящихся на нижележащих уровнях управления
- своевременную доставку очередных ключей для обслуживаемых изделий Швейцар-М, М-687 (М-687А, М-687В), а также для М-684, находящихся на нижележащих уровнях управления, с использованием каналов связи

Акционерное общество
ПЕНЗЕНСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ ИНСТИТУТ

440000, г. Пенза, ул. Советская, 9
info@pniei.penza.ru
сайт: пниэи.рф

(841-2) 59-33-50
приемная (841-2) 59-33-35
служба маркетинга (841-2) 59-33-43

- выполнение режима удаленного конфигурирования (реконфигурирования) и мониторинга изделий Швейцар-М, М-687 (М-687А, М-687В) в зашифрованном и имитозащищенном виде по каналам связи и отображение его результатов на мониторе:
  - контроль технического состояния готовности к работе
  - контроль наличия действующих и очередных ключей
  - контроль целостности программного обеспечения
  - контроль меток точного времени
  - функцию дистанционного конфигурирования (реконфигурирования)
  - дистанционное управления ключами (ввод ключевой информации, переход с действующего ключа на очередные, полное и выборочное стирание ключевой информации)
- выполнение следующих функций по управлению безопасностью:
  - ведение баз данных, обеспечивающих ввод, хранение и редактирование сведений о криптографической связанности абонентов, а также служебной информации об абонентах
  - ввод сведений о произошедших компрометациях, рассылка и доведение команд восстановления связи в целях исключения скомпрометированных абонентов из сети связи
- круглосуточную работу

Габаритные размеры: 370 х 313 х 70 мм

Акционерное общество
ПЕНЗЕНСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ ИНСТИТУТ

440000, г. Пенза, ул. Советская, 9
info@pniei.penza.ru
сайт: пниэи.рф

(841-2) 59-33-50
приемная (841-2) 59-33-35
служба маркетинга (841-2) 59-33-43

Изделие БиоЗамок

БиоЗамок-БВ

БиоЗамок-К

Изделие БиоЗамок предназначено для управления электромеханическим замком или защелкой входной двери с помощью смартфона. Электронное управление замком удобно, тем что позволяет отказаться от использования связки ключей. Это экономит время так как смартфон всегда под рукой.

Модификация БиоЗамок-БВ позволяет дополнительно проверить пользователя посредством биометрической аутентификации. В качестве биометрических характеристик могут использоваться изображение лица или отпечаток пальца. Изделие БиоЗамок поддерживает считывание бесконтактных карт и брелоков RFID, обеспечивает удаленный доступ к встроенной видеокамере.

Управление и конфигурирование изделия БиоЗамок может осуществляться через Web-интерфейс, как на ПК, так и с помощью смартфона.

Монтаж изделия БиоЗамок может производиться:

– на внешнюю сторону двери (БиоЗамок-БВ);

– в полость внутри каркаса двери (БиоЗамок-К);

– в стену (БиоЗамок-БВ).

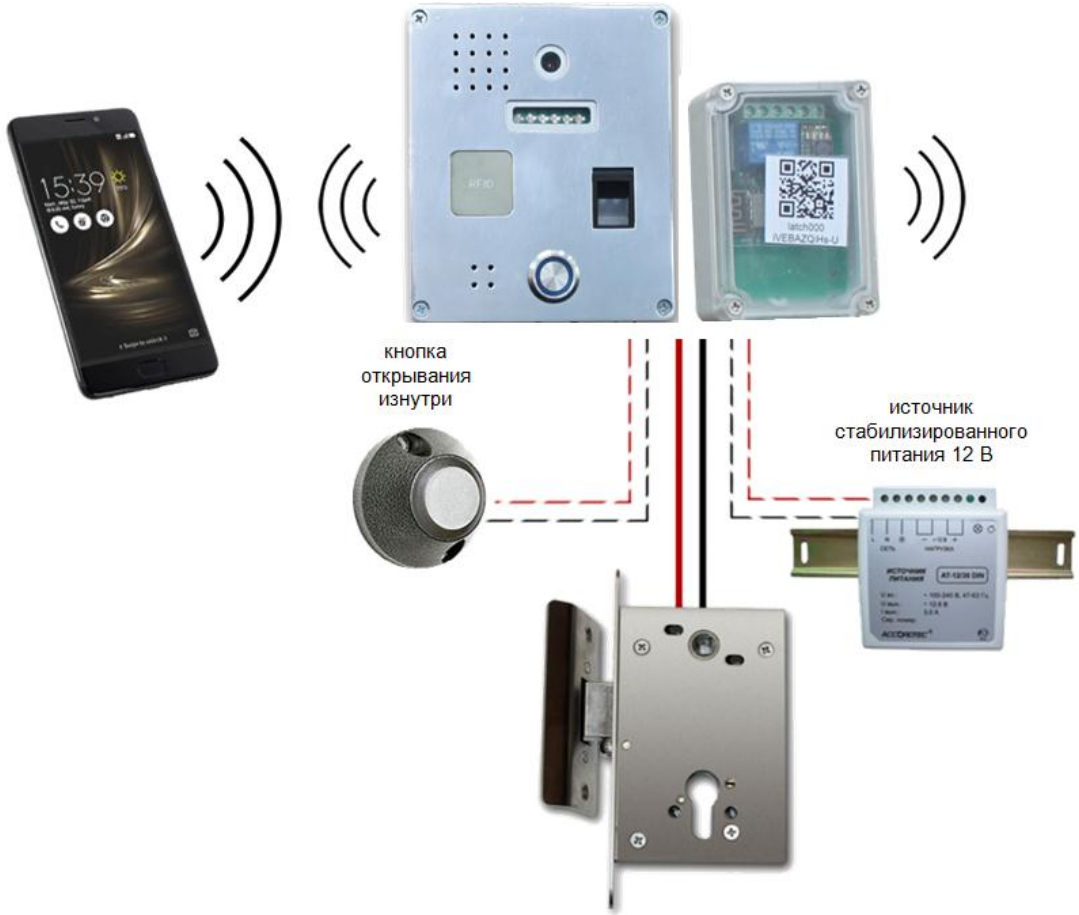
Акционерное общество
ПЕНЗЕНСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ ИНСТИТУТ

440000, г. Пенза, ул. Советская, 9
info@pniei.penza.ru
сайт: пниэи.рф

(841-2) 59-33-50
приемная (841-2) 59-33-35
служба маркетинга (841-2) 59-33-43

Технические характеристики изделия БиоЗамок

Напряжение питания	+12 В
Мощность потребления	6 Вт
Температура эксплуатации	от 0 до +50 °С
Допустимая влажность воздуха	не более 80 %
Вес (нетто)	0,5 кг
Габаритные размеры (ШхВхГ)	130x150x40 мм
Формат бесконтактных карт и брелоков	EM4100
Интерфейс беспроводной сети Wi-Fi	IEEE 802.11 b/g/n
Поддержка браузеров	Firefox, Chrome, Internet Explorer
Разрешение изображения	не менее 320 x 480 пикселей
Вероятность ошибочного предоставления доступа	менее 0,001 %
Вероятность ошибочного отказа в доступе	менее 1%



Акционерное общество
ПЕНЗЕНСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ ИНСТИТУТ

440000, г. Пенза, ул. Советская, 9
info@pniei.penza.ru
сайт: пниэи.рф

(841-2) 59-33-50
приемная (841-2) 59-33-35
служба маркетинга (841-2) 59-33-43

БиоТокен

ПРОГРАММНО-АППАРАТНОЕ СРЕДСТВО ДЛЯ ПРОВЕРКИ И ФОРМИРОВАНИЯ ЭЛЕКТРОННОЙ ПОДПИСИ С БИОМЕТРИЧЕСКИМ ПОДТВЕРЖДЕНИЕМ ЛИЧНОСТИ

Область применения

– системы обезличивания персональных данных медицинских учреждений (в составе изделия БиоГарант)

– системы электронного документооборота, торговли и услуг

и

– системы контроля и управления доступом с аппаратно-программным модулем доверенных вычислений, как отдельный фактор идентификации или для связывания биометрии с паролем доступа

– серверы децентрализованной идентификации пользователей

Функциональные возможности

– получение биометрических данных с графического планшета или сканера отпечатков пальцев

– создание и/или загрузка пары ключей формирования ЭП

– генерация псевдослучайных чисел с использованием естественной нестабильности биометрических образов

– формирование ЭП под электронными документами после биометрической авторизации пользователя

– связывание биометрии с личным ключом в процессе настройки БиоТокен с учетом требований пакета стандартов ГОСТ Р 52633 без выхода введенной биометрии и ключа пользователя из доверенной вычислительной среды БиоТокен

– обучение преобразователя биометрия-код (ГОСТ Р 52633.0–2006) на числе примеров биометрических образов «Свой» от 8 до 32 «Свой»

– хранение параметров связывания в защищенном биометрическом контейнере (ГОСТ Р 52633.4–2011)

– подтверждение критических операций загрузки данных в БиоТокен авторизованным пользователем

Электропитание устройства осуществляется от USB порта ПЭВМ. Устройство потребляет не более 150 мА.

Средство выполнено в форм-факторе USB, по условиям эксплуатации удовлетворяет требованиям климатического исполнения УХЛ4.2 ГОСТ 15150 с ограничением предельной пониженной температуры окружающей среды до минус 10°С.

Средний срок службы – не менее 5 лет.

Средняя наработка на отказ – не менее 10 000 ч.

Габаритные размеры – 72х40х17 мм.

Масса – не более 50 г.

Акционерное общество
ПЕНЗЕНСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ ИНСТИТУТ

440000, г. Пенза, ул. Советская, 9
info@pniei.penza.ru
сайт: пниэи.рф

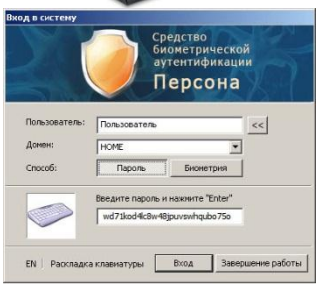
(841-2) 59-33-50
приемная (841-2) 59-33-35
служба маркетинга (841-2) 59-33-43

Персона

ПРОГРАММНОЕ СРЕДСТВО БИОМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ И АУТЕНТИФИКАЦИИ ПОЛЬЗОВАТЕЛЕЙ

Обеспечивает выполнение функций:

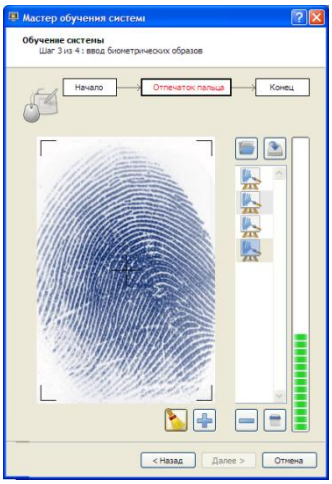
- защищённую биометрическую идентификацию субъектов доступа, проводимую путём создания нейросетевых биометрических контейнеров (НБК)
- простую защищённую и строгую биометрическую аутентификацию субъектов доступа с использованием НБК
- доступ к авторизованному запуску операционной системы Windows XP и контроль доступа к ее ресурсам с помощью средств криптографической защиты информации (СКЗИ)
- защиту файлов данных и контейнеров СКЗИ произвольного размера с помощью биометрических образов



Программное средство осуществляет преобразование легкозапоминаемого рукописного слова-пароля или отпечатка пальца в произвольный длинный пароль или ключ до 256 бит. Таким образом, пользователь избавлен от необходимости хранить надлежащим образом ключ или запоминать длинный случайный пароль. При подключении дополнительных модулей возможно связывание пароля (ключа) с голосовой фразой и другими биометрическими технологиями.

В программном средстве используются алгоритмы быстрого автоматического обучения искусственных нейронных сетей, параметры которых хранятся в нейросетевых биометрических контейнерах.

Преимуществом НБК является то, что сам ключ в них не хранится, не хранятся также биометрические образы пользователя.



Программное средство биометрической идентификации и аутентификации пользователей устанавливается на персональный компьютер с операционной системой семейства Windows, выполнено в соответствии с требованиями пакета стандартов ГОСТ Р 52633 и Федерального закона «О персональных данных».

Акционерное общество
ПЕНЗЕНСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ ИНСТИТУТ

440000, г. Пенза, ул. Советская, 9
info@pniei.penza.ru
сайт: пниэи.рф

(841-2) 59-33-50
приемная (841-2) 59-33-35
служба маркетинга (841-2) 59-33-43

Научное издание

БЕЗОПАСНОСТЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Сборник научных статей по материалам
I Всероссийской научно-технической конференции
(24 апреля 2019 г.)

Статьи печатаются в авторской редакции.

Компьютерная верстка *Р. Б. Бердниковой*
Дизайн обложки *А. А. Стаценко*

Подписано в печать 09.07.2019. Формат $60\times84^{1}/_{16}$.
Усл. печ. л. 11,86.
Заказ № 13693. Тираж 500.

Издательство ПГУ
440026, Пенза, Красная, 40
Тел./факс: (8412) 56-47-33; e-mail: iic@pnzgu.ru