http://vestnik.mrsu.ru
ISSN Print 2658-4123
ISSN Online 2658-6525

УДК 621.865.8:004.032.26 DOI: 10.15507/2658-4123.030.202002.300-312

Быстрая оценка энтропии длинных кодов с зависимыми разрядами на микроконтроллерах с малым потреблением и низкой разрядностью (обзор литературы по снижению размерности задачи)

А. И. Иванов[1*], **А. Г. Банных**[2]
[1]АО «Пензенский научно-исследовательский электротехнический институт» (г. Пенза, Россия)
[2]ФГБОУ ВО «Пензенский государственный университет» (г. Пенза, Россия)
[*]bio.ivan.penza@mail.ru

Введение. Целью работы является снижение требований к разрядности и производительности процессоров доверенной вычислительной среды при оценке энтропии длинных кодов с зависимыми разрядами.
Материалы и методы. Используются процедуры тестирования, рекомендованные национальными стандартами России. Используется переход от анализа обычных длинных кодов к расстояниям Хэмминга между случайными кодами «Чужой» и кодом образа «Свой».
Результаты исследования. Показано, что переход к представлению данных нормальным законом распределения в пространстве расстояний Хэмминга делает связь между математическим ожиданием и энтропией практически линейной. Построены низкоразрядные таблицы, связывающие первые статистические моменты распределения расстояний Хэмминга с энтропией длинных кодов. При вычислениях показатель коррелированности разрядов исследуемых кодов может изменяться в широких пределах.
Обсуждение и заключение. Вычисление математического ожидания и стандартного отклонения легковыполнимы на малоразрядных микроконтроллерах с низким потреблением. Пользуясь синтезированными таблицами, от младших статистических моментов расстояний Хэмминга удается легко переходить к энтропии длинных кодов. Задача вычисления энтропии многократно ускоряется в сравнении с процедурами Шеннона и становится реализуемой на дешевых малоразрядных процессорах.

Ключевые слова: микроконтроллеры с малым потреблением, тестирование нейронных сетей, энтропия длинных кодов, зависимые разряды, код

Для цитирования: Иванов, А. И. Быстрая оценка энтропии длинных кодов с зависимыми разрядами на микроконтроллерах с малым потреблением и низкой разрядностью (обзор литературы по снижению размерности задачи) / А. И. Иванов, А. Г. Банных. – DOI 10.15507/2658-4123.030.202002.300-312 // Инженерные технологии и системы. – 2020. – Т. 30, № 2. – С. 300–312.

Rapid Estimation of the Entropy of Long Codes with Dependent Bits on Low-Power, Low-Bit Microcontrollers (Review of Literature on Reducing the Dimension of a Problem)

A. I. Ivanov[a*], **A. G. Bannykh**[b]
[a]*Penza Research Electrotechnical Institute (Penza, Russia)*
[b]*Penza State University (Penza, Russia)*
[*]*bio.ivan.penza@mail.ru*

Introduction. The aim of the work is to reduce the requirements for bit depth and processor performance of a trusted computing environment when estimating the entropy of long codes with dependent bits.
Materials and Methods. Testing procedures recommended by the Russia national standards are used. The transition from the analysis of ordinary long codes to Hamming distances between random Alien codes and the Own image code is used.
Results. It is shown that the transition to the presentation of data by the normal distribution law in the space of Hamming distances makes the relationship between mathematical expectation and entropy almost linear. Low-bit tables are constructed that relate the first statistical moments of the distribution of Hamming distances to the entropy of long codes. In calculations, the correlation index of the digits of the studied codes can vary widely.
Discussion and Conclusion. The calculation of the mathematical expectation and standard deviation is easily feasible on low-discharge low-power microcontrollers. The use of the synthesized tables makes it possible to pass easily from the lower statistical moments of the Hamming distances to the entropy of long codes. The task of calculating entropy is accelerated many times in comparison with Shannon's procedures and becomes feasible on cheap low-bit processors.

Keywords: low-power microcontrollers, testing of neural networks, the entropy of long codes, dependent bits, code

For citation: Ivanov A.I., Bannykh A.G. Rapid Estimation of the Entropy of Long Codes with Dependent Bits on Low-Power, Low-Bit Microcontrollers (Review of Literature on Reducing the Dimension of a Problem). *Inzhenerernyye tekhnologii i sistemy* = Engineering Technologies and Systems. 2020; 30(2):300-312. DOI https://doi.org/10.15507/2658-4123.030.202002.300-312

Введение

Цифровая экономика должна базироваться на безопасных облачных сервисах обработки личных и корпоративных данных. Сегодня безопасность облачных сервисов строится на применении парольной аутентификации. К сожалению, человек не может запоминать длинные пароли из случайных символов для своей безопасной работы в интернет-облаках.

Для того чтобы избавить человека от необходимости запоминать длинный пароль доступа, в США, Канаде и странах Евросоюза пытаются применять так называемые «нечеткие экстракторы» [1–3]. Из-за того, что «нечеткие экстракторы» используют классические коды обнаружения и исправления ошибок с 20-кратной избыточностью, пароли, сцепленные с биометрией, оказываются короткими. Так, код пароля для папиллярного рисунка отпечатка пальца составляет примерно 16 бит, или 2 случайных символа в 8-битной кодировке. Очевидно, что вычисление энтропии пароля из двух случайных символов легко выполняется по Шеннону.

В России эта проблема решается использованием больших искусственных нейронных сетей, которые заранее

обучаются преобразовывать биометрические данные человека в длинный код его пароля доступа. При этом обучение выполняется автоматически алгоритмом ГОСТа Р 52633.5-2011[1]. За один бит кода пароля пользователя отвечает один нейрон нейросети, по этой причине нейросетевые преобразователи биометрии обычно строят под длину кода в 256 бит. Наиболее распространенные сегодня операционные системы Linux, Windows, Android воспринимают пароли длиной не больше 256 бит (32 случайных знака в 8-битной кодировке). Нейросетевые преобразователи биометрия-код выгодно строить под максимально возможную длину пароля доступа, характерную для той или иной операционной системы. По этой причине нейросетевую технологию защиты иногда называют «высоконадежной»[2]. Хакер, ничего не знающий о биометрическом образе пользователя «Свой», вынужден перебирать все состояния очень длинного кода доступа [4].

Как правило, после получения длинного личного ключа пользователя или его длинного пароля доступа запускается некоторый криптографический протокол выполнения аутентификации. Криптографические протоколы принято считать надежными, если энтропия блока шифротекста длинной в 256 бит будет составлять ровно 256 бит. Если энтропия оказывается больше или меньше, то криптосхема защиты информации может оказаться дефектной. Эти соображения с некоторой натяжкой можно перенести на защиту криптографического ключа «нечеткими экстракторами» или размещением его данных в параметрах обученной нейронной сети. Знание энтропии состояний кода ключа оказывается эффективным контрольным параметром при анализе уровня защиты от попыток его подбора.

В случае, если разрядность кода доступа мала, то расчет энтропии этих кодов можно выполнить по Шеннону. В частности, при кодах длиной 16 бит «нечеткого экстрактора» энтропию следует вычислять по следующей формуле [2; 3]:

$$H(x_1, x_2, ..., x_{16}) = -\sum_{i=1}^{65536} p_i \cdot \log_2(p_i), \quad (1)$$

где p_i – вероятность появления одного из $2^{16} = 65\,536$ состояний кодов.

Для того чтобы оценить вероятность появления 2^{16} состояний кода, необходимо иметь базу из 2^{16+4} папиллярных рисунков отпечатков пальцев «Чужой». Собрать базу из почти миллиона рисунков отпечатков пальцев сложно, но технически возможно. По этой причине процедуры вычисления энтропии по Шеннону для «нечетких экстракторов» вполне применимы.

Положение коренным образом меняется, если мы переходим к кодам длинной в 256 бит:

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

Для прямых оценок очень малых вероятностей появления 2^{256} состояний кода потребуется использовать базу из 2^{256+4} биометрических образов «Чужой». Технически невозможно создать и использовать столь большую тестовую базу.

Обзор литературы

Для решения проблемы больших тестовых баз и сложных вычислений энтропии по Шеннону в России разра-

[1] ГОСТ Р 52633.5-2011. Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа.

[2] ГОСТ Р 52633.3-2011. Защита информации. Техника защиты информации. Тестирование стойкости средств высоконадежной биометрической защиты к атакам подбора.

ботан стандарт ГОСТ Р 52633.5[3]. По рекомендациям этого стандарта требуется применение малых тестовых баз образов «Чужой» объемом от 21 до 64 примеров.

Применение стандарта дает экспоненциальное снижение объема тестовой выборки. Этот выигрыш обусловлен тем, что стандарт рекомендует переходить от анализа обычных кодов в пространство расстояний Хэмминга:

$$h = \sum_{i=1}^{256} \begin{bmatrix} " c_i " \\ \oplus \\ " x_i " \end{bmatrix}, \quad (3)$$

где "c_i" – дискретное состояние i-го разряда кода «Свой»; "x_i" – дискретное состояние i-го разряда случайного кода образа «Чужой»; $\oplus$ – операция сложения по модулю 2. Если пространство обычных 256-битных кодов имеет огромное число состояний, то пространство расстояний Хэмминга будет иметь всего 257 состояний, наблюдается экспоненциальное снижение размерности задачи.

По малой выборке от 21 до 64 опытов можем вычислить математическое ожидание $E(h)$ и стандартное отклонение $\sigma(h)$. Знание о значениях этих двух статистических моментов позволяет оценить вероятность угадывания кода «Свой» P_2. Оценка вероятности выполняется в рамках гипотезы нормальности:

$$P_2 \approx \frac{1}{\sigma(h)\sqrt{2\pi}} \int_{-\infty}^{1} \exp\left\{ \frac{-\left(E(h)-u\right)^2}{2\left(\sigma(h)\right)^2} \right\} \cdot du. \quad (4)$$

В этом случае энтропия нейросетевого преобразователя оценивается следующим образом:

$$H(" x_1, x_2, ..., x_{256} ") \approx -\log_2(P_2). \quad (5)$$

Применение тройки преобразований (3), (4), (5) позволяет уйти от экспоненциальной вычислительной сложности оценок энтропии длинных кодов по Шеннону[4]. Более того, эта совокупность вычислительных процедур обеспечивает линейную вычислительную сложность и снимает проблемы привлечения процессоров, поддерживающих 32- и 64-разрядные вычисления[5] [5].

Кроме того, отпадает необходимость тратить значительные объемы памяти на хранение больших объемов тестовой базы образов «Чужой». Достаточно помнить порядка 30 тестовых образов «Чужой». Столь малое число образов «Чужой» может быть использовано как образы-родители. Их морфинг-скрещиванием могут быть получены сотни тысяч синтетических тестовых образов-потомков с помощью процедур, регламентируемых национальным стандартом ГОСТ Р 52633.2-2010[6]. В одном поколении морфинг-скрещиванием удается увеличивать число естественных тестовых примеров «Чужой» примерно в 20 раз, соответственно, для получения больших тестовых баз приходится размножать данные в нескольких поколениях [6–9]. Процедуры размножения данных применимы и к примерам образа «Свой» в рамках бутстрап-идео-

[3] ГОСТ Р 52633.5-2011. Защита информации...

[4] Быстрые алгоритмы тестирования нейросетевых механизмов биометрико-криптографической защиты информации / А. Ю. Малыгин [и др.]. Пенза: Изд-во Пензенского государственного университета, 2006. 161 с. URL: https://scholar.google.ru/citations?user=iEZ66cUAAAAJ&hl=ru (дата обращения: 29.04.2020).

[5] Нейросетевая защита персональных биометрических данных / под ред. Ю. К. Язова. М.: Радиотехника, 2012. 157 с. URL: http://www.radiotec.ru/book/170 (дата обращения: 29.04.2020).

[6] ГОСТ Р 52633.2-2010. Защита информации. Техника защиты информации. Требования к формированию синтетических биометрических образов, предназначенных для тестирования средств высоконадежной биометрической аутентификации.

логии, заполнения пустых интервалов реальных гистограмм распределения[7].

К сожалению, применение морфинг-размножения биометрических образов в нескольких поколениях приводит к их вырождению. В частности, возникают дефекты «кошмара Дженкина» [10–12] и дефекты вырождения корреляционных матриц[8]. Крайне важным является то, что, применяя специальное программное обеспечение, удается практически полностью устранить дефекты размножения данных в первых нескольких поколениях[9].

Материалы и методы

Еще одним направлением работ, связанных с упрощением вычислений, является симметризация корреляционных связей в длинном коде с зависимыми разрядами. Идея метода сводится к замене изначально асимметричной корреляционной матрицы на ее симметричный аналог:

$$\begin{bmatrix} 1 & r_1 & r_2 & \vdots & r_n \\ r_1 & 1 & r_{n+1} & \vdots & r_{2n-2} \\ r_2 & r_{n+1} & 1 & \vdots & r_{3n-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ r_n & r_{2n-2} & r_{3n-3} & \vdots & 1 \end{bmatrix} \Rightarrow$$

$$\Rightarrow \begin{bmatrix} 1 & \tilde{r} & \tilde{r} & \vdots & \tilde{r} \\ \tilde{r} & 1 & \tilde{r} & \vdots & \tilde{r} \\ \tilde{r} & \tilde{r} & 1 & \vdots & \tilde{r} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \tilde{r} & \tilde{r} & \tilde{r} & \vdots & 1 \end{bmatrix}, \qquad (6)$$

где параметр симметричной матрицы $\tilde{r}$ вычисляется простым усреднением модулей всех коэффициентов корреляции исходной асимметричной матрицы:

$$\tilde{r} \approx E(|r_i|) \approx \frac{2}{n^2 - 2n} \cdot \sum_{i=1}^{0,5n^2-n} |r_i|. \quad (7)$$

При росте размерностей задачи (число нейронов в нейросети: 32, 64, 128, 256) ошибка приближения (7) быстро падает. По этой причине оценивать энтропию можно через вычисления показателя средней коррелированности (7) [13; 14]. Результаты в этом направлении исследований отражены в работах [15–18].

Очевидно, что при рассматриваемых классах преобразований крайне важным является то, насколько быстро сходятся вычислительные процессы и как быстро снижаются ошибки приближений. Эти вопросы рассматриваются в работах [19–22].

Первоначально задача вычисления энтропии длинных кодов с зависимыми разрядами рассматривалась как некоторая достаточно быстрая и эффективная процедура тестирования. Однако в 2009 г. с привлечением рассматриваемых в данной статье процедур была итерационно решена обратная задача нейросетевой биометрии. Через быстрое вычисление энтропии удалось извлечь знания из нейросети. Это послужило началом работ, улучшающих процедуры вычисления энтропии в контексте решения обратных задач нейросетевой биометрии [23–26].

---

[7] **Качалин С. В., Иванов А. И.** Заполнение пробелов биометрических данных генетическим алгоритмом размножения реальных примеров образа «Свой» без использования «мутаций» // Компьютерные науки и информационные технологии: материалы Международной науч. конф. (30 июня–03 июля 2014 г.). Саратов: Изд-во «Наука», 2014. С. 154–157.

[8] **Туреев С. В., Малыгина Е. А., Солопов А. И.** Методика формирования тестовых баз для проверки качества обучения нейросетевых преобразователей биометрия-код // Сборник научных статей по материалам I Всероссийской науч.-техн. конф. «Безопасность информационных технологий» (24 апреля 2019 г.). Пенза, 2019. С. 90–101.

[9] Свидетельство о государственной регистрации программы для ЭВМ № 2019662112 «Саморазвивающийся эмбрион-архив тестовой базы биометрических образов «чужой» / А. В. Безяев [и др.]. Дата государственной регистрации: 17.09.2019.

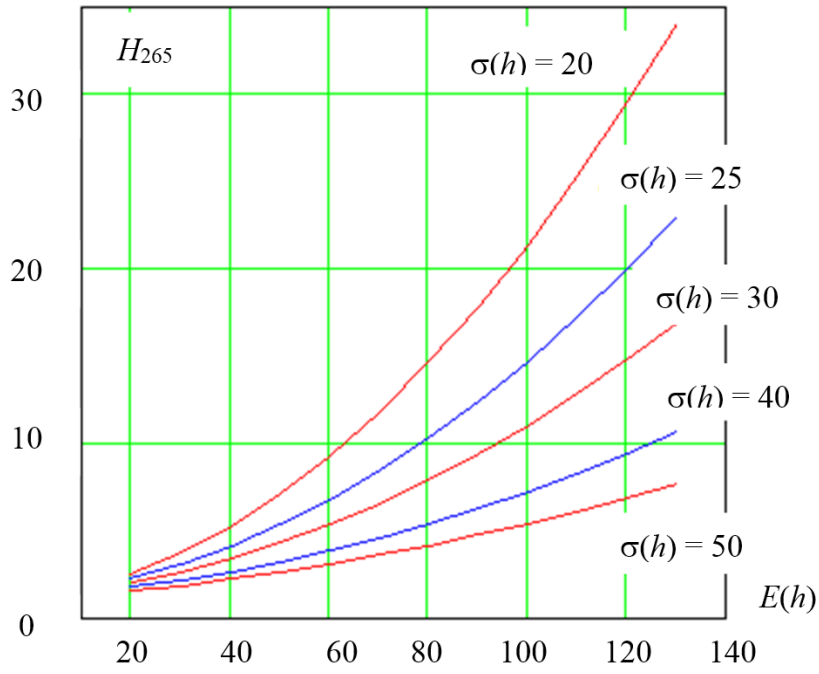
Результаты исследования

Одной из проблем нейросетевой биометрии является необходимость использования малогабаритной физически и криптографически защищенной доверенной вычислительной среды [27]. Как правило, доверенную вычислительную среду приходится создавать с использованием малопотребляющих, низкоразрядных микроконтроллеров при ограниченном объеме памяти. Система биометрико-криптографической защиты оказывается надежной, только если все криптографические операции над данными и нейросетевые операции будут выполняться внутри доверенной малогабаритной вычислительной среды.

Последнее означает, что на малопотребляющих, низкоразрядных процессорах нужно уметь точно вычислять интегралы вероятности (4). Вычислить подобный интеграл на компьютерах с 32- или 64-разрядной арифметикой несложно. При вычислении интеграла (4) на 4-, 8-разрядном микроконтроллере без привлечения специализированных программ (MathCAD, MatLAB, Maple и т. д.) возникают проблемы из-за низкой разрядности двоичных чисел и отсутствия поддержки повышения точности (разрядности) вычислений, обычно применяемой в типовом программном обеспечении математических пакетов.

Проведенные исследования показали, что если при 32-разрядных вычислениях зафиксировать стандартное отклонение расстояний Хэмминга $\sigma(h)$ и изменять только математическое ожидание расстояний Хэмминга $E(h)$, то мы получим почти линейную связь с оцениваемой энтропией (рис. 1).

При использовании малоразрядных процессоров экономически выгодно применять таблицы преобразований



Р и с. 1. Почти линейная связь энтропии с математическим ожиданием расстояний Хэмминга при фиксированном стандартном отклонении

F i g. 1. Almost linear relationship between entropy and mathematical expectation of Humming distances at fixed standard deviation

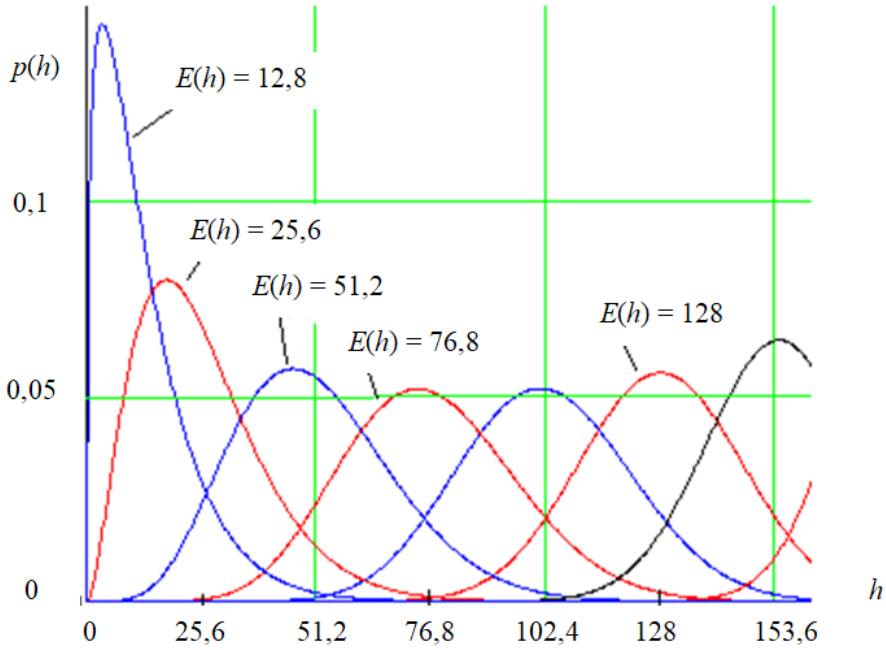
с шагом записи данных $20 + 10 \cdot i$, где строчный индекс i меняется от 0 до 13 для математического ожидания расстояний Хэмминга $E(h)$. Таблицы должны быть двухмерными, например, иметь 30 столбцов для значений стандартного отклонения $\sigma(h)$, изменяющегося на единицу. В ячейках двухмерной таблицы должны лежать значения энтропии номограммы, отображенной на рисунке 1.

Еще одним важным моментом является использование таблиц поправок значений энтропии DH_{256}, которые компенсируют ошибку от проявления асимметрии закона распределения расстояний Хэмминга $p(h)$.

Из рисунка 2 видно, что при значениях математического ожидания $E(h)$=12,8, $E(h)$=25,6 распределение расстояний Хэмминга становится существенно асимметричным [28; 29]. Это означает, что номограмма данных рисунка 1 при малых значениях математических ожиданий нуждается в значительных поправках DH_{256}. Вычисление поправок выполняется путем нормирования расстояний Хэмминга, приводящего все возможные состояния к интервалу от 0 до 1. В итоге получается 8-разрядная двоичная таблица преобразований. Далее под заданное значение асимметрии подбираются два параметра бета-распределения. Уже в пространстве бета-распределений по формуле, аналогичной формуле (4), находится вероятность ошибок второго рода на 32-разрядной вычислительной машине:

$$P_2 = \int_0^{\frac{1}{256}} \frac{1}{B(\beta 1, \beta 2)} \cdot \tilde{h}^{\beta 1 - 1} \cdot (1 - \tilde{h})^{\beta 2 - 1} \cdot d\tilde{h}, \quad (8)$$

где $\beta 1$, $\beta 2$ – параметры бета-распределения, обеспечивающие нуж-



Р и с. 2. Распределение расстояний Хэмминга для одинаково коррелированных данных $\tilde{r} = 0{,}33$ и монотонно изменяющихся значений математического ожидания

F i g. 2. Distribution of Humming distances for equally correlated data $\tilde{r} = 0.33$ and monotonically changing values of mathematical expectation

ный показатель асимметрии полученного распределения Хэмминга; $B(\beta 1, \beta 2) = \int_0^1 x^{\beta 1-1} \cdot (1-x)^{\beta 2-1} \cdot dx$ – бета функция; $\tilde{h} = h/256$ – нормированное расстояние Хэмминга.

Это и позволяет в конечном итоге синтезировать вторую таблицу поправок DH_{256} для малых значений математического ожидания для $E(h) < 70$ бит.

Обсуждение и заключение

Таким образом, использование стандартизованых в России нейросетевых преобразователей биометрии в длинный код позволяет снизить вычислительную сложность оценки энтропии с экспоненциальной вычислительной сложности до линейной и одновременно экспоненциально снизить объем памяти, необходимой для хранения тестовых образов «Чужой». Кроме того, в новых условиях перехода в пространство расстояний Хэмминга удается сделать вычисление интегралов вероятности табличным (4), (8) (используются две таблицы размерами 12x30 и 7x30, которые хранят 8-битные значения энтропии и ее поправок).

Следует отметить, что биометрические приложения являются одними из самых глубокопроработанных и стандартизованных приложений искусственного интеллекта. Формально это обусловлено тем, что международный технический комитет по стандартизации ISO/IEC JTC 1/SC 37 (Биометрия) создан в 2002 г., а технический комитет ISO/IEC JTC 1/SC 42 (Искусственный интеллект) создан только в 2017 г. По этой причине международных стандартов по биометрии на текущий момент разработано и находится в разработке 160, тогда как стандартов по искусственному интеллекту разработано и находится в разработке только 11.

Все, что изложено в данной статье, проверено только на биометрических данных, однако авторы уверены в том, что нейросетевое распознавание биометрических образов является частным случаем распознавания образов произвольной природы нейросетевым искусственным интеллектом. При этом перенос достигнутого уровня тестирований и доверия биометрических приложений корректен только в том случае, когда применяются «широкие» нейронные сети, стандартизованные пакетом отечественных стандартов с номерами ГОСТов Р 52633. xx-20xx. Нейронные сети других классов, в том числе сверточные нейронные сети «глубокого» обучения, тестировать с применением быстрых алгоритмов вычисления энтропии нельзя[10]. Как следствие, доверие к ним должно быть значительно ниже, чем доверие к отечественным стандартизованным нейросетевым конструкциям.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. **Juels, A.** A Fuzzy Commitment Scheme / A. Juels, M. Wattenberg. – DOI 10.1145/319709.319714 // CCS'99: Proceedings of the 6th ACM Conference on Computer and Communications Security. – 1999. – Pp. 28–36. – URL: https://dl.acm.org/doi/10.1145/319709.319714 (дата обращения: 29.04.2020).

2. **Ramírez-Ruiz, J. A.** Cryptographic Keys Generation Using Finger Codes / J. A. Ramírez-Ruiz, C. F. Pfeiffer, J. Nolazco-Flores. – DOI 10.1007/11874850_22 // Advances in Artificial Intelligence – IBERAMIA-SBIA. – 2006. – Pp. 178–187. – URL: https://link.springer.com/chapter/10.1007/11874850_22 (дата обращения: 29.04.2020).

[10] **Гудфеллоу Я., Бенджио И., Курвиль А.** Глубокое обучение. М.: ДМК Пресс, 2017. 652 с.

3. **Ушмаев, О. В.** Алгоритмы защищенной верификации на основе бинарного представления топологии отпечатка пальцев / О. В. Ушмаев, В. В. Кузнецов // Информатика и ее применения. – 2012. – Т. 6, № 1. – С. 132–140. – URL: http://www.ipiran.ru/journal/issues/2012_01_eng/ (дата обращения: 29.04.2020). – Рез. англ.

4. **Иванов, А. И.** Протоколы биометрико-криптографического рукопожатия: защита распределенного искусственного интеллекта интернет-вещей нейросетевыми методами / А. И. Иванов, П. А. Чернов / Системы безопасности. – 2018. – № 6. – С. 50–59. – URL: http://cs.groteck.ru/SS_6_2018/index.html?pn=&pageNumber= (дата обращения: 29.04.2020).

5. **Ivanov, A.** Statistical Description of Output States of the Neural Network “Biometrics-code” Transformers / A. Ivanov, B. Akhmetov, V. Funtikov [et al.] // Progress in Electromagnetics Research Symposium : PIERS Proceedings. – Moscow, 2012. – Pp. 62–65. – URL: https://piers.org/pierspublications/PIERS2012MoscowFinalProgram.pdf (дата обращения: 29.04.2020).

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. – 2014. – № 4 (104). – С. 194–199. – URL: https://official.satbayev.university/download/document/7138/ВЕСТНИК-2014 №4.pdf (дата обращения: 29.04.2020). – Рез. англ.

7. **Малыгин, А. Ю.** Требования к синтетическим базам биометрических образов и генераторам для их формирования / А. Ю. Малыгин, В. В. Федулаев, Д. Н. Надеев [и др.] // Нейрокомпьютеры: разработка, применение. – 2007. – № 12. – С. 60–64. – URL: http://www.radiotec.ru/article/3747 (дата обращения: 29.04.2020). – Рез. англ.

8. **Akhmetov, B.** Morph-Reproduction Examples of Parents in Several Generations of Examples Descen / B. Akhmetov, A. Ivanov, A. Malyghin [et al.] // International Conference on Global Trends in Academic Research. – Lumpur, 2014. – Pp. 188–190. – URL: https://globalilluminators.org/conferences/icmrp-2014-kuala-lumpur-malaysia/icmrp-full-paper-proceeding-2014/ (дата обращения: 29.04.2020).

9. **Волчихин, В. И.** Регуляризация вычисления энтропии выходных состояний нейросетевого преобразователя биометрия-код, построенная на размножении малой выборки исходных данных / В. И. Волчихин, А. И. Иванов, А. Г. Банных. – DOI 10.21685/2072-3059-2017-4-2 // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2017. – № 4. – С. 14–23. – URL: https://izvuz_tn.pnzgu.ru/tn2417 (дата обращения: 29.04.2020).

10. **Akhmetov, B.** Solving the Inverse Task of Neural Network Biometrics without Mutations and Jenkins “Nightmare” in the Implementation of Genetic Algorithms / B. Akhmetov, S. Kachalin, A. Ivanov [et al.] // International Conference “Computational and Informational Technologies in Science, Engineering and Education”. – 2015. – URL: http://conf.ict.nsc.ru/citech-2015/en/reportview/261166 (дата обращения: 29.04.2020).

11. **Качалин, С. В.** Алгоритм генетического обращения матриц нейросетевых функционалов без дефектов «кошмара» Дженкина / С. В. Качалин // Евразийский Союз Ученых. – 2015. – № 4 (13). – С. 59–62. – URL: https://euroasia-science.ru/tehnicheskie-nauki/алгоритм-генетического-обращения-ма/ (дата обращения: 29.04.2020).

12. **Качалин, С. В.** Направленное морфинг-размножение биометрических образов, исключающее эффект вырождения их популяции / С. В. Качалин, А. И. Иванов // Вопросы радиоэлектроники. – 2015. – № 1. – С. 76–85.

13. **Надеев, Д. Н.** Связь энтропии выходных состояний нейросетевых преобразователей биометрия-код с коэффициентами парной корреляции / Д. Н. Надеев, В. А. Фунтиков, А. И. Иванов // Нейрокомпьютеры: разработка, применение. – 2012. – № 3. – С. 74–77. – URL: http://www.radiotec.ru/article/10426 (дата обращения: 29.04.2020). – Рез. англ.

14. **Ахметов, Б. С.** Моделирование длинных биометрических кодов, воспроизводящих корреляционные связи выходных данных нейросетевого преобразователя / Б. С. Ахметов, В. И. Волчихин, С. В. Куликов [и др.] // Нейрокомпьютеры: разработка, применение. – 2012. – № 3. – С. 40–43. – URL: http://www.radiotec.ru/article/10418 (дата обращения: 29.04.2020). – Рез. англ.

15. **Bezyaev, V.** On the Issue of Modeling Long Biometric Codes with Dependent Bit States / V. Bezyaev, I. Serikov, A. Kruchinin [et al.] // Progress in Electromagnetics Research Symposium : PIERS Proceedings. – Moscow, 2012. – Pp. 62–65. – URL: https://piers.org/pierspublications/PIERS2012MoscowAbstracts.pdf (дата обращения: 29.04.2020).

16. **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. – DOI 10.1007/s10559-016-9838-x // Cybernetics and Systems Analysis. – 2016. – Issue 52. – Pp. 379–385. – URL: https://link.springer.com/article/10.1007%2Fs10559-016-9838-x (дата обращения: 29.04.2020).

17. **Волчихин, В. И.** Быстрый алгоритм симметризации корреляционных связей биометрических данных высокой размерности / В. И. Волчихин, Б. Б. Ахметов, А. И. Иванов // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2016. – № 1. – С. 3–7. – URL: https://izvuz_tn.pnzgu.ru/tn1116 (дата обращения: 29.04.2020).

18. **Ivanov, A. I.** Simplification of Statistical Description of Quantum Entanglement of Multidimensional Biometric Data Using Simmetrization of Paired Correlation Matrices / A. I. Ivanov, A. V. Bezyaev, A. I. Gazin // Journal of Computational and Engineering Mathematics. – 2017. – Vol. 4, Issue 2. – Pp. 3–13. – URL: https://jcem.susu.ru/jcem/article/view/110 (дата обращения: 29.04.2020).

19. **Качалин, С. В.** Оценка устойчивости алгоритмов обучения больших искусственных нейронных сетей биометрических приложений / С. В. Качалин // Вестник СибГАУ. – 2014. – № 3 (55). – С. 68–72. – URL: https://vestnik.sibsau.ru/vestnik/897/ (дата обращения: 29.04.2020). – Рез. англ.

20. **Иванов, А. И.** Номограммы оценки погрешности, коэффициентов корреляции, вычисленных на малых выборках биометрических данных / А. И. Иванов, Ю. И. Серикова // Вопросы радиоэлектроники. – 2015. – № 12. – С. 123–130.

21. **Волчихин, В. И.** Компенсация методических погрешностей вычисления стандартных отклонений и коэффициентов корреляции, возникающих из-за малого объема выборок / В. И. Волчихин, А. И. Иванов, Ю. И. Серикова // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2016. – № 1. – С. 45–49. – URL: https://izvuz_tn.pnzgu.ru/tn9116 (дата обращения: 29.04.2020).

22. **Иванов, А. И.** Корректировка методической погрешности вычисления статистических моментов четвертого порядка для малых выборок биометрических данных / А. И. Иванов, Ю. И. Серикова, А. Г. Банных // Модели, системы, сети в экономике, технике, природе и обществе. – 2016. – № 4 (20). – С. 108–114. – URL: https://mss.pnzgu.ru/mss416 (дата обращения: 29.04.2020).

23. **Akhmetov, B.** Evaluation of Multidimensional Entropy on Short Strings of Biometric Codes with Dependent Bits / B. Akhmetov, A. Ivanov, V. Funtikov // Progress in Electromagnetics Research Symposium : PIERS Proceedings. – Moscow, 2012. – Pp. 66–69. – URL: https://piers.org/pierspublications/PIERS2012MoscowFinalProgram.pdf (дата обращения: 29.04.2020).

24. **Иванов, А. И.** Биометрическая аутентификация личности: обращение матриц нейросетевых функционалов в пространстве метрики Хемминга / А. И. Иванов, Е. А. Малыгина // Вопросы защиты информации. – 2015. – № 1. – С. 23–29. – URL: http://izdat.ntckompas.ru/editions/magazine_news/detail.php?ELEMENT_ID=20164&SECTION_ID=155&ID=184 (дата обращения: 29.04.2020).

25. **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 [и др.]. – DOI 10.21685/2072-3059-2017-1-4 // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2017. – № 1. – С. 43–55. – URL: https://izvuz_tn.pnzgu.ru/tn4117 (дата обращения: 29.04.2020).

26. **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. – DOI 10.15507/0236-2910.027.201704.518-529 // Вестник Мордовского университета. – 2017. – Т. 27, № 4. – С. 518–529. – URL: http://vestnik.mrsu.ru/index.php/en/articles2-en/56-17-4/358-10-15507-0236-2910-027-201704-04 (дата обращения: 29.04.2020). – Рез. англ.

27. **Гулов, В. П.** Перспектива нейросетевой защиты облачных сервисов через биометрическое обезличивание персональной информации на примере медицинских электронных историй болезни (краткий обзор литературы) / В. П. Гулов, А. И. Иванов, Ю. К. Язов [и др.]. – DOI 10.12737/article_5947d5509f0411.58967456 // Вестник новых медицинских технологий. – 2017. – Т. 24, № 2. – С. 220–225. – URL: https://naukaru.ru/en/nauka/article/17188/view (дата обращения: 29.04.2020). – Рез. англ.

28. **Иванов, А. И.** Оценка вероятности ошибок биометрической аутентификации на малых выборках, использующая гипотезу бета-распределения расстояний Хэмминга / А. И. Иванов, А. В. Безяев, А. В. Елфимов [и др.] // Специальная техника. – 2017. – № 1. – С. 48–51.

29. **Ivanov, A. I.** A Simple Nomogram for Fast Computing the Code Entropy for 256-Bit Codes That Artificial Neural Networks Output / A. I. Ivanov, P. S. Lozhnikov, A. G. Bannykh // Journal of Physics: Conference Series. – 2019. – Vol. 1260, Issue 2. – URL: https://iopscience.iop.org/article/10.1088/1742-6596/1260/2/022003/meta (дата обращения: 29.04.2020).

Поступила 15.01.2020; принята к публикации 20.02.2020; опубликована онлайн 30.06.2020

Об авторах:

Иванов Александр Иванович, начальник лаборатории биометрических и нейросетевых технологий АО «Пензенский научно-исследовательский электротехнический институт» (440000, Россия, г. Пенза, ул. Советская, д. 9), доктор технических наук, доцент, Researcher ID: R-4514-2019, ORCID: https://orcid.org/0000-0003-3475-2182, bio.ivan.penza@mail.ru

Банных Андрей Григорьевич, аспирант кафедры информационной безопасности систем и технологий ФГБОУ ВО «Пензенский государственный университет» (440026, Россия, г. Пенза, ул. Красная, д. 40), ORCID: https://orcid.org/0000-0003-4776-5273, ibst@pnzgu.ru

Заявленный вклад соавторов:

А. И. Иванов – формулирование основной концепции, цели и задачи исследования, проведение расчетов, подготовка текста, формирование выводов; А. Г. Банных – проведение численного эксперимента, синтез таблиц для быстрого вычисления энтропии длинных кодов на малоразрядных процессорах.

Все авторы прочитали и одобрили окончательный вариант рукописи.

REFERENCES

1. Juels A., Wattenberg M. A Fuzzy Commitment Scheme. In: CCS'99: Proceedings of the 6th ACM Conference on Computer and Communications Security; 1999. Pp 28-36. (In Eng.) DOI: https://doi.org/10.1145/319709.319714

2. Ramírez-Ruiz J.A., Pfeiffer C.F., Nolazco-Flores J. Cryptographic Keys Generation Using Finger Codes. *Advances in Artificial Intelligence – IBERAMIA-SBIA*. 2006; Pp. 178–187. (In Eng.) DOI: https://doi.org/10.1007/11874850_22

3. Ushmaev O.S., Kuznetsov V.V. Secured Biometric Verification Based on Fingerprint Topology Binary Representation. *Informatika i eyo primeneniya* = Informatics and Applications. 2012; 6(1):132-140. Available at: http://www.ipiran.ru/journal/issues/2012_01_eng/ (accessed 29.04.2020). (In Russ.)

4. Ivanov A.I., Chernov P.A. Protocols of Biometric Cryptological Handshaking: Protecting Artificial Intelligence of the Internet of Things by Neural Network Methods. *Sistemy bezopasnosti* = Safety Systems. 2018; (6):50-59. Available at: http://cs.groteck.ru/SS_6_2018/index.html?pn=&pageNumber= (accessed 29.04.2020). (In Russ.)

5. Ivanov A., Akhmetov B., Funtikov V., et al. Statistical Description of Output States of the Neural Network "Biometrics-code" Transformers. In: Progress in Electromagnetics Research Symposium: PIERS Proceedings. Moscow; 2012. Pp. 62-65. Available at: https://piers.org/pierspublications/PIERS2012MoscowFinalProgram.pdf (accessed 29.04.2020). (In Eng.)

6. Akhmetov B.S., Kachalin S.V., Ivanov A.I. Supplement Fuzzy Biometric Morphing Reproduction Examples of Parents in Several Generations of Descendants Examples. *Vestnik KazNTU* = Vestnik KazNTU. 2014; (4):194-199. Available at: https://official.satbayev.university/download/document/7138/ВЕСТНИК-2014 №4.pdf (accessed 29.04.2020). (In Russ.)

7. Malygin A.Yu, Fedulaev V.V., Nadeev D.N., et al. Requirements for Synthetic Bases of Biometric Images and Generators for Their Formation. *Neyrokompyutery: razrabotka, primenenie* = Neurocomputers. 2007; (12):60-64. Available at: http://www.radiotec.ru/article/3747 (accessed 29.04.2020). (In Russ.)

8. Akhmetov B., Ivanov A., Malyghin A., et al. Morph-Reproduction Examples of Parents in Several Generations of Examples Descen. In: International Conference on Global Trends in Academic Research. Lumpur; 2014. Pp. 188-190. Available at: https://globalilluminators.org/conferences/icmrp-2014-kuala-lumpur-malaysia/icmrp-full-paper-proceeding-2014/ (accessed 29.04.2020). (In Eng.)

9. Volchikhin V.I., Ivanov A.I., Bannykh A.G. Regularizing Calculations of the Output Entropy of a Neural Network "Biometrics-Code" Converter though Multiplication of a Small Sample of Original Data. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Tekhnicheskie nauki* = University Proceedings. Volga Region. Engineering Sciences. 2017; (4):14-23. (In Russ.) DOI: https://doi.org/10.21685/2072-3059-2017-4-2

10. Akhmetov B., Kachalin S., Ivanov A., et al. Solving the Inverse Task of Neural Network Biometrics without Mutations and Jenkins "Nightmare" in the Implementation of Genetic Algorithms. In: International Conference "Computational and Informational Technologies in Science, Engineering and Education". 2015. Available at: http://conf.ict.nsc.ru/citech-2015/en/reportview/261166 (accessed 29.04.2020). (In Eng.)

11. Kachalin S.V. Algorithm of Genetic Circulation of Neural Network Functional Matrices without Jenkins' Nightmare Defects. *Yevraziyskiy Soyuz Uchenykh* = Eurasian Union of Scientists. 2015; (4):59-62. Available at: https://euroasia-science.ru/tehnicheskie-nauki/алгоритм-генетического-обращения-ма/ (accessed 29.04.2020). (In Russ.)

12. Kachalin S.V., Ivanov A.I. Direction of the Morph-Reproduction Biometric Images, the Exclusuve Degeneracy Effects of Their Populations. *Voprosy radioelektroniki* = Questions of Radio Electronics. 2015; (1):76-85. (In Russ.)

13. Nadeev D.N., Funtikov V.A., Ivanov A.I. Connection Between the Entropy of Neural Network Converters "Biometrics-Code" Output States and the Pair Correlation Coefficients. *Neyrokompyutery: razrabotka, primenenie* = Neurocomputers. 2012; (3):74-77. Available at: http://www.radiotec.ru/article/10426 (accessed 29.04.2020). (In Russ.)

14. Akhmetov B.S., Volchikhin V.I., Kulikov S.V., et al. Modeling of Long Biometric Codes, Reproducing Correlation of Neural Network Converter Output Data. *Neyrokompyutery: razrabotka, primenenie* = Neurocomputers. 2012; (3):40-43. Available at: http://www.radiotec.ru/article/10418 (accessed 29.04.2020). (In Russ.)

15. Bezyaev V., Serikov I., Kruchinin A., et al. On the Issue of Modeling Long Biometric Codes with Dependent Bit States. In: Progress in Electromagnetics Research Symposium: PIERS Proceedings. Moscow; 2012. Pp. 62-65. Available at: https://piers.org/pierspublications/PIERS2012MoscowAbstracts.pdf (accessed 29.04.2020). (In Eng.)

16. Ivanov A.I., Lozhnikov P.S., Serikova Yu.I. Reducing the Size of a Sample Sufficient for Learning Due to the Symmetrization of Correlation Relationships Between Biometric Data. *Cybernetics and Systems Analysis*. 2016; (52):379-385. (In Eng.) DOI: https://doi.org/10.1007/s10559-016-9838-x

17. Volchikhin V.I., Akhmetov B.B., Ivanov A.I. A Fast Symmetrization Algorithm for Correlations of Biometric Data of High Dimension. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Tekhnicheskie nauki* = University Proceedings. Volga Region. Engineering Sciences. 2016; (1):3-7. Available at: https://izvuz_tn.pnzgu.ru/tn1116 (accessed 29.04.2020). (In Russ.)

18. Ivanov A.I., Bezyaev A.V., Gazin A.I. Simplification of Statistical Description of Quantum Entanglement of Multidimensional Biometric Data Using Simmetrization of Paired Correlation Matrices. *Journal of Computational and Engineering Mathematics.* 2017; 4(2):3-13. Available at: https://jcem.susu.ru/jcem/article/view/110 (accessed 29.04.2020). (In Eng.)

19. Kachalin S.V. Assessment of Stability Learning Algorithms Large Artificial Neural Networks of Biometric Application. *Vestnik SibGAU* = Vestnik SibGAU. 2014; (3):68-72. Available at: https://vestnik.sibsau.ru/vestnik/897/ (accessed 29.04.2020). (In Russ.)

20. Ivanov A.I., Serikova Yu.I. Nomograms Error Estimates of Correlation Rates, Calculated for Small Samples of Biometric Data. *Voprosy radioelektroniki* = Questions of Radio Electronics. 2015; (12):123-130. (In Russ.)

21. Volchikhin V.I., Ivanov A.I., Serikova Yu.I. Compensation of Methodological Errors in Calculations of Standard Deviations and Correlation Coefficients Caused by Small Sample Sizes. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Tekhnicheskie nauki* = University Proceedings. Volga Region. Engineering Sciences. 2016; (1):45-49. Available at: https://izvuz_tn.pnzgu.ru/tn9116 (accessed 29.04.2020). (In Russ.)

22. Ivanov A.I., Serikova Yu.I., Bannykh A.G. Adjustment of Methodological Error Calculating the Statistical Moments of the Fourth Order for Small Samples of Biometric Data. *Modeli, sistemy, seti v ekonomike, tekhnike, prirode i obshchestve* = Models, Systems, Networks in Economics, Engineering, Nature and Society. 2016; (4):108-114. Available at: https://mss.pnzgu.ru/mss416 (accessed 29.04.2020). (In Russ.)

23. Akhmetov B., Ivanov A., Funtikov V. Evaluation of Multidimensional Entropy on Short Strings of Biometric Codes with Dependent Bits. In: Progress in Electromagnetics Research Symposium: PIERS Proceedings. Moscow; 2012. Pp. 66-69. Available at: https://piers.org/pierspublications/PIERS2012MoscowFinalProgram.pdf (accessed 29.04.2020). (In Eng.)

24. Ivanov A.I., Malygina Ye.A. Biometric Identity Vertification: Inversion of Matrices of Neural Network Functionals in the Hamming Space. *Voprosy zashchity informatsii* = Information Security Questions. 2015; (1):23-29. Available at: http://izdat.ntckompas.ru/editions/magazine_news/detail.php?ELEMENT_ID=20164&SECTION_ID=155&ID=184 (accessed 29.04.2020). (In Russ.)

25. Volchikhin V.I., Ivanov A.I., Bezyaev A.V., et al. Evaluation of the Calculation Acceleration Effect, Caused by the Support of Quantum Superposition States during Adjustment of Output Conditions of a "Biometrics-Code" Neural Network Converter. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Tekhnicheskie nauki* = University Proceedings. Volga Region. Engineering Sciences. 2017; (1):43-55. (In Russ.) DOI: https://doi.org/10.21685/2072-3059-2017-1-4

26. Volchikhin V.I., Ivanov A.I. Neural Network Molecule: a Solution of the Inverse Biometry Problem through Software Support of Quantum Superposition on Outputs of the Network of Artificial Neurons. *Vestnik Mordovskogo universiteta* = Mordovia University Bulletin. 2017; 27(4):518-529. (In Russ.) DOI: https://doi.org/10.15507/0236-2910.027.201704.518-529

27. Gulov V.P., Ivanov A.I., Yazov Yu.K., et al. Perspective of Neuro Network Protection of Cloud Services through Biometric Deployment of Personal Information on the Example of Medical Electronic History of Disease (Brief Review of the Literature). *Vestnik novykh meditsinskikh tekhnologiy* = Journal of New Medical Technologies. 2017; 24(2):220-225 (In Russ.) DOI: https://doi.org/10.12737/article_5947d5509f0411.58967456

28. Ivanov A.I., Bezyaev A.V., Yelfimov A.V., et al. Estimating the Probability of Biometric Authentication Errors on Small Samples using the Hypothesis of Beta Distribution Hamming Distances. *Spetsialnaya tekhnika* = Special equipment. 2017; (1):48-51. (In Russ.)

29. Ivanov A.I., Lozhnikov P.S., Bannykh A.G. A Simple Nomogram for Fast Computing the Code Entropy for 256-Bit Codes That Artificial Neural Networks Output. *Journal of Physics: Conference Series*. 2019; 1260(2). Available at: https://iopscience.iop.org/article/10.1088/1742-6596/1260/2/022003/meta (accessed 29.04.2020). (In Eng.)

Received 15.01.2020; revised 20.02.2020; published online 30.06.2020

About the authors:

Aleksandr I. Ivanov, Head of Laboratory of Biometric and Neural Network Technologies, Penza Research Electrotechnical Institute (9 Sovetskaya St., Penza 440000, Russia), D.Sc. (Engineering), Associate Professor, Researcher ID: R-4514-2019, ORCID: https://orcid.org/0000-0003-3475-2182, bio.ivan.penza@mail.ru

Andrey G. Bannykh, Postgraduate Student of Chair of Information Security of Systems and Technologies, Penza State University (40 Krasnaya St., Penza 440026, Russia), ORCID: https://orcid.org/0000-0003-4776-5273, ibst@pnzgu.ru

Contribution of the authors:

A. I. Ivanov – articulating the basic concept, goals and objectives of the study, performing the calculations, preparing the text, drawing the conclusions; A. G. Bannykh – conducting a numerical experiment, synthesis of the tables for rapid calculation of the entropy of long codes on low-bit processors.

All authors have read and approved the final manuscript.