УДК 004.056.053
DOI 10.21685/2587-7704-2018-4-1-5

Open Access | RESEARCH ARTICLE

Вычисление энтропии выходных состояний нейронных преобразователей биометрического кода

В. И. Семенов
Пензенский государственный университет, Россия, 440026 г. Пенза, ул. Красная, 40

А. И. Иванов
Пензенский государственный университет, Россия, 440026 г. Пенза, ул. Красная, 40

Аннотация. Целью работы было выяснить, какие бывают нейронные преобразователи биометрического кода, а также способы вычисления энтропии выходных состояний нейронных преобразователей биометрического кода. В результате было решено, что для вычисления энтропии логичнее использовать пространство сверток Хэмминга.

Ключевые слова: энтропия, биометрический код, нейронный преобразователь.

Output entropy calculation of neural biometric code converters

V. I. Semenov
Penza State University, 40 Krasnaya Street, 440026, Penza, Russia

A. I. Ivanov
Penza State University, 40 Krasnaya Street, 440026, Penza, Russia

Abstract. The purpose of the work was to find out various types of neural biometric-code converters, and methods for calculating their output entropy. As a result, it was decided that to calculate the entropy, it is more logical to use the Hamming convolution space.

Keywords: entropy, biometric code, neural converter.

Биометрическая аутентификация является одной из технологий обеспечения высокого уровня информационной безопасности. В России этому вопросу посвящена серия стандартов высоконадежной биометрической аутентификации ГОСТ Р 52633. Ключевым моментом в данной линейке стандартов является разработка нейросетевых преобразователей «Биометрия – код доступа».

Преобразователь биометрического кода – преобразователь, способный преобразовывать вектор нечетких, неоднозначных биометрических параметров «Свой» в четкий однозначный код ключа (пароля), откликающийся случайным выходным кодом на воздействие случайного входного вектора, не принадлежащего множеству образов «Свой».

Нейросетевые преобразователи биометрического кода классифицируют по виду входных биометрических параметров следующим образом:

1) нейросетевые преобразователи, ориентированные на работу с непрерывными биометрическими параметрами, имеющими, как правило, низкое среднее входное качество;

2) нейросетевые преобразователи, ориентированные на работу с дискретными биометрическими параметрами, имеющими, как правило, хорошее среднее входное качество.

Специализация нейронных сетей на обработку непрерывных или дискретных биометрических параметров обусловлена значительным различием алгоритмов их обучения и автоматов, реализующих эти алгоритмы.

Нейросетевые преобразователи биометрического кода также классифицируют по числу слоев нейронов, содержащихся в их нейронной сети.

Различают нейросетевые преобразователи с однослойной нейронной сетью и с двухслойной нейронной сетью.

Двухслойные нейронные сети, как правило, способны решать любые задачи высоконадежной биометрической идентификации и аутентификации. Дальнейшее увеличение числа слоев возможно, но для большинства биометрических приложений является избыточным и не рассматривается в рамках настоящего стандарта.

Нейронные сети должны осуществлять обогащение (повышение качества) исходных биометрических данных. Если удается решить задачу повышения качества исходных биометрических данных до приемлемого качества однослойной нейронной сетью, то применение для решения той же задачи двухслойной нейронной сети не рекомендуется. Корректировка незначительного числа ошибок однослойной нейронной сети допускается классическим избыточным кодом с обнаружением и исправлением ошибок.

Важнейшей характеристикой любого преобразователя «биометрия–код» является энтропия его выходных состояний, вычисленная при условии воздействия на преобразователь «биометрия–код» образами «Все чужие».

После получения данных от нейронных преобразователей биометрического кода необходимо вычислить энтропию выходных состояний.

Информационная энтропия – мера неопределенности или непредсказуемости информации. Это количество информации, приходящейся на одно элементарное сообщение источника, вырабатывающего статистически независимые сообщения.

Как правило, энтропию объекта исследования оценивают через наблюдение вероятности появления возможных состояний заранее заданного алфавита по Шеннону:

$$S(x) = -\sum_{i=1}^{n} p(i) \log_2 p(i) = \sum_{i=1}^{n} p(i) \log_2 \frac{1}{p(i)}. \quad (1)$$

Эта величина также называется средней энтропией сообщения. Энтропия в формуле Шеннона является средней характеристикой – математическим ожиданием распределения случайной величины [1].

Когда число состояний мало, проблем с оценкой энтропии не возникает. Этот классический подход вполне применим к оценке энтропии состояний преобразователей биометрического кода. 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. Можно показать, что для оценки энтропии одного бита с инженерной точностью достаточно 2^{1+8} примеров состояний контролируемого кода. Для оценки энтропии двух бит потребуется 2^{2+8} примеров состояний контролируемого кода. Мы наблюдаем экспоненциальный рост вычислительной сложности задачи и объема необходимой выборки исследуемых кодов. Попытка прямого вычисления энтропии 256 разрядов кода потребует обработки 2^{256+8} примеров. Выходом является вычисление многомерной энтропии в пространстве мер Хемминга. Это позволяет перенести задачу оценки биометрических образов энтропии из поля 2^n в пространство мер Хемминга размера $n + 1$. Очевидно, что переход от обычной классической энтропии к энтропии в пространстве мер Хэмминга является нелинейным преобразованием. Энтропия в пространстве мер Хэмминга является нелинейной функцией при малых размерностях n менее 100, при более высоких размерностях она практически линейно зависит от размерности, как и классическая энтропия [2].

В связи с этим происходит переход в пространство сверток Хэмминга:

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

где $\text{"}c_i\text{"}$ – i-й разряд проверяемого разряда кода синтезированного ключа; $\text{"}x_i\text{"}$ – эталонная последовательность «не белого» шума.

Принципиальным преимуществом перехода от обработки обычных кодов к работе со свертками Хэмминга (2) состоит в том, что размер выборки кодов сокращается логарифмически до величины $\log_2(2^{256+8}) = 256 + 8.$

Если мы будем иметь ключ "$\bar{c}$" с независимыми разрядами, то множество сверток Хэмминга (2) с фрагментами идеального «белого» шума должно давать следующие статистические моменты, исходя из распределения Бернулли для 256 независимых опытов:

$$\begin{cases} E(h) = 128 \text{ бит}, \\ \sigma(h) = 8 \text{ бит}. \end{cases} \quad (3)$$

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 $\sigma(h)$ будет существенно больше 8 бит, то мы должны сделать вывод о наличии сильной корреляционной связи разрядов синтезированного кода ключа [3].

Библиографический список

1. Иванов, А. И. Многомерная нейросетевая обработка биометрических данных с программным воспроизведением эффектов квантовой суперпозиции. – Пенза : Изд-во Пенз. научн.-исследоват. электротехнич. ин-та (АО «ПНИЭИ»), 2016. – 133 с.
2. 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 // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2017. – № 4 (44). – С. 14–23.
3. Основы биометрической аутентификации личности : учеб. пособие / Б. С. Ахметов, А. И. Иванов, А. Ю. Малыгин, В. А. Фунтиков. – Алматы : Учебно-издат. центр КазНТУ, 2014. – 150 с.

Образец цитирования:

Семенов, В. И. Вычисление энтропии выходных состояний нейронных преобразователей биометрического кода / В. И. Семенов, А. И. Иванов // Инжиниринг и технологии. – 2019. – Vol. 4(1). – С. 1–3. – DOI 10.21685/2587-7704-2018-4-1-5.