DOI 10.37882/2223–2966.2021.02.04

О ТРЕБОВАНИЯХ К ФОРМИРОВАНИЮ СИНТЕТИЧЕСКИХ ОБРАЗОВ ЭЛЕКТРОЭНЦЕФАЛОГРАММЫ ДЛЯ ЗАДАЧ ВЫСОКОНАДЕЖНОЙ БИОМЕТРИЧЕСКОЙ АУТЕНТИФИКАЦИИ[1]

REQUIREMENTS FOR THE FORMATION OF SYNTHETIC IMAGES OF THE ELECTROENCEPHALOGRAM FOR THE TASKS OF HIGHLY RELIABLE BIOMETRIC AUTHENTICATION

A. Borshevnikov
S. Goncharov
Yu. Dobrzhinskii

Summary. The article discusses algorithms of generation of synthetic images of the electroencephalogram used in the tasks of highly reliable biometric authentication. The paper describes the indicators of stability, uniqueness and quality of biometric parameters used to train the neural network converter "Biometrics — access code". Within the framework of the article, the requirements for the generation of synthetic biometric images of an electroencephalogram are formed using the example of the P300 potential.

Keywords: biometric authentication, neural net transformer "Biometrics — access code", key recovery, electroencephalogram, P300, natural biometric images, synthetic biometric images.

Боршевников Алексей Евгеньевич
Старший преподаватель, Дальневосточный федеральный университет (г. Владивосток)
borshevnikov.ae@dvfu.ru
Гончаров Сергей Михайлович
К.ф.-м.н., профессор, Морской государственный университет имени Г. И. Невельского (г. Владивосток)
sgprim143@gmail.com
Добржинский Юрий Вячеславович
К.т.н, профессор, Дальневосточный федеральный университет (г. Владивосток)
dobrzhinskii.yv@dvfu.ru

Аннотация. В статье рассматриваются алгоритмы генерации синтетических образов электроэнцефалограммы, использующейся в задачах высоконадежной биометрической аутентификации. В работе описываются показатели стабильности, уникальности и качества биометрических параметров, использующихсядляобучениянейросетевогопреобразователя«Биометрия— код доступа». В рамках статьи сформированы требования к генерации синтетических биометрических образов электроэнцефалограммы на примере потенциала P300.

Ключевые слова: биометрическая аутентификация, нейросетевой преобразователь «Биометрия — код доступа», восстановление ключа, электроэнцефалограмма, P300, естественные биометрические образы, синтетические биометрические образы.

Создание новых механизмов защиты информации является важной задачей для технологического развития государств. В Российской Федерации в рамках национальной технологической инициативы разрабатывается план создания рынка SafeNet, который предполагает разработку новых методов и устройств обеспечения информационной безопасности. В частности, особая роль в создании рынка SafeNet отводится системам высоконадежной биометрической аутентификации, что означает надежность данных систем, сравнимой с задачей полного перебора злоумышленником некоторого криптографического ключа. Также подобные системы принято называть биокриптографическими.

Создание биокриптографических систем на данном этапе развития технологий может основываться на двух подходах:

- первый подход основывается на использовании «нечетких» экстракторов и кодов исправляющих ошибки [1];
- второй подход базируется на использовании нейронных сетей для восстановления криптографического ключа пользователя. Данный подход получил в литературе название нейросетевых преобразователей «Биометрия — код доступа» [2,3].

Для обеспечения высокого уровня безопасности подобных технологий целесообразно использовать

[1] Исследование выполнено при финансовой поддержке Минобрнауки России (грант ИБ) № 9/2020

Таблица 1. Сравнение показателей уникальности, стабильности и качества естественных и синтетических данных электроэнцефалограммы

Характеристика	Характеристики естественных данных	Характеристики синтетических данных
$E(c(a_i))$	1.36	0.58
$\sigma(c(a_i))$	0.51	0.35
$\min(c(a_i))$	0.85	0.14
$\max(c(a_i))$	4.18	2.67
$E(u(a_i))$	0.29	4.04
$\sigma(u(a_i))$	0.12	1.76
$\min(u(a_i))$	0.14	0.69
$\max(u(a_i))$	0.68	8.40
$E(q(a_i))$	0.14	0.71
$\sigma(q(a_i))$	0.04	0.24
$\min(q(a_i))$	0.08	0.19
$\max(q(a_i))$	0.24	1.47

биометрические характеристики, обладающие высокой конфиденциальностью. Такой характеристикой является электроэнцефалограмма (ЭЭГ).

Одной из проблем, возникающей в ходе проведения исследований методов биометрической аутентификации, является проблема сложности сбора большого количества естественных образцов биометрической характеристики. Особенно остро подобная задача стоит как раз для высококонфиденциальных биометрических характеристик, так как их уровень конфиденциальности обычно напрямую связан с трудоемкостью сбора данных этих характеристик. Для решения данной проблемы принято использовать вместо естественных данных специально сгенерированные синтетические образцы. Однако для сохранения корректности исследуемой модели биометрической системы необходимо, чтобы характеристики синтетических биометрических параметров соответствовали значениям естественных параметров.

Для биометрических параметров вводится ряд характеристик, позволяющих охарактеризовать их: стабильность биометрического параметра $c(a_i)$, показатель уникальности биометрического параметра $u(a_i)$, показатель качества биометрического параметра $q(a_i)$, а также важными параметрами являются математические ожидания образов $E(a_i)$ и их стандартное отклонение $\sigma(a_i)$. Вычисление показателей стабильности, уникальности и качества указаны в работе [5].

Для заданных величин очевидно, что они зависят от значений математического ожидания и стандартного отклонения биометрических параметров образов «Свой» и «Чужой».

Для введения ограничений для формирования синтетических образцов необходимо описать данную процедуру. Стандарт ГОСТ Р 52633.2–2010 предусматривает генерацию синтетических образцов с помощью четырех алгоритмов, имеющих определенные недостатки: синтез, мутация, морфинг и перестановка фрагментов [3].

Для анализа алгоритмов генерации синтетических образцов, используемого в задачах высоконадежной биометрической аутентификации на основе ЭЭГ, был создан генератор синтетических образов, основанный на алгоритмах ГОСТ Р 52633.2.

Для проведения исследований были использованы данные опыта, проведенного с нейросетевым преобразователем «Биометрия — код доступа» ранее [4]. В част-

ности, использовались данные ЭЭГ для потенциалов P300. Дополнительно была сгенерирована база синтетических образов, состоящая из 104 образцов. Для первоначальной базы естественных образов «Чужой», содержащей 1000 примеров и дополненной синтетическими образами были рассчитаны характеристики биометрических параметров образов. Результаты расчетов этих данных приведены в таблице (таблица 1).

Как можно увидеть дополнение базы данных «Чужой» ведет к уменьшению значений показателей стабильности, и качества биометрических параметров ЭЭГ. Однако можно отметить рост уникальности биометрических параметров.

Для проверки результатов генерации образов было проведено исследование на нейросетевом преобразователе. При тестировании использовались параметры, использованные в рамках экспериментов [4,6]. Был проведен опыт по возможности получения злоумышленником секретного ключа при условии знания весовых коэффициентов и «мысленного пароля» легитимного пользователя. Точность аутентификации в рамках данного опыта составила 100%.

Полученные результаты говорят о том, что необходимо контролировать стандартное отклонение, которое должно уменьшаться или оставаться таким же, как и стандартное отклонение параметров естественных образов. Также важно контролировать уменьшение разности между математическими ожиданиями параметров естественных и синтетических образов. Данная величина должна увеличиваться, что обеспечит рост качества биометрических параметров. Также данная мера позволяет увеличить уникальность параметров. Если говорить о стабильности биометрических параметров, то данная характеристика может варьироваться, хотя желательно, чтобы она была равна 1.

Таким образом, установленные требования позволяют обеспечивать высокие показатели работы преобразователя «Биометрия — код доступа». Для более точной оценки качества работы генератора можно выделить следующие направления исследований:

- расширение базы естественных образов и изучение их характеристик;
- генерация базы синтетических образов, согласно определенных требований для генерации синтетических образов, и изучение их характеристик;
- проверка корректности введенных требований для генерации синтетических образов на полученных ранее более больших базах биометрических образов.

Также стоит отметить, что одним из теоретически эффективных подходов, которые могут позволить увеличить качество биометрической аутентификации, является генерация синтетических примеров электроэнцефалограммы на основе генеративно-состязательных сетей [7]. Эффективность построения алгоритма на основе указанного перспективного подхода требует отдельного исследования.

ЛИТЕРАТУРА

1. Dodis Y. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data / Y. Dodis, R. Ostrovsky, L. Reyzin, A. Smith // SIAM Journal on Computing.— 2008.— Vol. 38, № 1.— P. 97–139.
2. Защита информации. Техника защиты информации. Требования к формированию синтетических биометрических образов, предназначенных для тестирования средств высоконадежной биометрической аутентификации: ГОСТ Р 52633.2–2010.— Введен впервые; Введ. 30.09.2010.— М.: Стандартинформ, 2011.— 17 с.
3. Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия — код доступа: ГОСТ Р 52633.5–2011.— Введен впервые; Введ. 01.12.2011.— М.: Стандартинформ, 2012.— 20 с.
4. Гончаров С. М. Использование вейвлет-преобразования для выделения биометрических характеристик потенциала P300 в задачах высоконадежной биометрической аутентификации / Гончаров С. М., Боршевников А. Е. // Журнал «Информация и безопасность». Том. 19, часть 4. Воронеж: ВГТУ, 2016.— С. 527–530.
5. 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.
6. Гончаров С. М., Боршевников А. Е. Нейросетевой преобразователь «Биометрия — код доступа» на основе электроэнцефалограммы в современных криптографических приложениях. // Вестник СИБГУТИ: — Новосибирск: Изд-во СИБГУТИ, 2016.— № 1.— С. 17–22.
7. T. Piplani, N. Merill and J. Chuang, "Faking it, Making it: Fooling and Improving Brain-Based Authentication with Generative Adversarial Networks," 2018 IEEE9th International Conference on Biometrics Theory, Applications and Systems (BTAS), 2018, pp. 1–7.