УДК 681.322.067

С.М. Гончаров, А.Е. Боршевников

Построение нейросетевого преобразователя «Биометрия – код доступа» на основе параметров визуального вызванного потенциала электроэнцефалограммы

Рассматривается возможность построения нейросетевого преобразователя «Биометрия – код доступа» на основе ЭЭГ. Описывается структура нейросетевого преобразователя «Биометрия – код доступа». Предлагаются направления дальнейших исследований по разработке преобразователя «Биометрия – код доступа» на основе электроэнцефалограммы.

Ключевые слова: нейросетевой преобразователь «Биометрия – код доступа», секретный криптографический ключ, восстановление ключа, электроэнцефалограмма, визуальные вызванные потенциалы, биометрическая аутентификация.

В настоящее время идет активное развитие биометрических технологий. Одним из направлений развития данных технологий является биометрическая криптография. Основной задачей биометрической криптографии является привязка некоторой секретной информации (пароля или ключа) к определенной биометрической характеристике. Особенный интерес для использования в критических системах или приложениях, в которых используются элементы аутентификации или криптографической защиты информации, представляют характеристики деятельности мозга. Распространенной биометрией, характеризующей деятельность мозга, является электроэнцефалограмма, или ЭЭГ. Использование ЭЭГ в качестве биометрической характеристики дает несколько преимуществ. Данные электроэнцефалограммы конфиденциальны, их сложно подделать, а также они обеспечивают дополнительную меру защищенности от перехвата злоумышленником, заключающуюся в том, что снятие электроэнцефалограммы возможно на расстоянии не более 0,001 м от головы, что означает невозможность незаметного для пользователя съема данных. Помимо указанных преимуществ, внедрение технологии восстановления ключа из нечетких данных может обеспечить легкую смену «мысленного пароля» [1].

Один из эффективных подходов надежного хранения и восстановления секретного ключа был предложен в России. Для хранения секретных ключей (паролей) используются нейросетевые преобразователи «Биометрия – код доступа». Описанию данных преобразователей посвящена линейка стандартов ГОСТ Р 52633. Использование подобных преобразователей показывает хорошие результаты в вероятностях ошибок первого и второго рода [2].

В данной статье рассматривается нейросетевой преобразователь «Биометрия – код доступа» на основе электроэнцефалограммы с использованием бегущих цифр на экране, процедура восстановления ключа с помощью данного преобразователя, рассчитываются вероятности ошибок первого и второго рода.

Стимуляция на основе визуальных вызванных потенциалов. Опишем процедуру стимуляции деятельности мозга, при которой снимается электроэнцефалограмма для восстановления секретного ключа.

Используемая стимуляция для создания выглядит, как поочередно меняющиеся цифры от «0» до «9». Стимуляция для эксперимента вызывает визуальный вызванный потенциал [3]. Фрагмент стимуляции изображен на рис. 1.

Пользователи выбирают 1 или 2 символа и при их появлении на экране концентрируются на них. Данные символы являются «мысленным паролем».

Рис. 1. Фрагмент визуальной стимуляции

Съем ЭЭГ производился в течение 10 с. Для каждой секунды было использовано разбиение данной секунды на 128 частей, что соответствует синхронизации с нейрогарнитурой, используемой для съема ЭЭГ, и обеспечивает съем в реальном времени. Для случая когда пользователь запоминает 2 символа, съем ЭЭГ разбивается на два этапа по 5 с. В течение первого этапа пользователь концентрируется на одном символе, а в течение второго – на втором символе.

Биометрические характеристики визуального вызванного потенциала. В качестве биометрической характеристики a используется разница между уровнем ЭЭГ при стимуляции и усредненным значенем ЭЭГ в состоянии покоя. Обозначим уровень электроэнцефалограммы при стимуляции через $a_{\text{стим}}$, а усредненный уровень электроэнцефалограммы в состоянии покоя через $\overline{a}_{\text{покой}}$. Тогда

$$a = a_{\text{стим}} - \overline{a}_{\text{покой}}. \quad (1)$$

Однако в силу высокой сложности математического описания формы сигнала ЭЭГ [4] было принято решение производить выборку пятнадцати максимальных значений, вычисляемых по формуле (1). Целесообразно говорить об использовании характеристики a в векторном виде:

$$\overline{\mathbf{a}}_i = \{a_{ij}\}, \quad i = 1,...,14, \quad j = 1,...,15, \quad (2)$$

где $\overline{\mathbf{a}}_i$ – вектор биометрических данных, используемый в нейросетевом преобразователе; i – номер электрода, с которого снята электроэнцефалограмма; j – номер максимального значения a с канала i.

Построение и обучение нейросетевого преобразователя «Биометрия – код доступа» на основе ЭЭГ. В качестве структуры данного преобразователя выбрана двухслойная нейронная сеть сигмоидального типа.

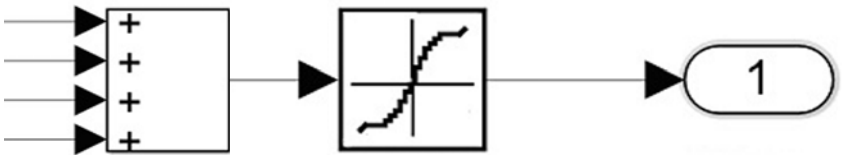
Для обучения выбрана стандартная процедура обучения нейросетевых преобразователей «Биометрия – код доступа», описанная в стандарте ГОСТ Р 52633.5–2011 [2]. Для обучения необходимо сформировать базу электроэнцефалограмм при воздействии стимуляции образов «Чужой», т.е. образов для которых нейросетевой преобразователь будет выдавать случайный криптографический ключ. Данную базу можно использовать для последующих процессов обучения преобразователя. Также необходимо сформировать базу электроэнцефалограмм образов «Свой» при состоянии покоя и при воздействии стимуляции. Данную базу необходимо удалить сразу после обучения преобразователя, в целях предотвращения её кражи и использования для компрометации секретного ключа. Результатом выполнения данной процедуры будут являться весовые коэффициенты нейронной сети:

$$\overline{\mathbf{M}}_i = \{M_{ij}\}, \quad i = 1,...,14, \quad j = 1,...,15, \quad (3)$$

$$\overline{\mathbf{M}} = \{M_k\}, \quad k = 1,...,320, \quad (4)$$

где $\overline{\mathbf{M}}_i$ – вектор весовых коэффициентов первого слоя нейронной сети, соответствующий вектору $\overline{\mathbf{a}}_i$; j – номер соответствующего компонента вектора $\overline{\mathbf{a}}_i$; $\overline{\mathbf{M}}$ – вектор весовых коэффициентов второго слоя нейронной сети; k – номер соответствующего нейрона первого слоя.

Нейроны первого и второго слоя сходны по строению (рис. 2), однако имеют различие в обрабатываемых данных и получаемых результатах.



Сумматор Передаточная функция Результат

Рис. 2. Строение нейрона нейросетевого преобразователя

Каждый нейрон первого слоя можно описать следующим образом:

$$x_1 = \sum \Delta \cdot \overline{\mathbf{M}}_i \cdot \overline{\mathbf{a}}_i = \sum \Delta \cdot \sum_{j=1}^{15} M_{ij} \cdot a_{ij}, \quad i = 1,...,14, \quad (5)$$

$$y_1 = y_1(x_1) = \frac{2}{1+e^{x_1}} - 1, \quad (6)$$

$$f_1(y_1) = \begin{cases} 1, & y_1 \geq 0, \\ -1, & y_1 < 0, \end{cases} \quad (7)$$

где x_1 – это результат работы сумматора нейрона первого слоя; Δ – коэффициент использования вектора $\overline{\mathbf{a}}_i$ в нейроне. Если $\overline{\mathbf{a}}_i$ используется в данном нейроне, то $\Delta = 1$ и $\Delta = 0$ в противном случае;

y_1 – передаточная функция первого слоя нейронной сети; $f_1(y_1)$ – решающее правило для нейрона первого слоя.

Используемые в сумматорах нейрона векторыбиометрических данных определяются следующим образом. В любом сумматоре обязательно используется 1 из 4 векторов, соответствующих векторам данных, снятых с электродов, расположенных на затылочной области головы. Данные электроды снимают ЭЭГ с области, в которой возникает наиболее сильный визуальный вызванный потенциал [4]. Для оставшихся трех входов сумматора используются 3 из 14 неиспользованных векторов биометрических данных.

Составим результирующий вектор работы первого слоя нейронной сети $\bar{t}$:

$$\bar{\mathbf{t}} = \{t_k\},\ k = 1,..., 320\ . \tag{8}$$

Каждый нейрон второго слоя можно описать следующим образом:

$$x_2 = \sum \Delta \cdot \overline{\mathbf{M}} \cdot \bar{t} = \sum \Delta \cdot M_k \cdot t_k,\quad k = 1,...,320\ , \tag{9}$$

$$y_2 = y_2(x_2) = \frac{2}{1+e^{x_2}} - 1\ , \tag{10}$$

$$f_2(y_2) = \begin{cases} 1, & y_2 \ge 0, \\ 0, & y_2 < 0, \end{cases} \tag{11}$$

где x_2 – это результат работы сумматора нейрона второго слоя; Δ – коэффициент использования компонента t_k в нейроне. Если t_k используется в данном нейроне, то $\Delta = 1$ и $\Delta = 0$ в противном случае; y_2 – передаточная функция второго слоя нейронной сети; $f_2(y_2)$ – решающее правило для нейрона второго слоя.

Используемые в сумматорах нейрона выходы первого слоя определяются согласно процедуре, описанной в ГОСТ Р 52633.5–2011 [2].

Результат работы каждого нейрона второго слоя является битом восстанавливаемого секретного криптографического ключа.

Полученные результаты работы нейросетевого преобразователя. Для проведения исследования построенного преобразователя была создана база из 10 различных биометрических образов, для каждого из которых было снято 20 примеров ЭЭГ в состоянии покоя и 80 примеров ЭЭГ под воздействием стимуляции. Один образ был выбран в качестве образа «Свой», остальные девять сформировали базу образов «Чужой».

Был проведен опыт по возможности получения злоумышленником секретного ключа при условии знания злоумышленником весовых коэффициентов. Наиболее интересными являются следующие результаты:

1. В случае, когда злоумышленник угадывает «мысленный пароль», расстояние Хэмминга от полученного злоумышленником ключа до секретного ключа пользователя было равно 7.

2. Во всех опытах по восстановлению ключа пользователем преобразователь безошибочно восстанавливал секретный ключ.

Приведем расчет ошибок первого и второго рода на основе результатов, полученных в ходе проведения опытов.

Для случаев, когда тестирующая выборка является небольшой и ошибка первого рода не была выбрана, данную ошибку можно вычислить по следующей формуле [4]:

$$P_1 \approx \int_1^{\infty} \frac{1}{2^{\frac{\Omega}{2}} \cdot \Gamma\left(\frac{\Omega}{2}\right)} \cdot x^{\frac{\Omega}{2}-1} \cdot e^{-\frac{x^2}{2}} \cdot dx\ , \tag{12}$$

где Ω – количество степеней свободы в распределении X^2.

В случае, когда в проведенной серии испытаний по предъявлению биометрической характеристики образа «Свой», состоящей из m опытов, не обнаружен факт отказа в доступе, число степеней свободы в распределении X^2 вычисляется по формуле

$$\Omega = \frac{1}{m+1}\ . \tag{13}$$

По формуле (12) получим ошибку первого рода: $P_1 = 6 \cdot 10^{-4}$.

Прогноз вероятности ошибок второго рода P_2 вычисляют приближенно, исходя из гипотезы нормального закона распределения значений вероятности ошибок по формуле [2]:

$$P_2 \approx \frac{1}{\sqrt{2\pi}} \int_{\sqrt{n}E(q(v))}^{\infty} e^{-\frac{x^2}{2}} dx, \qquad (14)$$

где n – число учитываемых преобразователем биометрических параметров; $E(q(v))$ – среднее качество всех учитываемых преобразователем биометрических параметров.

В построенном преобразователе использовалось 210 параметров, а среднее качество было получено равным 2,3. Тогда по формуле (14) получим ошибку второго рода: $P_2 \le 10^{-50}$.

Кроме описанных опытов, был проведен опыт по использованию в качестве «мысленного пароля» PIN-кода, состоящего из четырех символов. В ходе опыта было получено, что ошибка первого рода осталась на прежнем уровне, но среднее качество всех учитываемых параметров увеличилось до 3,6, и по формуле (14) было получено, что ошибка второго рода $P_2 \le 10^{-50}$. Полученный результат на порядки отличается от существующих средств биометрической аутентификации [5, 6].

Заключение. В данной работе описано построение нейросетевого преобразователя «Биометрия – код доступа» на основе данных электроэнцефалограммы и получены ошибки первого и второго рода для него.

Полученные результаты показывают, что необходимо дальнейшее исследование работы данного нейросетевого преобразователя. Необходимо:

1) увеличить размер базы электроэнцефалограмм не только по количеству биометрических образов в ней, но и по количеству биометрических образцов для каждого образа;

2) исследовать возможность увеличения расстояния Хэмминга от секретного ключа, получаемого злоумышленником, до ключа пользователя без утраты восстановительной способности преобразователя для пользователей;

3) оптимально подобрать коэффициенты обучения нейронной сети для преобразователя.

Однако уже сейчас полученные результаты показывают большие перспективы развития данной технологии.

Литература

1. Гончаров С.М. Идентификация пользователей на основе электроэнцефалографии с использованием технологий «Интерфейс мозг – компьютер» / С.М. Гончаров, М.С. Вишняков // Доклады Томского государственного университета систем управления и радиоэлектроники. – 2012. – № 1 (25), ч. 2. – С. 166–170.

2. ГОСТ Р 52633.5–2011. Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрии. – М.: Стандартинформ, 2012. – 20 с.

3. Гнездицкий В.В. Обратная задача ЭЭГ и клиническая электроэнцефалография (картирование и локализация источников электрической активности мозга). – М.: МЕДпрессинформ, 2004. – 624 с.

4. Оценка вероятностей появления ошибок нейросетевых преобразователей биометрия-код на основе малых выборок / Б.С. Ахметов, А.И. Иванов, А.Ю. Малыгин, Т.С. Картбаев // Труды II Международной научной конференции «Высокие технологии – залог устойчивого развития». – Алматы, Казахстан. – 2013. – Т. 1. – С. 234–237.

5. Мещеряков Р.В. Биометрические методы идентификации / Р.В. Мещеряков, А.А. Шелупанов, В.П. Бондаренко // Известия Южного федерального университета. Технические науки. – 2003. – Т. 33, № 4. – С. 176–177.

6. Костюченко Е.Ю. Распознавание пользователя по клавиатурному почерку на фиксированной парольной фразе в компьютерных системах / Е.Ю. Костюченко, Р.В. Мещеряков // Известия Южного федерального университета. Технические науки. – 2003. – Т. 33, № 4. – С. 177–178.

Гончаров Сергей Михайлович
Канд. физ.-мат. наук, доцент, зав. каф. «Безопасность информации и телекоммуникационных систем»
Морского государственного университета им. адм. Г.И. Невельского, Владивосток
Эл. почта: sgprim@smtp.ru, goncharov@msun.ru

Боршевников Алексей Евгеньевич
Инженер-программист Дальневосточного регионального учебно-научного центра
по проблемам информационной безопасности, Владивосток
Тел.: 8 (924-1) 31-67-97
Эл. почта: LAdG91@mail.ru

Goncharov S.M., Borshevnikov A.E.
Construction of neural network transformer «Biometrics – access code» based on the parameters of the visual evoked potential electroencephalogram

The construction of neural network transformer «Biometrics – access code» based on EEG is researched. The structure of neural network transformer «Biometrics – access code» is described. The directions for further research of transformer «Biometrics – access code» based on electroencephalogram is offered.
Keywords: neural network transformer «Biometrics – access code», secret cryptographic key, key recovery, electroencephalogram, visual evoked potentials, biometric authentication.