Труды научно-технической конференции кластера пензенских предприятий,
обеспечивающих **БЕЗОПАСНОСТЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ.**
Том 9, Пенза-2014, с. 46-50 (http://www.pniei.penza.ru/RV-conf/T9/C46)

НЕЙРОСЕТЕВОЙ ПРЕОБРАЗОВАТЕЛЬ "БИОМЕТРИЯ - КОД ДОСТУПА" НА ОСНОВЕ МЫСЛЕННОГО PIN-КОДА

Гончаров С.М., Боршевников А.Е. (г. Владивосток)
Морской государственный университет им. адм. Г.И. Невельского
Дальневосточный федеральный университет

Биометрическая аутентификация является одной из технологий обеспечения высокого уровня информационной безопасности. Классическое использование технологий биометрической аутентификации вызывает ряд проблем [1]. С целью устранения имеемых недостатков разрабатываются методы генерации ключевых последовательностей на основе биометрических данных. В России этому вопросу посвящена серия стандартов высоконадежной биометрической аутентификации ГОСТ Р 52633 [2,3]. Ключевым моментом в данной линейке стандартов является разработка нейросетевых преобразователей "Биометрия - код доступа" [2]. Помимо задачи аутентификации подобные средства решают задачу безопасного хранения секретных криптографических ключей.

Значительный интерес представляет разработка нейросетевых преобразователей на основе деятельности мозга. Распространенной биометрией, характеризующей деятельность мозга, является электроэнцефалограмма или ЭЭГ. Использование ЭЭГ в качестве биометрической характеристики дает несколько преимуществ. Данные ЭЭГ конфиденциальны, их сложно подделать, а процедура получения ЭЭГ практически исключает перехват информации по беспроводным каналам. Помимо указанных преимуществ внедрение технологии восстановления ключа из нечетких данных может обеспечить легкую смену "мысленного пароля" [4].

Стимуляция для выделения визуального вызванного потенциала

Опишем процедуру стимуляции деятельности мозга, при которой снимается электроэнцефалограмма для восстановления секретного ключа.

Используемая стимуляция выглядит, как поочередно меняющиеся цифры от «0» до «9». Ответом на стимуляцию является визуальный вызванный потенциал (VEP, ВВП) [5]. Фрагмент стимуляции изображен на рисунке (рис. 1).

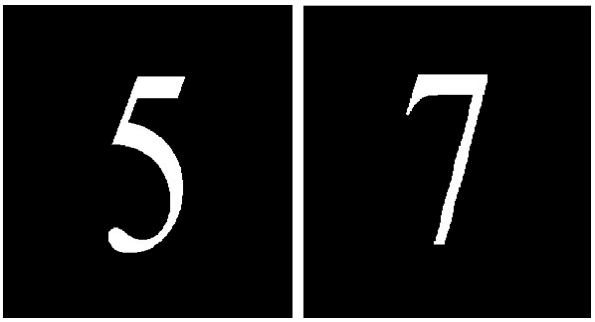

Рис. 1. Фрагмент визуальной стимуляции

Пользователи выбирают один, два или четыре символа, и при их появлении на экране концентрируются на них. Данные символы являются "мысленным паролем".

Съем ЭЭГ производился в течение 10 секунд. Для каждой секунды было использовано разбиение данной секунды на 128 частей, что обеспечивает синхронизацию с нейрогарнитурой, используемой для съема ЭЭГ. Для случая, когда пользователь запоминает два символа, съем ЭЭГ разбивается на два этапа по 5 секунд. В течение первого этапа пользователь концентрируется на одном символе, а в течение второго - на втором символе. В случае, когда пользователь запоминает четыре символа, структура эксперимента аналогична, но съем ЭЭГ производится 20 секунд.

Описание биометрического параметра

В качестве биометрического параметра a используется разница между уровнем ЭЭГ при стимуляции и усредненного значения 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:

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

В силу высокой сложности математического описания формы сигнала ЭЭГ [5] было принято решение производить выборку пятнадцати максимальных значений, вычисляемых по формуле (1). Целесообразно говорить об использовании параметра a в векторном виде:

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

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

Нейросетевой преобразователь "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]. 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 создать базу электроэнцефалограмм образов "Свой" при состоянии покоя и при воздействии стимуляции. Эту 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:

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

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

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

$\bar{a}_i$; $\overline{\mathrm{M}}$ – вектор весовых коэффициентов второго слоя нейронной сети; k – номер соответствующего нейрона первого слоя.

Используемые в сумматорах нейронов первого слоя вектора биометрических данных определяются следующим образом. В любом сумматоре обязательно используется 1 из 4 векторов, соответствующих векторам данных, снятых с электродов, расположенных на затылочной области головы (соответственно электроды P7, O1, O2, P8). Расположение электродов нейрогарнитуры указано на рисунке (рис. 2). Данные электроды снимают ЭЭГ с области, в которой возникает наиболее сильный визуальный вызванный потенциал [5]. Для оставшихся трех входов сумматора используются 3 из 10 не использованных векторов биометрических данных, и они распределяются согласно процедуре описанной в ГОСТ Р 52633.5-2011. Используемые в сумматорах нейрона входы второго слоя также определяются согласно процедуре, описанной в ГОСТ Р 52633.5-2011 [2].

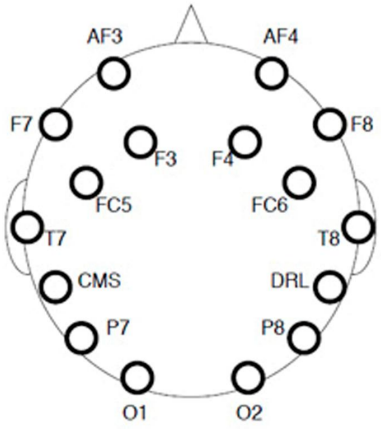

Рис. 2. Расположение электродов нейрогарнитуры

Каждый нейрон первого слоя можно описать следующим образом:

$$x_1 = \sum \Delta \cdot \overline{\mathrm{M}}_i \cdot \bar{a}_i = \sum \Delta \cdot \sum_{j=1}^{15} \mathrm{M}_{ij} \cdot a_{ij}, i = 1,...,14, \qquad (5)$$

$$y_1 = \mathrm{y}_1(x_1) = \frac{2}{1+e^{x_1}} - 1, \quad (6) \qquad \mathrm{f}_1(y_1) = \begin{cases} 1, y_1 \geq 0 \\ -1, y_1 < 0 \end{cases}, \qquad (7)$$

где x_1 – это результат работы сумматора нейрона первого слоя; Δ – коэффициент использования вектора $\bar{a}_i$ в нейроне. Если $\bar{a}_i$ используется в данном нейроне, то $\Delta = 1$ и $\Delta = 0$ в противном случае; y_1 – передаточная функция первого слоя нейронной сети; $\mathrm{f}_1(y_1)$ – решающее правило для нейрона первого слоя.

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

$$\bar{t} = \{t_k\}, k = 1,...,320. \qquad (8)$$

Каждый нейрон второго слоя можно описать следующим образом:

$$x_2 = \sum \Delta \cdot \overline{\mathrm{M}} \cdot \bar{t} = \sum \Delta \cdot \mathrm{M}_k \cdot t_k, k = 1,...,320, \qquad (9)$$

$$y_2 = \mathrm{y}_2(x_2) = \frac{2}{1+e^{x_2}} - 1, \qquad (10)$$

$$\mathrm{f}_2(y_2) = \begin{cases} 1, y_2 \geq 0 \\ 0, y_2 < 0 \end{cases}, \qquad (11)$$

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; $\mathrm{f}_2(y_2)$ – решающее правило для нейрона второго слоя.

Также был проведен опыт по использованию во втором слое преобразователя линейной передаточной функции. В этом случае каждый нейрон второго слоя можно описать следующим образом.

$$x_2 = \sum \Delta \cdot \overline{\mathrm{M}} \cdot \bar{t} = \sum \Delta \cdot \mathrm{M}_k \cdot t_k, k = 1,...,320, \quad (12)$$

$$y_2 = \mathrm{y}_2(x_2) = x_2, \quad (13)$$

$$\mathrm{f}_2(y_2) = \begin{cases} 1, y_2 \geq 0 \\ 0, y_2 < 0 \end{cases}, \quad (14)$$

Результатом работы каждого нейрона второго слоя является бит восстанавливаемого секретного криптографического ключа. Нейросетевой преобразователь выдает на выходе 256-битовый ключ.

Полученные результаты

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".

Проводились исследования по возможности получения злоумышленником секретного ключа при условии знания злоумышленником PIN-кода и без такового.

Таблица 1. Расстояния Хэмминга до исходного ключа в случае, когда злоумышленник угадал "мысленный пароль"

h_i - расстояние Хэмминга при "мысленном пароле" i.

№ образа «Чужой»	Нелинейный второй слой нейросетевого преобразователя			Линейный второй слой нейросетевого преобразователя		
	h_0	h_{17}	h_{2468}	h_0	h_{17}	h_{2468}
1	29	21	97	67	67	-
2	19	7	88	80	75	-
3	7	12	54	65	40	-
4	23	32	34	31	49	-
5	32	7	27	71	40	-
6	24	25	101	69	45	-
7	25	13	76	46	57	-
8	8	21	116	21	29	-
9	12	15	99	35	53	-

Как видно из таблицы даже тогда, когда злоумышленник угадывает "мысленный пароль", расстояние Хэмминга от полученного злоумышленником ключа до секретного ключа пользователя в случае линейного второго слоя нейронной сети весьма значительно. Работы по PIN-коду «2468» для линейного второго слоя не проводились.

Также были проведены опыты по восстановлению ключа пользователем, которому принадлежал этот ключ. При этом часть опытов проводилась с недельным разрывом. Во всех опытах преобразователь безошибочно восстанавливал секретный ключ.

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

$$P_2 \approx \frac{1}{\sqrt{2\pi}} \int_{\sqrt{n}E(q(v))}^{\infty} e^{-\frac{x^2}{2}} dx, \qquad (15)$$

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

В построенном преобразователе использовались 210 параметров, а среднее качество было получено равным 2,3 и 3,6 соответственно. Тогда по формуле (15) получим ошибки второго рода $P_2 \le 10^{-12}$.

Проведен анализ порядка 5000 ЭЭГ для 3 типов экспериментов и 15 пользователей. Для легитимных пользователей вырабатываемый секретный ключ всегда совпадал с истинным. Для злоумышленников при линейном втором слое нейросетевого преобразователя различие в ключе даже в случае угадывания «мысленного пароля» составило минимум 21 бит в 256 битном ключе.

ЛИТЕРАТУРА:

1. Харин Е.А. Построение систем биометрической аутентификации с использованием генератора ключевых последовательностей на основе нечетких данных / Е.А. Харин, С.М. Гончаров, П.Н. Корнюшин // Матер. 50-й Всерос. межвуз. науч.-техн. конф. – Владивосток: ТОВМИ, 2007. – С. 112–115.

2. 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.

3. Защита информации. Техника защиты информации. Требования к формированию баз естественных биометрических образов, предназначенных для тестирования средств высоконадежной биометрической аутентификации: ГОСТ Р 52633.1-2009. – Введен впервые; Введ. 15.12.2009. – М.: Стандартинформ, 2010. – 24 с.

4. 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: Научный журнал. –Томск: Изд-во ТУСУР, 2012. – № 1-2. – С. 166–170.

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

Материал поступил 09.07.2014, опубликовано по положительной рецензии доктора технических наук Иванова А.И.