Ирина В. Машкина¹, Елена П. Белова²

ФГБОУ ВО «Уфимский государственный авиационный технический университет»,
ул. К. Маркса, 12, г. Уфа, 450008, Республика Башкортостан, Россия
¹e-mail: mashkina.vtzi@gmail.com, <https://orcid.org/0000-0002-3096-3102>
²e-mail: super.yelenar@yandex.ru, <https://orcid.org/0000-0001-5488-943X>

РАЗРАБОТКА НЕЙРОСЕТЕВОЙ БАЗЫ ДАННЫХ БИОМЕТРИЧЕСКИХ ОБРАЗОВ
НА ОСНОВЕ НЕСКОЛЬКИХ ПАРАМЕТРОВ СПЕКТРОВ ГЛАСНЫХ ЗВУКОВ
ДЛЯ СИСТЕМЫ АУТЕНТИФИКАЦИИ И АВТОРИЗАЦИИ ПО ГОЛОСУ

DOI: <http://dx.doi.org/10.26583/bit.2019.3.08>

Аннотация. Данное исследование посвящено методам биометрической аутентификации и идентификации администраторов сети по голосу. В работе предложен метод представления уникальных параметров голоса, а также принцип построения искусственной нейронной сети для системы аутентификации пользователя и контроля доступа. Представлена структура системы биометрической аутентификации / авторизации пользователей информационной системы с использованием уникальных параметров голоса: частотного диапазона четвертой форманты гласного звука, частоты четвертой форманты и частоты лидирующей форманты спектральной характеристики. Построены две нейросетевые базы данных биометрических образов: первая базируется только на характеристиках четвертой форманты гласного звука, вторая – кроме них использует и частоту лидирующей форманты. Рассмотрены возможные решения специфических задач построения и обучения искусственной нейронной сети для построения базы биометрических образов пользователей и уменьшения ошибок первого и второго рода.

Ключевые слова: биометрический образ личности, нейросетевая база данных биометрических образов, характеристики четвертой форманты гласного звука, частота лидирующей форманты спектральной характеристики, аутентификация по голосу, авторизация по голосу.

Для цитирования: МАШКИНА, Ирина В.; БЕЛОВА, Елена П. РАЗРАБОТКА НЕЙРОСЕТЕВОЙ БАЗЫ ДАННЫХ БИОМЕТРИЧЕСКИХ ОБРАЗОВ НА ОСНОВЕ НЕСКОЛЬКИХ ПАРАМЕТРОВ СПЕКТРОВ ГЛАСНЫХ ЗВУКОВ ДЛЯ СИСТЕМЫ АУТЕНТИФИКАЦИИ И АВТОРИЗАЦИИ ПО ГОЛОСУ. Безопасность информационных технологий, [S.l.], v. 26, n. 3, p. 90-102, 2019. ISSN 2074-7136. Доступно на: <<https://bit.mephi.ru/index.php/bit/article/view/1221>>. Дата доступа: 17 sep. 2019. doi:<http://dx.doi.org/10.26583/bit.2019.3.08>.

Irina V. Mashkina¹, Yelena P. Belova²

Ufa State Aviation Technical University,
K. Marx Street, 12, Ufa, 450008, Republic of Bashkortostan, Russia
¹e-mail: mashkina.vtzi@gmail.com, <https://orcid.org/0000-0002-3096-3102>
²e-mail: super.yelenar@yandex.ru, <https://orcid.org/0000-0001-5488-943X>

Development of neural network database for biometric images based on several parameters of voice sound spectrum for authentication and authorization by voice

DOI: <http://dx.doi.org/10.26583/bit.2019.3.08>

Abstract. This study focuses on the methods of biometric authentication and identification of network administrators by voice. The authors propose a method for presenting unique voice parameters, as well as the principle of constructing an artificial neural network for a user authentication system and access control. The structure of the biometric authentication / authorization system of information system users is presented using unique voice parameters: the frequency range of the vowel fourth formant, the fourth formant frequency and the leading formant frequency of spectral characteristic. Two neural network databases of biometric images were built: the first is based only on the characteristics of the vowel fourth formant, the second uses the frequency of the leading formant as well. Possible solutions for specific tasks of building

and teaching ANNs are considered to build a base of user biometric images and reduce first and second types errors.

Keywords: biometric image of a person, neural network database of biometric images, characteristics of the fourth formant of the vowel sound, frequency of the leading spectral characteristic formants, voice authentication, voice authorization.

For citation: MASHKINA, Irina V.; BELOVA, Yelena P. Development of neural network database for biometric images based on several parameters of voice sound spectrum for authentication and authorization by voice. IT Security (Russia), [S.l.], v. 26, n. 3, p. 90-102, 2019. ISSN 2074-7136. Available at: <<https://bit.mephi.ru/index.php/bit/article/view/1221>>. Date accessed: 17 sep. 2019. doi:<http://dx.doi.org/10.26583/bit.2019.3.08>.

Введение

В настоящее время в индустрии информационных технологий наблюдается развитие биометрических систем аутентификации (БСА) [1], но при этом недостаточно широко освещается проблема использования этих систем с точки зрения достоверности распознавания личности, зависящей от возможных ошибок первого и второго рода. Аутентификация по голосу представляет собой одно из актуальных и развивающихся направлений в области распознавания личности по биометрическим характеристикам. При этом уникальность выбранного биометрического признака обеспечивает высокую достоверность, т.е. приемлемые значения ошибок распознавания личности [2].

Особое значение распознавание личности имеет в сфере компьютерной безопасности. Процедуры аутентификации и авторизации субъектов являются важнейшим механизмом защиты, от качества которого зависит безопасность информационной системы (ИС) [3].

Главным достоинством БСА по голосу, согласно источникам [4, 5], является относительно низкая стоимость, поскольку система может быть построена исключительно с использованием имеющихся на компьютере стандартных средств и разработанных программных модулей. Ещё одно достоинство, выявленное авторами статьи в работе [6], является возможность хранения в БСА биометрического образа пользователя, причём в случае использования искусственной нейронной сети (ИНС) в качестве базы биометрических образов этот эталон невозможно каким-либо способом выявить или скопировать.

К основным недостаткам БСА по голосу относится недостаточная (в случае однофакторной аутентификации) точность распознавания и влияние на результаты психофизиологического состояния личности [7]. Однако точность распознавания можно повысить путём использования различных дополнительных методов обработки голосовых биометрических данных. А зависимость результатов аутентификации по голосу от психофизиологического состояния личности можно использовать при реализации процедуры допуска к работе в информационных системах с высокой ценой ошибки оператора.

Построение искусственной нейронной сети – базы данных биометрических образов пользователей

В работе предлагается система аутентификации / авторизации пользователей, имеющих административные привилегии. В такой системе объекты инфраструктуры разделены на группы и сферы администрирования. Предложен механизм усиленной аутентификации и делегирования полномочий, который позволяет исключить суперпользователя путём создания отдельных ролей: администратора сети, администратора безопасности, администратора виртуальной инфраструктуры, администратора АСУТП.

На основе построения нейросетевой базы биометрических образов пользователей-администраторов реализуется механизм усиленной аутентификации администраторов, имеющих доступ к управлению конфигурацией.

В данной работе в качестве биометрического образа личности используются сразу несколько спектральных характеристик гласных звуков: характеристики четвёртой форманты и частота лидирующей форманты гласного звука речи, произнесённого пользователем. Такое решение обусловлено тем, что, согласно [8, 9], комбинирование нескольких признаков всегда обеспечивает существенное уменьшение величин ошибок 1-го и 2-го рода.

Под формантой принято понимать концентрацию энергии в ограниченной частотной области [10, 11]. Из данного определения следует, что четвёртая форманта гласного звука – это четвёртый по счёту всплеск энергии в ограниченной частотной области, который фиксируется на спектрограмме гласного звука.

В данной работе используются следующие *характеристики четвёртой форманты*:

- частота $f_{4\max}$, соответствующая максимальному значению амплитуды выброса энергии в четвёртой частотной области на спектрограмме;
- частота f_{4s} , определяющая начало четвёртой ограниченной частотной области на спектрограмме;
- частота f_{4f} , определяющая конец четвёртой ограниченной частотной области на спектрограмме.

В работах [4, 12] отмечается, что «четвёртая форманта характеризует речевой тракт человека». Эффективность её использования доказана авторами данной статьи в [6].

Частота лидирующей форманты f_l – это частота, соответствующая наибольшему значению амплитуды выброса любой из фиксируемых на спектрограмме формант.

Комбинации лидирующих формант гласных звуков различных пользователей имеют существенные отличия, что позволяет повысить точность распознавания личности.

При проведении исследований использовались характеристики четвёртых формант и частоты лидирующих формант, определённые по спектрограммам, полученным при произнесении пользователем трёх гласных звуков: "А", "О" и "Э". Спектрограммы именно этих звуков характеризуются чётко выраженными областями повышенной концентрации энергии в 4-й ограниченной частотной области. Примеры спектрограмм гласных звуков «А», «О» и «Э» представлены на рис. 1-3. На каждой спектрограмме жёлтым цветом отмечены: частота, определяющая начало четвёртой ограниченной частотной области, частота четвёртой форманты и частота, определяющая конец четвёртой ограниченной частотной области. Красным цветом показана частота лидирующей форманты. Спектрограммы звуков «И», «У» и «Ы» не имеют чётко выраженной четвёртой частотной ограниченной области. На рис. 4-5 в качестве примера приведены спектрограммы звуков «И» и «У».

К исследованию привлечено 4 человека: 3 женщины и 1 мужчина. При этом для эксперимента были приглашены женщины именно со схожими голосами, так как схожесть голосов пользователей приводит к близости спектральных характеристик, тем самым создавая одну из главных проблем современных систем аутентификации.

При создании алгоритма нахождения формант авторы статьи опирались на [13], а при написании кода программных модулей – на источники [10, 14]. Выделение характеристик четвёртой форманты гласного звука осуществляется при помощи двух

программных модулей [15]. Назначение первого программного модуля – получение спектрограммы гласного звука, произнесённого пользователем для выбранных в эксперименте фонем (рис. 1-5). Второй модуль реализует алгоритм выделения характеристик четвёртой форманты гласного звука.

Рис. 1. Спектрограмма звука «А», произнесённого одним из пользователей
(Fig. 1. Spectrogram of sound “A”, pronounced by one of the users)

Рис. 2. Спектрограмма звука «О», произнесённого одним из пользователей
(Fig. 2. Spectrogram of sound “O”, pronounced by one of the users)

Рис. 3. Спектрограмма звука «Э», произнесённого одним из пользователей
(Fig. 3. Spectrogram of sound “Э”, pronounced by one of the users)

Рис. 4. Спектрограмма звука «И», произнесённого одним из пользователей
(Fig. 4. Spectrogram of sound “I”, pronounced by one of the users)

Рис. 5. Спектрограмма звука «У», произнесённого одним из пользователей
(Fig. 5. Spectrogram of sound “U”, pronounced by one of the users)

Модуль выделения частоты лидирующей форманты (f_1) в настоящее время не реализован авторами. Поэтому процедура выделения частоты лидирующей форманты осуществлялась при помощи программы bard [16] в полуавтоматическом режиме.

Для реализации механизма аутентификации и авторизации пользователей в работе используется искусственная нейронная сеть (ИНС). В [17, 18] отмечаются следующие достоинства ИНС: быстрота обучения и самообучения, высокий уровень отказоустойчивости и помехоустойчивости, высокая скорость решения задач и широкая область применения.

Применительно к задаче подтверждения личности пользователя по предъявленным им биометрическим данным, процесс построения ИНС включает в себя следующие этапы: постановку задачи, выбор входных и выходных переменных ИНС, выбор архитектуры и структуры ИНС, построение обучающей выборки, получение и обработку результатов обучения и тестирования.

1. Постановка задачи

Задача данного исследования – создать *нейросетевую базу данных* биометрических образов (НБДБО) для усиленной аутентификации и авторизации пользователей ИС с административными полномочиями.

2. Выбор входных и выходных переменных ИНС

В качестве входных переменных используется вектор X , включающий в себя следующие показатели: номер пользователя x_1 , предъявившего свой идентификатор, 12 входов ($x_2, \dots, x_{13}$) для предъявленных пользователем биометрических данных, прошедших предварительную обработку по выделению характеристик четвёртых формант и частот лидирующих формант каждого из выбранных гласных звуков. Значения характеристик $x_2, \dots, x_{13}$ получены после обработки спектрограмм звуков, произнесённых пользователем в соответствии с тремя выбранными гласными фонемами. В качестве выходных переменных ИНС рассматривается вектор Y , размерность которого равна четырём (по числу администраторов).

3. Выбор архитектуры и структуры ИНС

Выбрана архитектура сети, позволяющая в полном объёме решить поставленные в работе задачи, а именно: – двухслойный персептрон прямого распространения. Количество нейронов во внутреннем слое выбрано с учётом объёма обучающей выборки, количества входов и выходов [19]:

$$N = \left[\frac{Q * n}{m + n} \right], \quad (1)$$

где: N – количество нейронов в скрытом слое,

m – количество входов,

n – количество выходов,

Q – объём обучающей выборки,

$[]$ – операция округления до целого числа.

Аргументировать выбор данной архитектуры ИНС можно тем, что она представляет собой универсальный аппроксиматор, позволяющий реализовать нелинейное отображение F :

$$X \longrightarrow Y,$$

где: $X = (x_1, \dots, x_m)$ – входной вектор размерности m ;

$Y = (y_1, \dots, y_n)$ – выходной вектор размерности $n=4$.

В рамках данного исследования была создана ИНС, базирующаяся только на характеристиках четвёртой форманты. При этом на входы ИНС *единовременно* подавались результаты обработки спектрограмм произнесённых пользователем гласных звуков всех заявленных фонем (рис. 6).

Входной нейрон 1 – номера идентификаторов пользователей с 1 по 4; входные нейроны 2–4 – значения характеристик четвёртой форманты гласного звука «А»; входные нейроны 5–7 – значения характеристик четвёртой форманты гласного звука «О»; входные нейроны 8–10 – значения характеристик четвёртой форманты гласного звука «Э». Таким образом, $m = 10$. Выходные нейроны П1, П2, П3 и П4 – данные о пользователях с первого по четвёртый. Активный выходной нейрон свидетельствует о том, биометрические данные какого пользователя были поданы на вход нейросети.

Во втором варианте построения ИНС введены дополнительные входы: входной нейрон 11 – значение частоты лидирующей форманты гласного звука «А»; входной нейрон 12 – значение частоты лидирующей форманты гласного звука «О»; входной нейрон 13 – значение частоты лидирующей форманты гласного звука «Э» (рис. 7). То есть $m = 13$.

Рис. 6. Архитектура искусственной нейронной сети, базирующейся на характеристиках четвёртой форманты гласных звуков «А», «О» и «Э»
 (Fig. 6. The Architecture of an Artificial Neural Network based on the Characteristics of the Fourth Formant of Vowel Sounds “A”, “O” and “Э”)

Рис. 7. Архитектура искусственной нейронной сети, базирующейся на характеристиках четвёртой форманты и частотах лидирующих формант гласных звуков «А», «О» и «Э»
 (Fig. 7. The Architecture of an Artificial Neural Network based on the Characteristics of the Fourth Formant and the Frequencies of the Leading Formant of the vowel sounds “A”, “O” and “Э”)

4. Построение обучающей выборки

Обучающая выборка для второго варианта ИНС приведена в таблицах 1 и 2.

Таблица 1. Входные данные обучающей выборки для ИНС, базирующейся на характеристиках четвёртой форманты и частотах лидирующих формант гласных звуков «А», «О» и «Э»

№ строки	№ №	Значения характеристик четвёртой форманты гласного звука									Значения частот лидирующей форманты гласного звука		
		$f_{4\max}(A)$	$f_{4s}(A)$	$f_{4f}(A)$	$f_{4\max}(O)$	$f_{4s}(O)$	$f_{4f}(O)$	$f_{4\max}(Э)$	$f_{4s}(Э)$	$f_{4f}(Э)$	$f_l(A)$	$f_l(O)$	$f_l(Э)$
1	1	815	710	908	781	720	961	821	732	900	815	579	647
2	1	849	779	926	808	747	877	822	738	889	199	585	616
...	...	...	...	...	...	...	...	...	...	...	...	...	...
10	1	835	733	920	759	666	802	818	737	879	220	574	610
11	1	774	636	831	902	828	996	762	667	850	196	221	570
12	1	784	710	866	864	784	926	770	675	852	1003	417	574
...	...	...	...	...	...	...	...	...	...	...	...	...	...
19	1	722	639	792	745	675	848	787	668	879	898	198	176
20	1	773	693	832	759	675	828	840	776	945	197	195	650
21	1	531	441	653	539	562	603	529	464	584	531	436	529
...	...	...	...	...	...	...	...	...	...	...	...	...	...
30	1	548	482	603	564	384	668	529	470	569	657	453	529
31	1	898	845	961	912	860	860	880	855	949	898	467	222
...	...	...	...	...	...	...	...	...	...	...	...	...	...
40	1	917	846	873	928	868	962	917	834	1004	917	486	457
41	2	774	636	831	902	828	996	762	667	850	196	221	570
42	2	784	710	866	864	784	926	770	675	852	1003	417	574
...	...	...	...	...	...	...	...	...	...	...	...	...	...
50	2	773	693	832	759	675	828	840	776	945	197	195	650
51	2	815	710	908	781	720	961	821	732	900	815	579	647
...	...	...	...	...	...	...	...	...	...	...	...	...	...
80	2	917	846	873	928	868	962	917	834	1004	917	486	457
81	3	531	441	653	539	562	603	529	464	584	531	436	529
82	3	518	428	563	544	462	604	547	464	611	616	435	547
...	...	...	...	...	...	...	...	...	...	...	...	...	...
90	3	548	482	603	564	384	668	529	470	569	657	453	529
91	3	815	710	908	781	720	961	821	732	900	815	579	647
...	...	...	...	...	...	...	...	...	...	...	...	...	...
120	3	917	846	873	928	868	962	917	834	1004	917	486	457
121	4	898	845	961	912	860	860	880	855	949	898	467	222
122	4	916	843	843	905	832	1027	912	810	1005	916	452	455
...	...	...	...	...	...	...	...	...	...	...	...	...	...
130	4	917	846	873	928	868	962	917	834	1004	917	486	457
131	4	815	710	908	781	720	961	821	732	900	815	579	647
...	...	...	...	...	...	...	...	...	...	...	...	...	...
160	4	548	482	603	564	384	668	529	470	569	657	453	529

Обучающая выборка для сети с использованием частот лидирующих формант (рис. 7) содержит 160 строк (обучающих примеров). Она включает в себя четыре набора строк (по одному набору на одного человека). При этом каждый набор состоит из двух частей: данных для обучения ИНС на распознавание легального пользователя, предоставившего свой идентификатор (примеры 1 — 10, 41 — 50, 81 — 90, 121 — 130), и данных для обучения распознаванию несанкционированных действий администраторов в

случаях их попыток осуществить аутентификацию под идентификатором другого администратора (примеры 11 — 40, 51 — 80, 91 — 120, 131 — 160) с целью получения его привилегий.

Таблица 2. Выходные данные обучающей выборки для ИНС, базирующейся на характеристиках четвёртой форманты и частотах лидирующих формант гласных звуков «А», «О» и «Э»

№ строки	Номер			
	1	2	3	4
1	1	0	0	0
2	1	0	0	0
...	...	...	...	...
10	1	0	0	0
11	0	1	0	0
12	0	1	0	0
...	...	...	...	...
19	0	1	0	0
20	0	1	0	0
21	0	0	1	0
...	...	...	...	...
30	0	0	1	0
31	0	0	0	1
...	...	...	...	...
40	0	0	0	1
41	0	1	0	0
42	0	1	0	0
...	...	...	...	...
50	0	1	0	0
51	1	0	0	0
...	...	...	...	...
80	0	0	0	1
81	0	0	1	0
82	0	0	1	0
...	...	...	...	...
90	0	0	1	0
91	1	0	0	0
...	...	...	...	...
120	0	0	0	1
121	0	0	0	1
122	0	0	0	1
...	...	...	...	...
130	0	0	0	1
131	1	0	0	0
...	...	...	...	...
160	0	0	1	0

Разработанная ИНС на 160 примерах обучалась, на 160 примерах тестировалась или верифицировалась.

Обучающая выборка для сети с использованием только характеристик четвёртой форманты гласных звуков идентична вышеописанной, в ней лишь отсутствуют столбцы, соответствующие входам, на которые подаются частоты лидирующих формант гласных звуков.

5. Результаты обучения и тестирования

Создание и обучение ИНС осуществлялось при помощи программного комплекса Matlab R2015b.

Настройка сети, базирующейся только на характеристиках четвёртой форманты, произошла за 119 эпох. Величина лучшей среднеквадратичной ошибки сети составила $2,1152 \times 10^{-8}$.

Сеть, базирующаяся на характеристиках четвёртой форманты и частотах лидирующих формант гласных звуков «А», «О» и «Э», настроилась за 16 эпох. Величина лучшей среднеквадратичной ошибки сети составила $5,7485 \times 10^{-12}$, что значительно ниже, чем у ИНС, использующей только характеристики четвёртой форманты.

Вычисление ошибок 1-го и 2-го рода проводилось при помощи графического интерфейса nntool программного комплекса Matlab R2015b.

Среднее значение ошибки 1-го рода для сети, базирующейся на характеристиках четвёртой форманты гласных звуков «А», «О» и «Э» составило 10%, так как в 90% случаев легальный пользователь успешно прошёл аутентификацию (табл. 3), среднее значение ошибки 2-го рода – 5% (табл. 4).

Таблица 3. Процентное значение пропуска легального пользователя при использовании ИНС, базирующейся на характеристиках четвёртой форманты гласных звуков «А», «О» и «Э»

№ пользователя	1	2	3	4	Среднее значение:
Итого	90%	70%	100%	100%	90%

Таблица 4. Результаты вычисления ошибки 2-го рода при использовании ИНС, базирующейся на характеристиках четвёртой форманты гласных звуков «А», «О» и «Э»

№ пользователя	1	2	3	4	Среднее значение:
Итого	16,67%	3,33%	0,00%	0,00%	5,00%

Среднее значение ошибки 1-го рода у сети, базирующейся на характеристиках четвёртой форманты и частотах лидирующих формант гласных звуков «А», «О» и «Э», составило 7,5% (табл. 5).

Таблица 5. Процентное значение пропуска легального пользователя при использовании ИНС, базирующейся на характеристиках четвёртой форманты и частотах лидирующих формант гласных звуков «А», «О» и «Э»

№ пользователя	1	2	3	4	Среднее значение:
Итого	90%	80%	100%	100%	92,5%

Среднее значение ошибки 2-го рода составило 5% (табл. 6).

Таблица 6. Результаты вычисления ошибки 2-го рода при использовании ИНС, базирующейся на характеристиках четвёртой форманты и частотах лидирующих формант гласных звуков «А», «О» и «Э»

№ пользователя	1	2	3	4	Среднее значение:
Итого	13,33%	6,67%	0%	0%	5%

Благодаря одновременной подаче характеристик четвёртой форманты трех выбранных гласных звуков и использованию частот лидирующих формант этих же звуков удалось снизить ошибки даже при аутентификации пользователей, имеющих похожие голоса.

Согласно результатам проведённых исследований, целесообразно использовать ИНС, базирующуюся на характеристиках четвёртой форманты и частотах лидирующих формант гласных звуков «А», «О» и «Э». Введение дополнительного речевого параметра (частоты лидирующей форманты) во втором варианте построения ИНС позволило уменьшить среднее значение ошибки 2-го рода на 2,5% по сравнению с первым вариантом архитектуры ИНС. Оба варианта построения сетей прошли исследование на данных пользователей, имеющих близкие спектральные характеристики, то есть полученные результаты соответствуют наихудшему случаю. Полученные величины ошибок в обеих сетях являются приемлемыми по сравнению с величинами ошибок распознавания, основанного на использовании других методов и речевых параметров [4, 20].

Заключение

Аутентификация по голосу пользователей с похожими спектральными характеристиками является одной из важных задач распознавания личности. Искусственная нейронная сеть, базирующаяся на характеристиках четвёртой форманты гласных звуков «А», «О» и «Э», и искусственная нейронная сеть, базирующаяся на характеристиках четвёртой форманты и частотах лидирующих формант гласных звуков «А», «О» и «Э», могут стать одними из инструментов её решения. Полученные значения ошибок 1-го и 2-го рода у этих сетей показывают результативность идеи и возможность проведения дальнейших исследований в данном направлении. Планируется проведение исследований с целью распознавания «чужого».

СПИСОК ЛИТЕРАТУРЫ:

1. Мировой рынок биометрических систем 2015-2022 гг. Обзор рынка, январь 2017 г. // URL: http://json.tv/ict_telecom_analytics_view/mirovoy-rynok-biometricheskih-sistem-2015-2022-gg-20170119025618, свободный (дата обращения: 14.05.2019).
2. Сорокин В.Н. Верификация диктора по спектрально-временным параметрам речевого сигнала / В.Н. Сорокин, А.И. Цыплихин // Информационные процессы. 2010. Т. 10, №2. С. 87–104.
3. Актуальные аспекты информационной безопасности / под ред. О.Б. Макаревича: Монография. – Таганрог: Издательство ТТИ ЮФУ, 2011. – 448 с.
4. Сорокин В.Н. Распознавание личности по голосу: аналитический обзор [Текст] / В.Н. Сорокин, В.В. Вьюгин, А.А. Тананыкин // Информационные процессы. Том 12, №1. Институт проблем передачи информации, Российская академия наук, Москва. 2012. С. 1–30.
5. Первушин Е.А. Обзор основных методов распознавания дикторов / Е.А. Первушин // Математические структуры и моделирование. 2011. Вып. 24. С. 41–54.
6. Belova Ye.P., Mashkina I.V., Research Results of Artificial Neural Network for User Authentication According to Frequency of Fourth Formant of Vowel Sound Phoneme // Сборник научных трудов «2018 International Russian Automatisatation Conference (RusAutoCon)», издательство: институт IEEE. DOI: 10.1109/RUSAUTOCON.2018.8501680, 2018 г. URL: <https://ieeexplore.ieee.org/document/8501680> (дата обращения: 14.05.2019).
7. Новиков А.С., Нестеров К.С. Биометрическая система аутентификации с использованием голосовых данных / А.С. Новиков, К.С. Нестеров // Известия ТулГУ. Технические науки. 2016. №2. URL: <https://cyberleninka.ru/article/n/biometricheskaya-sistema-autentifikatsii-s-ispolzovaniem-golosovyh-dannyh> (дата обращения: 05.05.2019).
8. Матвеев Ю.Н. Исследование информативности признаков речи для систем автоматической идентификации дикторов [Текст] / Ю.Н. Матвеев // Известия вузов. Приборостроение. 2013. Т. 56, №2. С. 47–51.
9. Матвеев Ю.Н., Симончик К.К. Система идентификации дикторов по голосу для конкурса NIST SRE 2010 // Тр. 20-й Междунар. конф. по компьютерной графике и зрению «ГрафиКон 2010». СПб: СПбГУ ИТМО, 2010. С. 315–319.
10. Рабинер Л.Р., Шафер Р.В. Цифровая обработка речевых сигналов. – М.: Радио и связь, 1981. – 496 с.
11. Сидоренко И.А., Кускова П.А. О спектральном анализе фонем с использованием звуковых редакторов [Текст] / Научные ведомости БелГУ, серия История. Политология. Экономика. Информатика. 2013, №22 (165). С. 246–250.

12. Сорокин В.Н. Теория речеобразования. – М.: Радио и связь, 1985. – 312 с.
13. Оганесян А.Г. Эффективный алгоритм выделения формант из спектра речевого сигнала / А.Г. Оганесян // Вычислительная техника и информатика Том LIX, №1. Известия НАН РА и ГИУА, Москва. 2006. С. 177–183.
14. Лукин А. Введение в цифровую обработку сигналов [Электронный ресурс]. URL: <http://audio.rightmark.org/lukin/dspcourse/dspcourse.pdf> свободный (дата обращения: 26.04.2019).
15. Герасимов В.В., Белова Е.П., Машкина И.В. Выделение характеристик четвёртой форманты гласного звука // Свидетельство о государственной регистрации программы для ЭВМ от 3 апреля 2019 года №2019614367.
16. Bard 0.1.7. URL: <http://psi-logic.narod.ru/bard/bard.htm>, свободный (дата обращения: 26.04.2019).
17. Рутковский, Л. Методы и технологии искусственного интеллекта / Лешек Рутковский. – Пер. в польского И.Д. Рудинского. – М.: Горячая линия-Телеком, 2010. – 520 с.
18. Осовский С. Нейронные сети для обработки информации / Пер. с польского И. Д. Рудинского. – М.: Финансы и статистика, 2002. – 344 с.
19. Кафаров В.В. и др. К вопросу моделирования и управления непрерывными технологическими процессами с помощью нейронных сетей / В.В. Кафаров, Л.С. Гордеев, М.Б. Глебов, Го Цзинбао // ТОХТ. 1995, №2.
20. Кулибаба О.В., Привалов М.В. Выбор признаков для аутентификации по голосу в компьютеризированной системе контроля доступа [Текст] // Інформаційні управляючі системи та комп'ютерний моніторинг (ІУС та КМ-2010)/ Матеріали І всеукраїнської науково-технічної конференції студентів, аспірантів та молодих вчених – 19-21 травня 2010 р., Донецьк, ДонНТУ. 2010. С. 33–37.

REFERENCES:

- [1] The global market for biometric systems 2015-2022. Market review, January 2017. URL: http://json.tv/ict_telecom_analytics_view/mirovoy-rynok-biometricheskih-sistem-2015-2022-gg-20170119025618, free (accessed: 14.05.2019) (in Russian).
- [2] Sorokin V.N. Speaker verification by spectral parameters of a speech signal. V. N. Sorokin, A. I. Tsiplihin. Informatsionnye protsessi. 2010. Vol. 10, №2. S. 87–104 (in Russian).
- [3] Topical Aspects of Information Security by ed. O. B. Makarevich: Monograph. – Taganrog: Publishing TTI YuFU, 2011. – 448 p. (in Russian).
- [4] Sorokin V.N. Voice Recognition: An Analytical Review. V.N. Sorokin, V.V. Vjugin, A.A. Tananykin. Informatsionnye protsessi. Vol. 12, №1. Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow. 2012. P. 1–30. (in Russian).
- [5] Pervushin Ye.A. Overview of the main speaker recognition methods. Ye.A. Pervushin. Matematicheskie strukturi i modelirovaniye. 2011. Vol. 24. S. 41–54 (in Russian).
- [6] Belova Ye.P., Mashkina I.V. Research Results of Artificial Neural Network for User Authentication According to Frequency of Fourth Formant of Vowel Sound Phoneme. Collection of scientific papers «2018 International Russian Automatisatation Conference (RusAutoCon)», Publisher: IEEE Instituteo DOI: 10.1109/RUSAUTOCON.2018.8501680, 2018. URL: <https://ieeexplore.ieee.org/document/8501680> (accessed: 20.04.2019).
- [7] Novikov A.S., Nesterov K.S. «Biometric authentication system using voice data» [Text]. Izvestiya TulGU. Tehnicheskie nauki, 2016. №2. URL: <https://cyberleninka.ru/article/n/biometricheskaya-sistema-autentifikatsii-s-ispolzovaniem-golosovyh-dannyh> (дата обращения: 05.05.2019) (in Russian).
- [8] Matveev Yu.N. Investigation of the informativeness of speech signs for systems of automatic speaker identification [Text]. Yu.N. Matveev. Izvestiya Vuzov. Priborostroenie. 2013. Vol. 56, №2. S. 47–51. (in Russian).
- [9] Matveev Yu.N., Simonchik K.K. Voice Recognition Identification System for Competition NIST SRE 2010. Proceedings of the 20th Intern. conf. on computer graphics and vision «GrafCon 2010». Saint Petersburg:SPbSU ITMO, 2010. P. 315–319 (in Russian).
- [10] Rabiner L.R., Shafer R.V. Digital processing of speech signals. – М.: Радио и связь, 1981. – 496 s. (in Russian).
- [11] Sidorenko I.A., Kuskova P.A. About spectral analysis of phonemes using sound editors [Text]. Nauchnye vedomosti BelGU, series Istoriya. Politologiya. Economica. Informatika. 2013, №22 (165). S. 246–250. (in Russian).
- [12] Sorokin V.N. Theory of Speech Formation. – Moscow: Radio i svyaz, 1985. – 312 s. (in Russian).
- [13] Oganessian A.G. Efficient algorithm for formant extraction from speech spectrum. Oganessian A.G. Vichislitel'naya tekhnika i informatika Vol. LIX, No1. Izvestiya NAN RA i GIYA, Moscow. 2006. S. 177–183 (in Russian).

- [14] Lukin A. Introduction to digital signal processing [Electronic resource]. URL: <http://audio.rightmark.org/lukin/dspcourse/dspcourse.pdf> free (accessed: 26.04.2019).
- [15] Gerasimov V.V., Belova Ye.P., Mashkina I.V. Isolation of the characteristics of the fourth formant of vowel sound. Certificate of state registration of the computer program of April 3, 2019 No. 201414367.
- [16] Bard 0.1.7. URL: <http://psi-logic.narod.ru/bard/bard.htm>, free (accessed: 26.04.2019).
- [17] Rutkovsky, L. Methods and technologies of artificial intelligence. L. Rutkovsky – Trans. from Polish I.D. Rudinsky. – М.: Goryachaya Liniya-Telekom, 2010. – 520 s. (in Russian).
- [18] Osovsky S. Neural networks for information processing. Trans. from Polish I.D. Rudinsky. –М.: Finansy i Statistika, 2002. – 344 s. (in Russian).
- [19] Kafarov V.V. et al. On the issue of modeling and control of continuous technological processes using neural networks. Kafarov V.V., Gordeev L.S., Glebov M.B., Guo Jingbao. TOXT. 1995. No2 (in Russian).
- [20] Kulibaba O.V., Privalov M.V. The Selection of Characteristics for Voice Authentication in a Computerized Access Control System students, students and young people – 19-21 grass 2010, Donetsk, DonNTU. 2010. P. 33–37.

*Поступила в редакцию – 19 мая 2019 г. Окончательный вариант – 13 сентября 2019 г.
Received – May 19, 2019. The final version – September 13, 2019.*