

УДК 004.934.2

Д.И. Новохрестова, Е.Ю. Костюченко, Е.С. Катаева, А.Г. Буймов, Д.Н. Буинцев**Программный комплекс по оценке разборчивости речи**

Представлен программный комплекс, предназначенный для оценки качества речи на основе оценки разборчивости произношения слогов. Описаны изменения в структуре базы данных и ее наполнении, необходимом для функционирования программного комплекса, разработанные алгоритмы и функции программного продукта, а также изменения в модулях записи сеанса и оценивания слогов.

Ключевые слова: оценка качества речи, качество произнесения слогов, временная нормализация, программный комплекс, алгоритм детектирования голосовой активности.

doi: 10.21293/1818-0442-2019-22-3-43-48

В настоящий момент выявлено более 100 000 случаев раковых заболеваний органов речеобразующего аппарата, и каждый год это число увеличивается на 25 000 [1, 2]. Самый распространенный вариант лечения включает в себя хирургическое вмешательство, после которого необходимо пройти речевую реабилитацию. Очевидные недостатки метода оценки качества восстановления речи, основанного на ГОСТ Р 58040–95 [3], привели к необходимости разработки алгоритмов автоматической оценки и их реализации в рамках автоматизированных систем по оценке разборчивости речи. В настоящий момент существует ряд работ по автоматизации оценки разборчивости речи [4–6], однако все они появились позднее начала разработки и появления основных алгоритмов рассматриваемого программного комплекса. В рамках исследования по восстановлению речи с использованием технических методов такие алгоритмы были разработаны и реализованы. Они были включены в разработанный программный комплекс, описанный в [7]. После предварительного тестирования было принято решение о необходимости доработки программного комплекса.

Для внедрения программного комплекса по оценке качества речи в процесс речевой реабилитации больных в рамках лечения онкологических заболеваний в НИИ онкологии г. Томска требовалось решить следующие задачи:

1. Доработка структуры базы данных для возможности хранения нескольких вариантов оценки слоговой разборчивости, для хранения сегментации сигнала, возможность указывать тип сеанса (не только слоговая или фразовая разборчивость, но и отметка о моменте записи сеанса: до или после операции).
2. Добавление в базы данных разработанных ранее списков слогов для записи.
3. Разработка и внедрение в модуль оценивания алгоритмов подсчета оценок на основе ранее описанных подходов [8, 9] и алгоритма сегментации для возможности оценивать отдельные фонемы в слогах.
4. Доработка модуля записи слогов: разработка более точного алгоритма детектирования голосовой активности, добавление возможности настройки параметров алгоритма детектирования голосовой

активности, доработка интерфейса программы для возможности коррекции процесса записи слогов.

5. Добавление возможности оценки качества произношения слогов в режиме реального времени в процессе записи сеанса.

База данных

Сама база данных до изменения структуры представлена в [10]. В данной статье опишем только внесенные изменения. Ввиду необходимости в хранении нескольких оценок для одной записи слога в таблицу Slog были добавлены 5 полей для хранения оценок (типа double), а также поле для хранения сегментации слога (тип varchar). Сегментация представляет собой набор значений, обозначающих границу между фонемами, разделенными знаком табуляции. Изменения в таблице Slog представлены в табл. 1.

Таблица 1

Измененная структура таблицы Slog в базе данных

Название поля	Тип данных
Id_zap	INT(11)
Id_seans	INT(11)
Id_slog	INT(11)
Path	VARCHAR(255)
Listen	VARCHAR(255)
Correct1	DOUBLE(255,4)
Correct2	DOUBLE(255,4)
Correct3	DOUBLE(14,4)
Correct4	DOUBLE(14,4)
Correct5	DOUBLE(14,4)
Segmentation	VARCHAR(255)

Также в таблицу Seans, где хранится информация о всех имеющихся сеансах, было добавлено поле type (тип varchar) для возможности внесения комментариев или дополнительной информации о сеансе (предполагается, что в этом поле будет указано «время» записи сеанса, а именно до или после операции).

Для возможности сегментирования записей слогов в таблицу Gostslog_tr добавлено поле position_trouble (тип int), в котором отражено, на каком месте в рамках данного слога стоит проблемная фонема.

Помимо изменений в структуре базы данных, были добавлены наборы слогов, по которым в

настоящий момент можно проводить записи и оценку сеансов. В настоящий момент в базе данных для слоговой разборчивости имеется 5 наборов, по трем из которых возможна оценка сеанса. В табл. 2 представлено описание наборов слоговой оценки.

Таблица 2

Описание наборов слогов в базе данных

Название набора	Кол-во	Описание
Гост1	250	Первые 250 слогов из ГОСТ Р 58040–95
Счет	10	Числа от нуля до девяти
ОнкоЯзык	90	Набор слогов с проблемными фонемами [к][с][т] и их мягкими вариантами: по 5 слогов на каждое возможное расположение фонемы в слоге (в начале, в середине, в конце)
ОнкоЯзыкНК	60	Набор слогов с проблемными фонемами [к][с][т] и их мягкими вариантами: по 5 слогов на каждое крайнее расположение фонемы в слоге (в начале, в конце)
ОнкоЯзыкН	30	Набор слогов с проблемными фонемами [к][с][т] и их мягкими вариантами: по 5 слогов с расположением фонемы в начале слога

Наборы слогов ОнкоЯзыкНК и ОнкоЯзыкН были введены ввиду ограничений алгоритма сегментации (в данный момент он работает только при расположении фонемы в начале слога при любом количестве фонем в слоге и при расположении фонемы в конце слога при наличии трех фонем в слоге). Для каждого из наборов слогов также были добавлены транскрипции каждого слога из набора и номера проблемных фонем в них.

Параметры алгоритма детектирования голосовой активности и их настройка

Используемый в программе алгоритм детектирования голосовой активности работает на основе вычисления двух параметров записанного в буфер сигнала: мера спектральной плоскостности SMF (1) и энергия E :

$$SMF = 10 \times \log_{10} \frac{Gm}{Am}, \quad (1)$$

где Gm – среднее геометрическое спектра речи, Am – среднее арифметическое спектра речи.

Алгоритм принимает решение о наличии или отсутствии речи в буфере сигнала следующим образом: вычисляются параметры для буфера сигнала, сравниваются с минимальными значениями, если разница между полученными значениями и минимальные больше пороговых значений, то принимается решение о наличии голоса на данном отрезке. Для возможности настройки параметров алгоритма для различных записывающих устройств была добавлена новая форма, позволяющая провести запись звука без голоса (запись «тишины») и запись с постоянным присутствием голоса и на основе полученных параметров (минимальные значения и пороговые значения) проводить дальнейшую запись.

При построении модели шума проводится измерение уровня энергии сигнала на записях без голоса и с голосом, меры спектральной плоскостности, количества проанализированных при их подсчете фреймов и длительности записей. По итогам обработки проводится вычисление пороговых значений энергии сигнала и меры спектральной плоскостности.

Также была добавлена возможность корректировать пороговые значения в зависимости от чувствительности микрофона и речи пациента. Для этого было добавлено поле множителя, при помощи которого редактируется порог. В настоящий момент оптимальный множитель для чувствительных микрофонов 0,5, для остальных – 0,25. Оптимальное время для каждого из видов записи 4–8 с.

Настройка параметров записи голоса доступна как из главного окна приложения (окна со списком пациентов), так и из окна, в котором непосредственно производится запись сеанса.

Модуль записи слогов

Предполагается, что запись сеанса будет происходить в автоматизированном режиме, а именно логопед (или пациент) будет участвовать только в настройках записи, а обработка записей, подсчет оценки, если необходимо, и переключение на следующий слог будут автоматическими. Форма записи слогов представлена на рис. 1. В данной версии программного комплекса возможна настройка следующих параметров:

1. Возможность воспроизведения слога. Для наборов слогов ОнкоЯзык, ОнкоЯзыкНК, ОнкоЯзыкН предусмотрена возможность воспроизведения записанного эталонного произношения слога через наушники или колонки. Данная возможность должна помочь пациенту понять, как именно необходимо произнести отображенный на экране слог.

2. Отображение слога и его транскрипции на экране. В предыдущей версии программы и слог, и его транскрипция отображались с одинаковым размером (примерно 25 кеглей). Однако после первого тестового запуска программы в НИИ онкологии было принято решение о существенном увеличении шрифта для слога и уменьшении для транскрипции ввиду того, что пациент не понимал соответствия между слогом и транскрипцией. Также была возможность еще увеличить или уменьшить размер шрифта для отображения слога.

В рамках программного комплекса [7] при неправильном срабатывании алгоритма детектирования голосовой активности возврат к предыдущему слогу осуществлялся по нажатию клавиш Ctrl+Z, в новой версии программы была добавлена кнопка «Назад», выполняющая те же функции. Замена сочетания клавиш на кнопку обосновывается тем, что назначение кнопки с подобной надписью интуитивно понятно пользователю системы, что упрощает работу с программой. Кнопка «Следующий слог» остается для записей слогов пациентов, у которых после операции речь ухудшилась до состояния, при

котором невозможно корректно определить начало и конец речи. В этом случае параметры чувствительности алгоритма будут изменены на минимальные,

запись звука будет осуществляться постоянно, а переключение между слогами будет осуществляться специалистом-логопедом.

Рис. 1. Форма записи слогов

Для оценки вариативности речи до операции необходимо записать два сеанса, которые в дальнейшем будут использоваться как эталон. Однако для знакомства пациента с программным комплексом и облегчения дальнейшей работы с ним предлагается записывать три (и более, по необходимости) сеанса, а в качестве эталона выбрать два из них.

При записи сеансов после операции появляется возможность оценивать сеанс непосредственно во время записи. Для этого необходимо отменить «Оценивать слоги при записи», после этого станет доступна кнопка «Выбор сеанса для сравнения». Если отмечена необходимость в оценке, но при этом не выбраны эталонные сеансы, программа выдаст предупреждение и не начнет сеанс записи, пока не решится этот конфликт.

При выборе сеанса для работы выбор осуществляется на основе идентификатора и даты записи. Дополнительно предоставляется информация о типе проводимой оценки и полученных результатах оценки качества произнесения. В настоящий момент можно выбрать один или два сеанса для сравнения. Возможность выбора только одного сеанса для сравнения добавлена из-за того, что есть пациенты, которые в настоящий момент проходят лечение в НИИ Онкологии, но до операции у них получилось записать только один сеанс. Для выбора доступны сеансы только такого же типа (записанные по тому же набору слогов), что и записываемый сеанс.

Модуль оценки слогов

В настоящем комплексе предусмотрено три различных способа оценки разборчивости слога. Все три вычисляются как отношение среднего коэффи-

циента корреляций между парами эталонное произношение–оцениваемое произношение (здесь и далее описана ситуация с наличием двух эталонных сеансов, однако все оценки доступны при наличии только 1 сеанса) к коэффициенту корреляции между эталонными произношениями. Таким образом, оценка заключена в интервале от 0 до 1, чем ближе к 1 полученная оценка, тем качественнее, ближе к эталону произнесена фонема. В первом и третьем способах трансформация записей для подсчета коэффициента корреляции осуществляется за счет использования функции Resample из математического пакета Matlab, во втором способе используется алгоритм динамической трансформации временной шкалы. Отличие 1-го способа от 3-го заключается в том, что в первом способе оценивается не весь слог, а только проблемная фонема, поэтому в данном случае перед непосредственной оценкой необходимо произвести сегментацию слога и выделить проблемную фонему. Ввиду того, что алгоритм сегментации работает корректно только при определенных расположениях фонем в слоге, первый способ оценивания доступен только для определенных наборов слогов.

Оценка разборчивости слога возможна не только в процессе записи сеанса, но и для существующего сеанса. Форма просмотра сеанса и его оценки представлена на рис. 2. Для того чтобы провести сегментацию, необходимо нажать кнопку «Провести сегментацию», если для данного типа сеанса функция сегментации недоступна, о этом будет выведено сообщение. Результат сегментации хранится в колонке Segmentation и после нажатия кнопки

«Сохранить» записывается в базу данных. Далее необходимо выбрать настройки оценки, а именно тип оценки и эталонные сеансы. Эталонные сеансы выбираются аналогично форме записи слогов. После выбора эталонных сеансов кнопка «Оценить сеанс» станет доступна и при ее нажатии пройдет последо-

вательная оценка всех слогов в данном сеансе. Результат будет представлен в колонке correct1 и после нажатия кнопки «Сохранить» будет сохранен в базе данных. Также после окончания вывода всех оценок подсчитывается средняя оценка сеанса и также вносится в базу данных.

id_zap	id_seans	id_slog	correct1	correct2	segmentation
12708	1099	1	1		0,162
12709	1099	2	1		0,106
12710	1099	3	1		0,094
12711	1099	4	0,27		0,138
12712	1099	5	0,427		0,085
12713	1099	6	1		0,078
12714	1099	7	0,631		0,136
12715	1099	8	0,506		0,104
12716	1099	9	0,536		0,081
12717	1099	10	1		0,130
12718	1099	11	1		0,216
12719	1099	12	0,485		0,199

Рис. 2. Форма просмотра сеанса и оценки слогов

Заключение

В работе описаны выявленные в процессе тестирования программного комплекса по оценке разборчивости речи проблемные моменты и поставлены задачи по их решению. В рамках решения каждой задачи необходимые функции были реализованы, а также описаны имеющиеся в настоящий момент ограничения по работе с программным комплексом. Доработанный программный комплекс в настоящий момент проходит процедуру внедрения в НИИ онкологии г. Томска, проведены тестовые записи, а также первые записи пациентов. Результатом записей стали рекомендации по дополнительной доработке программного комплекса, а именно:

- добавление полей в таблицу с данными о пациентах для детализации медицинской информации;
- добавление возможности записи отдельных слогов для случаев позднего обнаружения проблем с записями сеанса (к примеру, во время дополнительного прослушивания в рамках слога обнаружен посторонний шум, мешающий его восприятию);
- возможность вывода нескольких оценок, полученных разными способами, в несколько столбцов с оценками слогов.

Дальнейшая модернизация программного комплекса предполагает использование систем распознавания и анализа речи [11–17] для оценки слоговой, словесной и фразовой разборчивости. Для обеспечения безопасности хранимой информации планируется к разработке модуль криптографической защиты, который не препятствовал бы получению оценок в режиме реального времени [18].

Работа выполнена при поддержке Российского научного фонда, проект «Восстановление речевой функции с использованием технических методов и математического моделирования у больных раком

полости рта и ротоглотки после хирургического лечения», № 1615-00038.

Литература

1. Злокачественные новообразования в России в 2016 году (заболеваемость и смертность) / под ред. А.Д. Каприна, В.В. Старинского, Г.В. Петровой. – М.: МНИОИ им. П.А. Герцена – филиал ФГБУ «НМИРЦ» Минздрава России, 2018. – 250 с.
2. Состояние онкологической помощи населению России в 2016 году / под ред. А.Д. Каприна, В.В. Старинского, Г.В. Петровой. – М.: МНИОИ им. П.А. Герцена – филиал ФГБУ «НМИРЦ» Минздрава России, 2017. – 236 с.
3. ГОСТ Р 50840–95. Передача речи по трактам связи. Методы оценки качества, разборчивости и узнаваемости. – М.: ИПК Изд-во стандартов, 1996. – 234 с.
4. Pamula H. et al. Parametric Assessment of Esophageal Speech in Post-Laryngectomy Patients // 2018 Joint Conference-Acoustics. – IEEE. – 2018. – P. 1–5.
5. Kalita S., Mahadeva Prasanna S. R., Dandapat S. Intelligibility assessment of cleft lip and palate speech using Gaussian posteriograms based on joint spectro-temporal features // The Journal of the Acoustical Society of America. – 2018. – Vol. 144, No. 4. – P. 2413–2423.
6. Sumita Y.I. et al. Digitised evaluation of speech intelligibility using vowels in maxillectomy patients / Y.I. Sumita, M. Hattori, M. Murase, M.E. Elbashti, H. Taniguchi // Journal of oral rehabilitation. – 2018. – Vol. 45, No. 3. – P. 216–221.
7. Костюченко Е.Ю. Программное обеспечение для объективной оценки качества произнесения слогов при речевой реабилитации / Е.Ю. Костюченко, Р.В. Мещеряков, Д.И. Новохрестова, А.В. Пятков, Л.Н. Балацкая, Е.Л. Чойзонов // Всерос. науч. конф. по проблемам управления в технических системах. Санкт-Петербургский гос. электротехн. Ун-т ЛЭТИ им. В.И. Ульянова (Ленина), 2017. – № 1. – С. 277–280.
8. Kostyuchenko E. Correlation normalization of syllables and comparative evaluation of pronunciation quality in speech rehabilitation / E. Kostyuchenko, R. Meshcheryakov,

D. Ignatieva, A. Pyatkov, E. Choyzonov, L. Balatskaya // International Conference on Speech and Computer. – Springer, Cham, 2017. – P. 262–271.

9. Новохрестова Д.И. Временная нормализация слогов алгоритмом динамической трансформации временной шкалы при оценке качества произнесения слогов в процессе речевой реабилитации // Доклады ТУСУР. – 2017. – Т. 20, № 4. – С. 142–145. DOI: 10.21293/1818-0442-2017-20-4-142-145

10. Костюченко Е.Ю. Формирование базы данных пациентов при речевой реабилитации после комбинированного лечения онкологических заболеваний органов речеобразующего тракта / Е.Ю. Костюченко, Д.И. Новохрестова, А.В. Пятков // Электронные средства и системы управления. – Томск: ТУСУР, 2017. – №2. – С. 245–247.

11. Кипяткова И.С., Карпов А.А. Разновидности глубоких искусственных нейронных сетей для систем распознавания речи // Труды СПИИРАН. – 2016. – Т. 6, № 49. – С. 80–103.

12. Кипяткова И.С., Карпов А.А. Аналитический обзор систем распознавания русской речи с большим словарем // Труды СПИИРАН. – 2010. – Т. 12, №. 1. – С. 7–20.

13. Ishikawa K., MacAuslan J., Boyce S. Toward clinical application of landmark-based speech analysis: Landmark expression in normal adult speech // The Journal of the Acoustical Society of America. – 2017. – Vol. 142, No. 5. – P. EL441–EL447.

14. Hossain M.S., Muhammad G. Emotion recognition using secure edge and cloud computing // Information Sciences. – 2019. – Vol. 504. – P. 589–601.

15. Kaya H., Karпов A.A., Salah A.A. Fisher vectors with cascaded normalization for paralinguistic analysis // Sixteenth Annual Conference of the International Speech Communication Association. – 2015.

16. Криптографические протоколы в системах с ограниченными ресурсами / С.К. Росошек, Р.В. Мещеряков, А.А. Шелупанов, М.А. Соськин // Вычислительные технологии. – 2007. – Т. 12, № S1. – С. 51–61.

17. Rakhmanenko I. Fusion of biLSTM and GMM-UBM systems for audio spoofing detection / I. Rakhmanenko, A. Shelupanov, E Kostyuchenko // International Journal of Advanced Trends in Computer Science and Engineering. – 2019. – Vol. 8, No. 4. – P. 1741–1746.

18. The evaluation process automation of phrase and word intelligibility using speech recognition systems / E. Kostyuchenko, D. Novokhrestova, M. Tirskaaya, A. Shelupanov, M. Mirovich-Danchenko, E. Choyzonov, L. Balatskaya // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). – 2019. – Vol. 11658 LNAI. – P. 237–246

Новохрестова Дарья Игоревна

Преп. каф. комплексной информационной безопасности электронно-вычислительных систем (КИБЭВС) Томского государственного университета систем управления и радиоэлектроники (ТУСУР) Ленина пр-т, д. 40, г. Томск, Россия, 634050
Тел.: +7 (382-2) 70-15-29 (внут. 29-66)
Эл. почта: ndi@fb.tusur.ru

Костюченко Евгений Юрьевич

Канд. техн. наук, доцент каф. КИБЭВС ТУСУР
Ленина пр-т, д. 40, г. Томск, Россия, 634050
Тел.: +7 (382-2) 70-15-29 (внут. 29-66)
Эл. почта: key@keva.tusur.ru

Катаева Елена Сергеевна

Ст. преп. каф. КИБЭВС ТУСУР
Ленина пр-т, д. 40, г. Томск, Россия, 634050
Тел.: +7 (382-2) 70-15-29 (внут. 29-66)
Эл. почта: kes@keva.tusur.ru

Буймов Аркадий Георгиевич

Д-р техн. наук, профессор каф. экономики ТУСУР
Ленина пр-т, д. 40, г. Томск, Россия, 634050
Тел.: +7 (382-2) 70-15-29 (внут. 2966)
Эл. почта: arkadii.g.buimov@tusur.ru

Буинцев Дмитрий Николаевич

Канд. техн. наук, проректор по развитию университетского комплекса и социальной работы ТУСУР
Ленина пр-т, д. 40, г. Томск, Россия, 634050
Тел.: +7 (382-2) 90-01-01
Эл. почта: buintsev-dn@tusur.ru

Novokhrestova D.I., Kostyuchenko E.Y., Kataeva E.S., Buimov A.G., Buintsev D.N.

Software package for assessing speech intelligibility

The article presents a software package designed to assess the quality of speech based on the evaluation of the syllables pronunciation intelligibility. Changes in the structure of the database and its filling, necessary for the functioning of the program complex, were described. The developed algorithms and functions of the software product were presented. Changes in session recording and syllable estimation modules were given.

Keywords: speech quality estimation, syllables pronunciation quality, temporal normalization, software complex, voice activity detection algorithm.

doi: 10.21293/1818-0442-2019-22-3-43-48

References

1. Kaprin A.D., Starinskiy V.V., Petrova G.V. *Zlo-kachestvennyye novoobrazovaniya v Rossii v 2015 godu (zabolevaemost' i smertnost')* [Malignant neoplasms in Russia in 2016 (morbidity and mortality)], 2018, 250 p. (in Russ.)
2. Kaprin A.D., Starinskiy V.V., Petrova G.V. *Sostoyaniye onkologicheskoy pomoshchi naseleniyu Rossii v 2016 godu* [The status of cancer care for the population of Russia in 2016], 2017, 236 p. (in Russ.)
3. *GOST R 50840 – Peredacha rechi po traktam svyazi. Metody ocenki kachestva, razborchivosti i uznavaemosti* [GOST R 50840 – Voice transmission over communication paths. Methods for assessing quality, legibility and recognition], 1996, 234 p. (in Russ.)
4. Pamula H. et al. Parametric Assessment of Esophageal Speech in Post-Laryngectomy Patients *2018 Joint Conference-Acoustics*, IEEE, 2018, pp. 1–5.
5. Kalita S., Mahadeva Prasanna S.R., Dandapat S. Intelligibility assessment of cleft lip and palate speech using Gaussian posteriors based on joint spectro-temporal features *The Journal of the Acoustical Society of America*, 2018, vol. 144, no. 4, pp. 2413–2423.
6. Sumita Y.I. et al. Digitised evaluation of speech intelligibility using vowels in maxillectomy patients *Journal of oral rehabilitation*, 2018, vol. 45, no. 3, pp. 216–221.
7. Kostyuchenko E. et al. Software for the objective assessment of the quality of pronunciation of syllables in speech rehabilitation *All-Russian Scientific Conference on Management Issues in Technical Systems. – Federal State Autonomous Educational Institution of Higher Education St. Petersburg*

State Electrotechnical University LETI named after VI Ulyanov (Lenin), 2017, no. 1, pp. 277–280 (in Russ.)

8. Kostyuchenko E. et al. Correlation normalization of syllables and comparative evaluation of pronunciation quality in speech rehabilitation *International Conference on Speech and Computer*, Springer, Cham, 2017, pp. 262–271.

9. Novokhrestova D. Temporary normalization of syllables by the dynamic transformation algorithm of the timeline in assessing the quality of the pronunciation of syllables in the process of speech rehabilitation *Proceedings of TUSUR University*, 2017, vol. 20, no. 4, pp. 142–145. DOI: 10.21293/1818-0442-2017-20-4-142-145 (in Russ.)

10. Kostyuchenko E., Novokhrestova D., Pyatkov A. Formation of a database of patients with speech rehabilitation after combined treatment of oncological diseases of the organs of the speech-forming tract *Electronic means and control systems*, Tomsk, 2017, no. 2, pp. 245–247 (in Russ.)

11. Kipyatkova I.S., Karpov A.A. Variants of Deep Artificial Neural Networks for Speech Recognition Systems. *Trudy SPIIRAN – SPIIRAS Proceedings*, 2016, vol. 6, no. 49, pp. 80–103 (in Russ.)

12. Kipyatkova I.S., Karpov A.A. Analytical review of Russian speech recognition systems with a large dictionary *Trudy SPIIRAN – SPIIRAS Proceedings*, 2010, vol. 12, no. 1, pp. 7–20 (in Russ.)

13. Ishikawa K., MacAuslan J., Boyce S. Toward clinical application of landmark-based speech analysis: Landmark expression in normal adult speech *The Journal of the Acoustical Society of America*, 2017, vol. 142, no. 5, pp. EL441–EL447.

14. Hossain M.S., Muhammad G. Emotion recognition using secure edge and cloud computing *Information Sciences*, 2019, vol. 504, pp. 589–601.

15. Kaya, Heysem, Alexey A. Karpov, and Albert Ali Salah. Fisher vectors with cascaded normalization for paralinguistic analysis. *Sixteenth Annual Conference of the International Speech Communication Association*. 2015.

16. Rososhek S.K., Meshcheryakov R.V., Shelupanov A.A., Sonkin M.A. *Kriptograficheskie protokoly v sistemah s ogranichennymi resursami* [Cryptographic protocols in systems with limited resources] *Computational Technologies*, vol. 12, no. S1, 2007, pp. 51–61 (in Russ.)

17. Rakhmanenko I., Shelupanov A., Kostyuchenko E. [Fusion of biLSTM and GMM-UBM systems for audio spoofing detection] *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 8, no. 4, 2019, pp. 1741–1746

18. Kostuchenko E., Novokhrestova D., Tirskaaya M., Shelupanov A., Nemirovich-Danchenko M., Choynzonov E., Balatskaya L. [The evaluation process automation of phrase and word intelligibility using speech recognition systems] *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11658 LNAI, 2019, pp. 237–246

Daria I. Novokhrestova

Lecturer, Department of Complex Information Security, Tomsk State University of Control Systems and Radioelectronics (TUSUR)
40, Lenin pr., 634050, Tomsk, Russia
Phone: +7 (382-2) 70-15-29 (main 29-66)
Email: ndi@fb.tusur.ru

Evgeny Y. Kostyuchenko

Candidate of Technical Sciences, Associate Professor, Department of Complex Information Security, Tomsk State University of Control Systems and Radioelectronics (TUSUR)
40, Lenin pr., 634050, Tomsk, Russia
Phone: +7 (382-2) 70-15-29 (main 29-66)
Email: key@keva.tusur.ru

Elena S. Kataeva

Senior Lecturer, Department of Complex Information Security, TUSUR
40, Lenin pr., 634050, Tomsk, Russia
Phone.: +7 (382-2) 70-15-29 (main 29-66)
Email: kes@keva.tusur.ru

Arkady G. Buymov

Dr. Tech. of Sciences, Professor,
Department of Economics, TUSUR
40, Lenin pr., 634050, Tomsk, Russia
Phone: +7 (382-2) 70-15-29 (main 2966)
Email: arkadii.g.buimov@tusur.ru

Dmitry N. Buintsev

Candidate of Technical Sciences, Vice-Rector for University Infrastructure Development and Social Work, TUSUR
40, Lenin pr., 634050, Tomsk, Russia
Phone: +7 (382-2) 90-01-01
Email: buintsev-dn@tusur.ru