

УДК 004.82

ПОДХОД К ФОРМИРОВАНИЮ НАБОРОВ ЭТАЛОНОВ РЕЧЕВЫХ КОМАНД С ИСПОЛЬЗОВАНИЕМ ОНТОЛОГИИ

А.И. Армер¹, В.С. Мошкин²

Ульяновский государственный технический университет, Ульяновск, Россия

¹a.armer@mail.ru, ²PostForVadim@ya.ru

Аннотация

В работе приводится метод формирования библиотеки эталонов речевых команд с использованием предметной онтологии с целью повышения их дальнейшего распознавания. Представлена формальная модель предметной OWL-онтологии, используемой в качестве базы знаний в процессе формирования и расширения библиотеки эталонных команд в процессе распознавания речевых команд, а также описан алгоритм извлечения списка синонимов входных речевых команд на основе предметной OWL-онтологии. Статья содержит описание методики проведения, а также результаты экспериментов по формированию вариантов библиотеки эталонов и выбора библиотеки с наиболее различимыми командами с использованием OWL-онтологии предметной области управления гражданского самолета с помощью речевых команд. Приводятся результаты экспериментов, доказывающие эффективность подхода.

Ключевые слова: *речевые команды, онтология, эталоны, распознавание речи.*

Цитирование: *Армер, А.И. Подход к формированию наборов эталонов речевых команд с использованием онтологии /А.И. Армер, В.С. Мошкин // Онтология проектирования. – 2016. – Т. 6, №3(21). - С. 270-277. – DOI: 10.18287/2223-9537-2016-6-3-270-277.*

Введение

При распознавании речевых команд (РК) из ограниченного словаря в ряде алгоритмов используется сравнение наблюдаемой РК с каждой эталонной РК, и из эталонных находится наиболее близкая к наблюдаемой. С этой эталонной РК наблюдаемая РК признаётся идентичной. В силу различных причин от состава библиотеки эталонных команд в значительной степени зависит вероятность правильного распознавания РК [1, 2].

При распознавании РК большое влияние на вероятность правильного распознавания оказывает различимость команд по длительности произнесения. Обычно по длительности произнесения производится предварительная сортировка эталонных команд, в результате которой в дальнейшем сравнении с наблюдаемой РК могут участвовать не все эталонные [3-5]. Однако, из-за изменчивости длительности произнесения РК даже при идентификации по длительности возможны ошибки первого и второго рода.

Ошибка первого рода возникает тогда, когда после предварительной идентификации в дальнейшей идентификации не участвует эталон, в действительности идентичный наблюдаемой РК. Предварительная идентификация целесообразна, если в результате вероятность ошибки первого рода близка к нулю.

Ошибки второго рода возникают, если в дальнейшей идентификации помимо действительно соответствующего эталона участвуют и другие эталонные РК, и сколько их, столько и ошибок второго рода. Поэтому при сохранении близкой к нулю вероятности ошибки первого рода снижение количества ошибок второго рода приводит к увеличению вероятности правильного распознавания речи [6, 7].

Целью настоящей работы является формирование библиотеки эталонных команд, в которой команды были бы максимально различимы по длительности произнесения, то есть в процессе распознавания давали минимальное количество ошибок второго рода при близкой к нулю вероятности ошибки первого рода, и при этом не терялся тематический и практический смысл библиотеки. Например, если библиотека используется для распознавания команд управления транспортным средством, то команды должны сохранять смысловую связь с вызываемыми ими автоматическими действиями бортового оборудования. Для достижения поставленной цели необходимо использовать некую базу знаний, определяющую семантику каждой РК из библиотеки эталонных команд и расширяющую её за счет формирования выражений, синонимичных РК. Одним из наиболее эффективных подходов при решении данной задачи является использование онтологии.

1 OWL-онтология формирования наборов эталонов РК

Онтология представляет собой систему, состоящую из множества понятий, их определений и аксиом, необходимых для ограничения интерпретации и использования понятий. Любая онтология основана на математическом аппарате – дескрипционной логике (ДЛ) определенного типа.

OWL (*Ontology Web Language*) – язык описания онтологий, позволяющий описывать классы и отношения между ними. В основе языка – представление действительности в модели данных «объект – свойство». OWL является переформулировкой ДЛ с использованием синтаксиса XML.

Онтология предметной области (ПрО) представляет собой совокупность RDF-троек: *subject – predicate – object* (субъект — отношение — объект). В рамках решения поставленной задачи использовалась OWL-онтология [8].

Онтологический подход к хранению знаний в рамках решения задачи получения полной библиотеки аналогов возможных РК предполагает представление их в виде

$$O = \langle T, R, F \rangle,$$

где:

- T – термины прикладной области, которую описывает онтология, при этом

$$T \subset \{T_{OBJ}, T_{ACT}, T_{CHAR}\},$$

где T_{OBJ} – множество классов, описывающих иерархию объектов затрагиваемой ПрО;

T_{ACT} – множество классов, описывающих возможные действия, совершаемые над элементами множества T_{OBJ} ;

T_{CHAR} – множество классов характеристик объектов ПрО или действий над ними.

- R – отношения между терминами заданной ПрО:

$$R \subset \{R_{CMN}, R_{OBJ}, R_{DATA}, R_{ANNOT}\},$$

где R_{CMN} – множество встроенных отношений синонимии и иерархии классов онтологии. Пример определения встроенного отношения идентичности (*sameAs*) разработанной OWL-онтологии:

```
<NamedIndividual rdf:about="&www;связь">
  <rdf:type rdf:resource="&www;Объект"/>
  <sameAs rdf:resource="&www;соединение"/>
  <sameAs rdf:resource="&www;сообщение"/>
</NamedIndividual>
```

R_{OBJ} – множество отношений объектов (Object Properties). Пример определения отношения объектов «частьВыражения»:

```
<NamedIndividual rdf:about="&www;больше">
```

```

<rdf:type rdf:resource="&www;Действие"/>
<www:частьВыражения rdf:resource="&www;мощнее"/>
<sameAs rdf:resource="&www;повысить"/>
</NamedIndividual>

```

R_{DATA} – множество отношений типа данных (Datatype Properties);

R_{ANNOT} – множество отношений аннотации (Annotation Properties).

- F – множество функций интерпретации (аксиоматизации), заданных на терминах и/или отношениях онтологии.

Для решения задачи получения максимального количества вариантов синонимичных РК в библиотеке эталонов была разработана OWL-онтология, включающая команды системы речевого управления режимами работы двигателя самолёта.

OWL-онтология состоит из 3 классов: «Действие», «Объект» и «Характеристика», объекты которых связаны между собой отношениями R_{CMN} и R_{OBJ} . Благодаря наличию этих отношений, онтология позволяет получать различные вариации написания команд, синонимичных анализируемому. Фрагмент онтологии, иллюстрирующий данный процесс, представлен на рисунке 1.

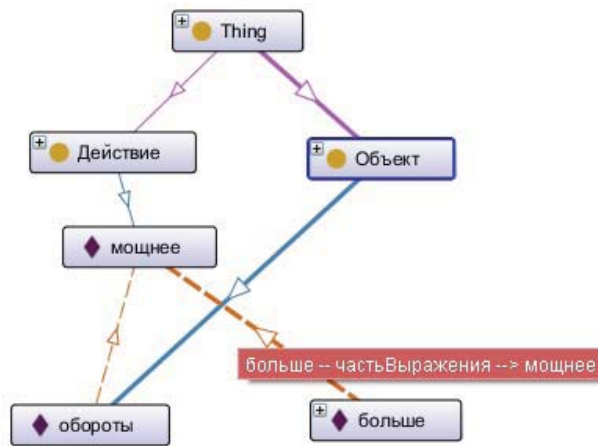


Рисунок 1 – Пример применения отношения «частьВыражения» OWL-онтологии

Как видно из рисунка 1, два объекта онтологии «обороты» и «больше» связаны с объектом «мощнее» класса «Действие» посредством отношения (Object Property) «частьВыражения», благодаря чему при формировании набора эталонов РК «обороты слева (справа) больше» можно заменить на семантически синонимичную РК «мощнее слева (справа)».

Аналогичным образом используются при формировании набора эталонов РК отношение «включает», являющееся обратным (свойство отношения «Inverse Of») отношению «частьВыражения», а также встроенное отношение идентичности «sameAs», которое определяет синонимию непосредственно между двумя объектами онтологии. На данный момент разработанная OWL-онтология содержит 89 объектов классов.

2 Алгоритм извлечения списка синонимов входных речевых команд на основе предметной OWL-онтологии

OWL-онтология позволяет, помимо непосредственного описания основных объектов рассматриваемой ПрО, связывать их между собой, благодаря чему в процессе обработки какого-либо семантически значимого выражения (в нашем случае им является РК), появляется возможность извлечения синонимичных выражений, необходимых при решении задачи определения и наполнения библиотеки эталонных команд.

Таким образом, алгоритм получения списка синонимов выражения на основе предметной OWL-онтологии включает в себя следующие этапы.

- 1) разбиение входного выражения на семантически значимые сочетания - синтагмы (с учётом употребления предлогов): $P: a \rightarrow B$, где a - входное выражение, B - множество семантически значимых единиц, на которое данное выражение разбивается.
- 2) поиск полученных объектов в предметной OWL-онтологии.
- 3) генерация набора выражений посредством смены порядка слов в сочетаниях $G_1: B \rightarrow B'$, $B' \subseteq \{B'_1, B'_2 \dots B'_n\}$, где B' - множество выражений, полученных из входного путём смены порядка семантически значимых единиц.
- 4) поиск синонимов для каждого слова входного сочетания: $S_1: B \rightarrow B_S$, где B_S - множество синонимов для каждой семантически значимой единицы входного выражения.
- 5) генерация набора выражений посредством смены порядка слов и с учётом их замены на синонимы: $G_2: B \rightarrow B'_S$, $B'_S \subseteq \{B'_{S_1}, B'_{S_2} \dots B'_{S_m}\}$, где B'_S - множество выражений, полученных из входного путём смены порядка слов с учётом их замены на синонимы.
- 6) поиск синонимов устойчивых выражений во входном выражении (количество слов может не совпадать): $S_2: B \rightarrow B_{EXP}$, где B_{EXP} - множество синонимов устойчивых выражений, содержащихся во входном выражении.
- 7) генерация набора выражений посредством смены порядка слов и с использованием синонимов устойчивых выражений: $G_3: B \rightarrow B'_{EXP}$, $B'_{EXP} \subseteq \{B'_{EXP_1}, B'_{EXP_2} \dots B'_{EXP_m}\}$, где B'_{EXP} - множество выражений, полученных из входного путём смены порядка слов с использованием найденных синонимов устойчивых выражений.
- 8) построение полного списка полученных синонимов: $B_{ALL} \subseteq \{B', B'_S, B'_{EXP}\}$.

Схематично данный алгоритм представлен на рисунке 2.

Таким образом, в результате анализа входного выражения, являющегося прототипом некоторой речевой команды, алгоритм формирует на основе анализа отношений в предметной онтологии список всех возможных синонимов данного выражения, тем самым расширяя возможность правильного выбора команды на этапе распознавания.

3 Создание вариантов библиотеки эталонов и выбор библиотеки с наиболее различимыми командами

Схема эксперимента по выявлению наиболее различимых по длительности команд, составляющих библиотеку эталонов, с использованием предметной OWL-онтологии выглядит следующим образом.

- 1) выбирается набор команд, рассматриваемый в качестве исходной библиотеки.
- 2) разрабатывается предметная OWL-онтология, позволяющая получать синонимы для команд из этой библиотеки.
- 3) для каждой команды из этой библиотеки с помощью OWL-онтологии создаётся максимально возможное количество синонимов так, что каждая команда из исходной библиотеки образует набор команд — синонимичных вариантов. Всевозможные сочетания без повторений команд из образованных наборов дают новые библиотеки эталонных команд.
- 4) для каждой новой библиотеки подготавливаются наборы тестовых команд для распознавания по длительности произнесения.
- 5) проводится эксперимент, в котором тестовые команды идентифицируются по длительности.

- б) выявляется библиотека, для которой в результате идентификации получено минимальное количество ошибок второго рода при отсутствии ошибок первого рода.

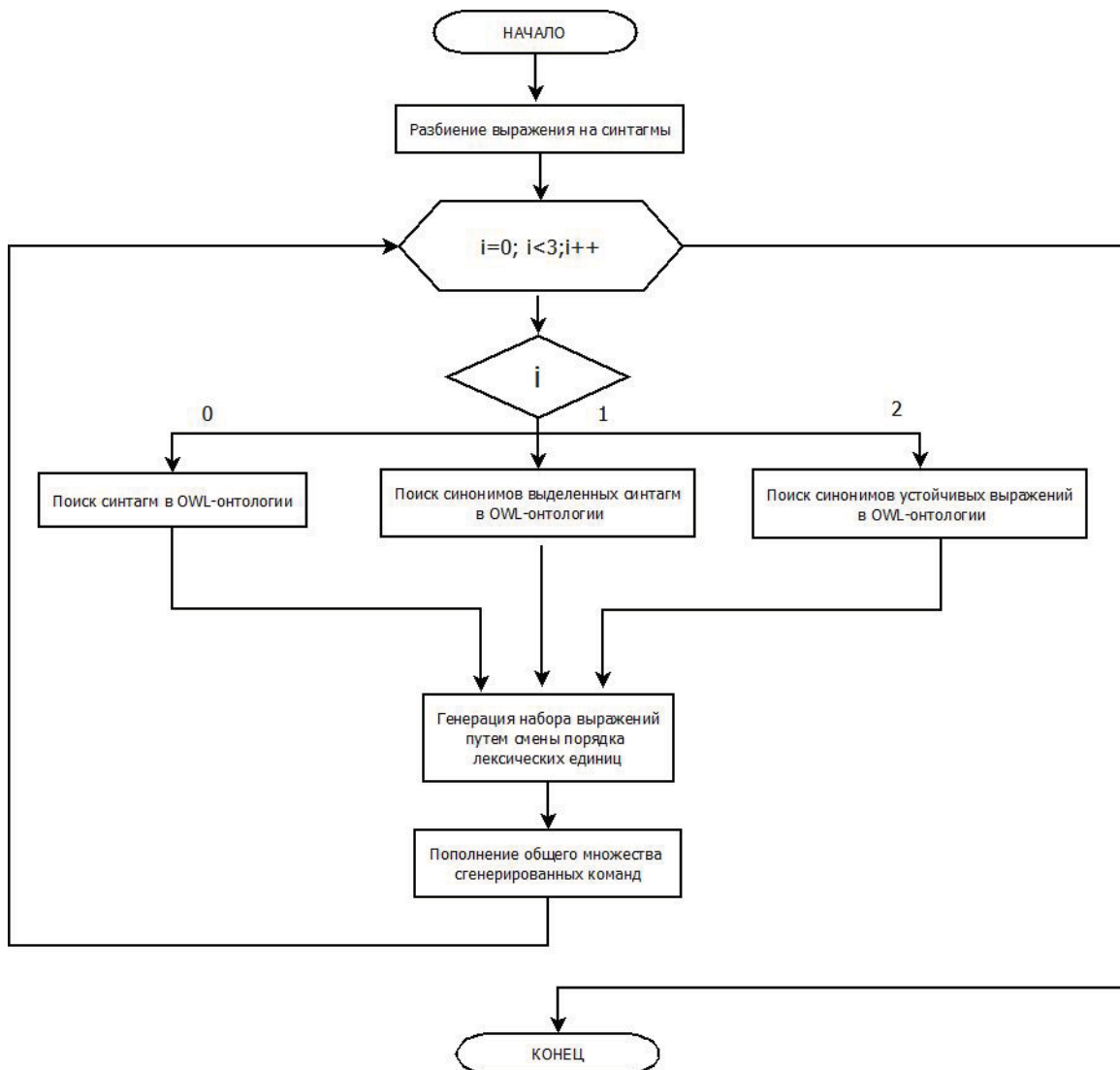


Рисунок 2 – Схема алгоритма расширения библиотеки эталонных команд

Для проведения испытаний взята библиотека из восьми команд, предназначенных для системы речевого управления режимами работы двигателя самолёта (см. таблицу 1). Для каждой команды с использованием алгоритма получения синонимов эталонных команд сгенерировано своё количество синонимов. Например, для команды «обороты слева меньше» сгенерировано 24 синонима, группа этих синонимов приведена в таблице 2.

Если из каждой группы брать по одному синониму, то получится 897122304 различных комбинаций синонимов, называемых сочетаниями. Каждое сочетание рассматривалось как библиотека эталонных РК. Для каждой из 897122304 библиотек определено количество ошибок идентификации по длительности произнесения. Та библиотека, для которой определено наименьшее количество ошибок второго рода, была признана «оптимальной». Идентификация производилась путём сравнения длины распознаваемой РК с длинами эталонных команд.

Таблица 1 – Формирование наборов эталонов речевых команд

№	Команда	Количество синонимов	Номер синонима в «оптимальной» библиотеке
1.	обороты слева меньше	24	1
2.	обороты слева больше	26	25
3.	обороты справа меньше	24	1
4.	обороты справа больше	26	25
5.	турбостартер слева	4	4
6.	турбостартер слева выключить	12	3
7.	турбостартер справа	4	3
8.	турбостартер справа выключить	12	5

Таблица 2 – Перечень синонимов команды «обороты слева меньше»

№	Синонимы	№	Синонимы
1.	обороты слева меньше	13.	мощность слева меньше
2.	обороты меньше слева	14.	мощность меньше слева
3.	слева обороты меньше	15.	слева мощность меньше
4.	слева меньше обороты	16.	слева меньше мощность
5.	меньше обороты слева	17.	меньше мощность слева
6.	меньше слева обороты	18.	меньше слева мощность
7.	обороты слева снизить	19.	мощность слева снизить
8.	обороты снизить слева	20.	мощность снизить слева
9.	слева обороты снизить	21.	слева мощность снизить
10.	слева снизить обороты	22.	слева снизить мощность
11.	снизить обороты слева	23.	снизить мощность слева
12.	снизить слева обороты	24.	снизить слева мощность

В результате сравнения, если длина распознаваемой команды попадает в интервал $\left[\frac{1}{a}L_i; aL_i\right], (i=1..8)$, где $L_i, (i=1..8)$ – длина i -й эталонной команды; a – коэффициент, подбираемый экспериментально, команда признаётся идентичной заданному эталону. Коэффициент a подбирался в процессе идентификации РК таким образом, чтобы при отсутствии ошибок первого рода количество ошибок второго рода было минимальным.

В результате эксперимента при идентификации восьми тестовых РК минимальное количество ошибок второго рода составило 46 при отсутствии ошибок первого рода, что говорит о хорошей эффективности разработанного алгоритма. Выявленная «оптимальная» библиотека РК приведена в таблице 3.

Заключение

Использование OWL-онтологии, позволяющей получать наборы синонимов для каждой речевой команды из библиотеки эталонов, позволило определить лучшую библиотеку эталонов, в которой команды в большей степени различимы по длительности произнесения, то

Таблица 3 – «Оптимальная» библиотека РК

№	Команда
1.	обороты слева меньше
2.	мощнее слева
3.	обороты справа меньше
4.	мощнее справа
5.	слева стартер
6.	слева турбостартер выключить
7.	стартер справа
8.	выключить турбостартер справа

есть в процессе идентификации РК с использованием этой библиотеки произошло минимальное количество ошибок второго рода при отсутствии ошибок первого рода. При этом не потерялся тематический и практический смысл библиотеки.

Благодарности

Работа выполнена при финансовой поддержке РФФИ (проект № 16-48-730305 «Разработка метода распознавания слитной речи по двумерным автокорреляционным изображениям с использованием онтологии, извлекаемой из текстов определённой предметной области»).

Список источников

- [1] **Крашенинников, В.Р.** Модели изменчивости речевых команд / В.Р. Крашенинников, А.И. Армер, И.Н. Синицин // Научные технологии. – М.: Радиотехника, 2007. – С 56-64.
- [2] **Крашенинников, В.Р.** Алгоритм выбора эталонов речевых команд при распознавании речи / В.Р. Крашенинников, Н.А. Крашенинникова, В.В. Кузнецов // Труды Российского научно-технического общества радиотехники, электроники и связи им. А.С. Попова. 2007. - № LXII. С. 158-159.
- [3] **Pienado A., Segura J.C.** Speech recognition over digital channels: robustness and standards / John Wiley & Sons Ltd., 2006. 257 P.
- [4] **Keshet J., Bengio S.** Automatic Speech and Speaker Recognition: Large Margin and Kernel Methods / John Wiley & Sons Ltd., 2009. 253 P.
- [5] **Gelbart D.** Ensemble Feature Selection for Multi-Stream Automatic Speech Recognition / Technical Report No. UCB/EECS-2008-160, University of California at Berkeley, 2008, 117 P.
- [6] **Галунов В.И.** Современные проблемы в области распознавания речи / В.И. Галунов, А.Н. Соловьев // Информационные технологии и вычислительные системы. Под. ред. С. В. Емельянова, Вып. 2. М.: «Рохос», 2004. - С.41-45.
- [7] **Rabiner L., Juang B.-H.** Fundamentals of speech recognition / Prentice-Hall International Inc., USA, 1993, 497 P.
- [8] **Гаврилова Т.А.** Базы знаний интеллектуальных систем: [учеб. пособие для студентов вузов]/ Т.А. Гаврилова, В.Ф. Хорошевский. – СПб: Изд-во «Питер», 2001. – 384 с.

APPROACH TO CREATING SETS OF REFERENCE SPEECH COMMANDS USING ONTOLOGY

A.I. Armer¹, V.S. Moshkin²

Ulyanovsk State Technical University, Ulyanovsk, Russia

¹a.armer@mail.ru, ²PostForVadim@ya.ru

Abstract

This article describes the method of formation of speech commands reference library using ontology to increase their further recognition. The article presents the formal model of a subject OWL-ontology, used as a base of knowledge in the process of formation and expansion of the library of reference teams in the process of recognition of voice commands. An algorithm for retrieval of the list of synonyms of input speech commands, based on the subject of OWL-ontology is also described. The article describes the research methodology, as well as the results of experiments on the formation of variants of the reference library and selection of a library with the most recognizable commands using

OWL-ontology of civil aircraft voice command control. In conclusion, the results of experiments demonstrating the effectiveness of this approach are presented.

Key words: *speech commands, ontology, model, speech recognition.*

Citation: *Armer AI, Moshkin VS. Approach to creating set of reference speech commands using ontology [In Russian]. Ontology of designing. 2016; v.6, 3(21): 270-277. - DOI: 10.18287/2223-9537-2016-6-3-270-277.*

References

- [1] *Krasheninnikov VR, Armer AI, Sinicin IN. Models variability voice commands [In Russian]. High Tech. – М.: Radiotekhnika, 2007. – pp. 56-64.*
- [2] *Krasheninnikov VR, Krasheninnikova NA., Kuznecov VV. The selection algorithm standards of voice commands at the speech recognition [In Russian]. Proceedings of the Russian Scientific and Technical Society of Radio Engineering, Electronics and Communication named after AS Popov. 2007. № LXII. pp. 158-159.*
- [3] *Pienado A, Segura JC. Speech recognition over digital channels: robustness and standards / John Wiley & Sons Ltd., 2006. 257 P.*
- [4] *Kesht J., Bengio S. Automatic Speech and Speaker Recognition: Large Margin and Kernel Methods / John Wiley & Sons Ltd., 2009. 253 P.*
- [5] *Gelbart D. Ensemble Feature Selection for Multi-Stream Automatic Speech Recognition / Technical Report No. UCB/EECS-2008-160, University of California at Berkeley, 2008, P. 117*
- [6] *Galunov VI., Solov'ev AN. Contemporary Issues in Speech Recognition [In Russian]. Information technology and computer systems. Ed. SV Emelyanov, 2004, Is.2. - Moscow: «Rojos», 2004. pp.41-45.*
- [7] *Rabiner L, Juang BH. Fundamentals of speech recognition / Prentice-Hall International Inc., USA, 1993, 497 P.*
- [8] *Gavrilova TA, Horoshevskiy VF. Knowledge Base of Intelligent Systems [In Russian]. - SPb: Pbl.«Piter», 2001. – 384 p.*

Сведения об авторах



Армер Андрей Игоревич, 1982 г. рождения, окончил Ульяновский государственный технический университет (2004) г., к.т.н. (2006 г.), доцент кафедры прикладной математики и информатики Ульяновского государственного технического университета (2011 г.), круг научных интересов: обработка и анализ речевых сигналов, распознавание образов. Имеет более 60 публикаций.

Armer Andrey Igorevich, (b.1982), education: Ulyanovsk State Technical University, 2004, PhD (modeling and recognition of speech signals against the background of high noise), 2006, main work: Ulyanovsk State Technical University, Associate professor (Dpt. of Applied Mathematics and Computer Science), 2011, Expert in modeling and recognition of speech signals against the background of high noise. He has more then 60 articles.



Мошкин Вадим Сергеевич, 1990 г. рождения. Окончил Ульяновский государственный технический университет (УлГТУ) в 2012 г., аспирант, ассистент кафедры информационных систем УлГТУ. Ведущий разработчик Центра разработки электронных мультимедиа технологий УлГТУ. Член Российской ассоциации искусственного интеллекта. В списке научных трудов более 30 статей в области интеллектуальной обработки знаний, автоматизации проектирования, а также построения прикладных интеллектуальных систем.

Moshkin Vadim Sergeevich (b. 1990) graduated from the Ulyanovsk State Technical University (UISTU) in 2012, post-graduate student at the Chair 'Information Systems' of Ulyanovsk State Technical University. A lead developer of Center of Multimedia Technologies of UISTU. He is Russian Association of Artificial Intelligence member. He is co-author of more than 30 publications in the field of intellectual knowledge processing, automation of design and construction of

intelligent systems applied.