

ISSN 2078-9181

РОССИЙСКАЯ АКАДЕМИЯ НАУК  
Отделение нанотехнологий и информационных технологий

САНКТ-ПЕТЕРБУРГСКИЙ  
ИНСТИТУТ ИНФОРМАТИКИ И АВТОМАТИЗАЦИИ РАН

# ТРУДЫ СПИИРАН

[proceedings.spiiras.nw.ru](http://proceedings.spiiras.nw.ru)



ВЫПУСК 5(36)



Санкт Петербург  
2014

18+

# Труды СПИИРАН

Выпуск № 5(36), 2014

Научный, научно-образовательный, междисциплинарный журнал с базовой специализацией в области информатики, автоматизации и прикладной математики

Журнал основан в 2002 году

## Учредитель и издатель

Федеральное государственное бюджетное учреждение науки  
Санкт-Петербургский институт информатики и автоматизации Российской академии наук  
(СПИИРАН)

## Главный редактор

Р.М. Юсупов, чл.-корр. РАН, д-р техн. наук, проф., С.-Петербург, РФ

## Редакционная коллегия

- |                                                                                                                        |                                                                                              |
|------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------|
| <b>А.А. Ашимов</b> , академик национальной академии наук Республики Казахстан д-р техн. наук, проф., Алматы, Казахстан | <b>А.Л. Ронжин</b> (зам. главного редактора), д-р техн. наук, проф., С.-Петербург, РФ        |
| <b>С.Н. Баранов</b> , д-р физ.-мат. наук, проф., С.-Петербург, РФ                                                      | <b>А.И. Рудской</b> , член-корр. РАН, д-р техн. наук, проф., С.-Петербург, РФ                |
| <b>Н.П. Веселкин</b> , академик РАН, д-р мед. наук, проф., С.-Петербург, РФ                                            | <b>В.А. Сарычев</b> , д-р техн. наук, проф., С.-Петербург, РФ                                |
| <b>В.И. Городецкий</b> , д-р техн. наук, проф., С.-Петербург, РФ                                                       | <b>В. Стурев</b> , академик Болгарской академии наук, д-р техн. наук, проф., София, Болгария |
| <b>О.Ю. Гусихин</b> , Ph.D., Диаборн, США                                                                              | <b>В.А. Скорин</b> , Ph.D., проф., Бингемптон, США                                           |
| <b>В. Делич</b> , д-р техн. наук, проф., Нови-Сад, Сербия                                                              | <b>А.В. Смирнов</b> , д-р техн. наук, проф., С.-Петербург, РФ                                |
| <b>А.Б. Долгий</b> , Dr. Habil., проф., Сент-Этьен, Франция                                                            | <b>Б.Я. Советов</b> , академик РАО, д-р техн. наук, проф., С.-Петербург, РФ                  |
| <b>М. Железны</b> , Ph.D., доцент, Пльзень, Чешская республика                                                         | <b>В.А. Соيفер</b> , член-корр. РАН, д-р техн. наук, проф., Самара, РФ                       |
| <b>Д.А. Иванов</b> , д-р экон. наук, проф., Берлин, Германия                                                           | <b>Б.В. Соколов</b> , д-р техн. наук, проф., С.-Петербург, РФ                                |
| <b>О.С. Ипатов</b> , д-р техн. наук, проф., С.-Петербург, РФ                                                           | <b>Л.В. Уткин</b> , д-р техн. наук, проф., С.-Петербург, РФ                                  |
| <b>В.П. Леонов</b> , д-р пед. наук, проф., С.-Петербург, РФ                                                            | <b>А.Л. Фрадков</b> , д-р техн. наук, проф., С.-Петербург, РФ                                |
| <b>Г.А. Леонов</b> , член-корр. РАН, д-р физ.-мат. наук, проф., С.-Петербург, РФ                                       | <b>Н.В. Хованов</b> , д-р физ.-мат. наук, проф., С.-Петербург, РФ                            |
| <b>К.П. Марков</b> , Ph.D., доцент, Аизу, Япония                                                                       | <b>Д.С. Черешкин</b> , д-р техн. наук, проф., Москва, РФ                                     |
| <b>Ю.А. Меркурьев</b> , член-корр. Латвийской академии наук, Dr. Habil., проф., Рига, Латвия                           | <b>Л.Б. Шереметов</b> , д-р техн. наук, Мехико, Мексика                                      |
| <b>Н.А. Молдовян</b> , д-р техн. наук, проф., С.-Петербург, РФ                                                         | <b>А.В. Язенин</b> , д-р техн. наук, профессор, Тверь, РФ                                    |
| <b>А.А. Петровский</b> , д-р техн. наук, проф., Минск, Беларусь                                                        |                                                                                              |
| <b>В.В. Попович</b> , д-р техн. наук, проф., С.-Петербург, РФ                                                          |                                                                                              |
| <b>В.А. Путилов</b> , д-р техн. наук, проф., Апатиты, Россия                                                           |                                                                                              |

## Адрес редакции

199178, Санкт-Петербург, 14-я линия, д. 39,

e-mail: [publ@iiias.spb.su](mailto:publ@iiias.spb.su), сайт: <http://www.proceedings.spiiras.nw.ru/>

Подписано к печати 14.11.2014. Формат 60×90 1/16. Усл. печ. л. 14,0. Заказ № 657. Тираж 200 экз., цена свободная  
Отпечатано в Редакционно-издательском центре ГУАП, 190000, Санкт-Петербург, Б. Морская, д. 67

Журнал зарегистрирован Федеральной службой по надзору в сфере связи и массовых коммуникаций,  
свидетельство ПИ № ФС77-41695 от 19 августа 2010 г.  
Подписной индекс 29393 по каталогу «Почта России»

Журнал входит в «Перечень ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные научные результаты диссертации на соискание ученой степени доктора и кандидата наук»

© Федеральное государственное бюджетное учреждение науки

Санкт-Петербургский институт информатики и автоматизации Российской академии наук, 2014

Разрешается воспроизведение в прессе, а также сообщение в эфир или по кабелю опубликованных в составе печатного периодического издания—журнала «Труды СПИИРАН» статей по текущим экономическим, политическим, социальным и религиозным вопросам с обязательным указанием имени автора статьи и печатного периодического издания—журнала «Труды СПИИРАН»

# SPIIRAS Proceedings

Issue № 5(36), 2014

Scientific, educational, and interdisciplinary journal primarily specialized  
in computer science, automation, and applied mathematics

Trudy SPIIRAN ♦ Founded in 2002 ♦ Труды СПИИРАН

---

## Founder and Publisher

Federal State Budget Institution of Science

St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences  
(SPIIRAS)

---

## Editor-in-Chief

R.M. Yusupov, Prof., Dr. Sci., Corr. Member of RAS, St. Petersburg, Russia

---

## Editorial Board Members

**A.A. Ashimov**, Prof., Dr. Sci., Academician  
of the National Academy of Sciences of the  
Republic of Kazakhstan, Almaty, Kazakhstan  
**S.N. Baranov**, Prof., Dr. Sci., St. Petersburg, Russia  
**N.P. Veselkin**, Prof., Dr. Sci., Academician of RAS,  
St. Petersburg, Russia  
**V.I. Gorodetski**, Prof., Dr. Sci., St. Petersburg, Russia  
**O.Yu. Gusikhin**, Ph. D., Dearborn, USA  
**V. Delic**, Prof., Dr. Sci., Novi Sad, Serbia  
**A. Dolgui**, Prof., Dr. Habil., St. Etienne, France  
**M. Zelezny**, Assoc. Prof., Ph.D., Plzen, Czech  
Republic  
**D.A. Ivanov**, Prof., Dr. Habil., Berlin, Germany  
**O.S. Ipatov**, Prof., Dr. Sci., St. Petersburg, Russia  
**V.P. Leonov**, Prof., Dr. Sci., St. Petersburg, Russia  
**G.A. Leonov**, Prof., Dr. Sci., Corr. Member of RAS,  
St. Petersburg, Russia  
**K.P. Markov**, Assoc. Prof., Ph.D., Aizu, Japan  
**Yu.A. Merkurjev**, Prof., Dr. Habil., Corr. Member  
of the Latvian Academy of Sciences, Riga, Latvia  
**N.A. Moldovian**, Prof., Dr. Sci., St. Petersburg, Russia  
**A.A. Petrovsky**, Prof., Dr. Sci., Minsk, Belarus  
**V.V. Popovich**, Prof., Dr. Sci., St. Petersburg, Russia  
**V.A. Putilov**, Prof., Dr. Sci., Apatity, Russia

**A.L. Ronzhin** (Deputy Editor-in-Chief),  
Prof., Dr. Sci., St. Petersburg, Russia  
**A.I. Rudskoi**, Prof., Dr. Sci., Corr. Member of RAS,  
St. Petersburg, Russia  
**V.A. Saruchev**, Prof., Dr. Sci., St. Petersburg,  
Russia  
**V. Sgurev**, Prof., Dr. Sci., Academician  
of the Bulgarian academy of sciences, Sofia,  
Bulgaria  
**V. Skormin**, Prof., Ph.D., Binghamton, USA  
**A.V. Smirnov**, Prof., Dr. Sci., St. Petersburg, Russia  
**B.Ya. Sovetov**, Prof., Dr. Sci., Academician of RAE,  
St. Petersburg, Russia  
**V.A. Soyfer**, Prof., Dr. Sci., Corr. Member of RAS,  
Samara, Russia  
**B.V. Sokolov**, Prof., Dr. Sci., St. Petersburg, Russia  
**L.V. Utkin**, Prof., Dr. Sci., St. Petersburg, Russia  
**A.L. Fradkov**, Prof., Dr. Sci., St. Petersburg, Russia  
**N.V. Hovanov**, Prof., Dr. Sci., St. Petersburg,  
Russia  
**D.S. Chereshekin**, Prof., Dr. Sci., Moscow, Russia  
**L.B. Sheremetov**, Assoc. Prof., Dr. Sci., Mexico,  
Mexico  
**A.V. Yazenin**, Prof., Dr. Sci. Tver, Russia

---

## Editorial Board's address

14-th line VO, 39, SPIIRAS, St. Petersburg, 199178, Russia,

e-mail: [publ@iias.spb.su](mailto:publ@iias.spb.su), web: <http://www.proceedings.spiiras.nw.ru/>

Signed to print 14.11.2014

Printed in Publishing center GUAP, 67, B. Morskaya, St. Petersburg, 190000, Russia

---

The journal is registered in Russian Federal Agency for Communications and Mass-Media Supervision,  
certificate ПИ № ФС77-41695 dated August 19, 2010 r.

Subscription Index 29393, Russian Post Catalog

© Federal State Budget Institution of Science

St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, 2014

## СОДЕРЖАНИЕ

Баранов С.Н., Тележкин А.М. МЕТРИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ПРОГРАММНЫХ РАЗРАБОТОК	5
Цветков М.В., Смирнова О.В., Гальяно Ф.Р. СИСТЕМА МОНИТОРИНГА ЛЕДОВОЙ ОБСТАНОВКИ И ОБЕСПЕЧЕНИЯ БЕЗОПАСНОГО СУДОХОДСТВА ПО СЕВЕРНОМУ МОРСКОМУ ПУТИ НА БАЗЕ ИНТЕЛЛЕКТУАЛЬНОЙ ГИС	28
Девятков В.В. РАЗВИТИЕ МЕТОДОЛОГИИ И ТЕХНОЛОГИИ ИМИТАЦИОННЫХ ИССЛЕДОВАНИЙ СЛОЖНЫХ СИСТЕМ	44
Рубцова Ю.В. РАЗРАБОТКА И ИССЛЕДОВАНИЕ ПРЕДМЕТНО НЕЗАВИСИМОГО КЛАССИФИКАТОРА ТЕКСТОВ ПО ТОНАЛЬНОСТИ	59
Созинова И.С., Романов А.С., Мещеряков Р.В. ОПРЕДЕЛЕНИЕ ПОИСКОВОГО СПАМА С ИСПОЛЬЗОВАНИЕМ МЕТОДА ОПОРНЫХ ВЕКТОРОВ	78
Важенина Д.А., Кипяткова И.С., Марков К.П., Карпов А.А. МЕТОДИКА ВЫБОРА ФОНЕМНОГО НАБОРА ДЛЯ АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ РУССКОЙ РЕЧИ	92
Евграфова К.В., Качковская ТВ ДИФФЕРЕНЦИАЛЬНЫЕ ПОРОГИ ВОСПРИЯТИЯ ДЛИТЕЛЬНОСТИ ЯЗЫКОВЫХ ЕДИНИЦ В ЗАВИСИМОСТИ ОТ ФОНЕТИЧЕСКОГО КОНТЕКСТА	114
Азаров И.С., Петровский А.А. ФОРМИРОВАНИЕ ПЕРСОНАЛЬНОЙ МОДЕЛИ ГОЛОСА ДИКТОРА С УНИВЕРСАЛЬНЫМ ФОНЕТИЧЕСКИМ ПРОСТРАНСТВОМ ПРИЗНАКОВ НА ОСНОВЕ ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ	128
Гениатулина Е.В., Гриф М.Г. МЕТОДЫ ПРОЕКТИРОВАНИЯ И МОДЕЛИРОВАНИЯ В ЗАДАЧАХ ОПТИМИЗАЦИИ ПРОЦЕССОВ ФУНКЦИОНИРОВАНИЯ ЧЕЛОВЕКО-МАШИННЫХ СИСТЕМ	151
Кулик Б. А., Курбанов В.Г., Фридман А.Я. ПАРАЛЛЕЛЬНАЯ ОБРАБОТКА ДАННЫХ И ЗНАНИЙ МЕТОДАМИ АЛГЕБРЫ КОРТЕЖЕЙ	168
Мусаев А.А. ЧИСЛЕННОЕ ИССЛЕДОВАНИЕ УПРАВЛЯЮЩИХ СТРАТЕГИЙ, ОСНОВАННЫХ НА АНАЛИЗЕ ТРЕНДОВ	180
Вихман В.В., Панков М.А. ПОВЫШЕНИЕ СТОЙКОСТИ ХЕШ-ФУНКЦИЙ В ИНФОРМАЦИОННЫХ СИСТЕМАХ НА ОСНОВЕ АЛГОРИТМА МНОГОИТЕРАЦИОННОГО ХЕШИРОВАНИЯ С НЕСКОЛЬКИМИ МОДИФИКАТОРАМИ	194
Щекотов М.С. АНАЛИЗ ПОДХОДОВ К ПОЗИЦИОНИРОВАНИЮ ВНУТРИ ПОМЕЩЕНИЙ С ИСПОЛЬЗОВАНИЕМ ТРИЛАТЕРАЦИИ СИГНАЛОВ WI-FI	206
Смокий О.И. ВЛИЯНИЕ АТМОСФЕРНОЙ ТУРБУЛЕНТНОСТИ, МОЛЕКУЛЯРНО-АЭРОЗОЛЬНОГО РАССЕЯНИЯ И ПАРАМЕТРОВ БОРТОВОЙ ОПТИЧЕСКОЙ АППАРАТУРЫ НА КАЧЕСТВО СПУТНИКОВОЙ ВИДЕОИНФОРМАЦИИ ПРИ СЪЕМКЕ ЗЕМЛИ ИЗ КОСМОСА	215

## CONTENTS

Baranov S.N., Telezhkin A.M. METRICS FOR SOFTWARE DEVELOPMENT	5
Tsvetkov M.V., Smirnova O.V., Galiano P.R. ICE SITUATION AND SAFE NAVIGATION MONITORING SYSTEM ON THE NORTHERN SEA ROUTE BASED ON INTELLIGENT GIS	28
Devyatkov V.V. THE DEVELOPMENT OF METHODOLOGY AND TECHNOLOGY OF SIMULATION STUDIES OF COMPLEX SYSTEMS	44
Rubtsova Y.V. RESEARCH AND DEVELOPMENT OF DOMAIN INDEPENDENT SENTIMENT CLASSIFIER	59
Sozina I.S., Romanov A.S., Meshcheryakov R.V. SEARCH SPAM IDENTIFICATION USING SUPPORT VECTOR MACHINE	78
Vazhenina D.A., Kipyatkova I.S., Markov K.P., Karpov A.A. TECHNIQUE FOR PHONEME SET SELECTION FOR AUTOMATIC RUSSIAN SPEECH RECOGNITION	92
Evgrafova K.V., Kachkovskaya T.V. DIFFERENTIAL THRESHOLDS OF PERCEPTION OF SOUND DURATION IN DIFFERENT PHONETIC CONTEXTS	114
Azarov E.S., Petrovsky A.A. TRAINING PERSONAL VOICE MODEL OF A SPEAKER WITH UNIFIED PHONETIC SPACE OF FEATURES USING ARTIFICIAL NEURAL NETWORK	128
Geniatulina E.V., Grif M.G. MAN-MACHINE SYSTEMS FUNCTIONING DESIGN AND MODELING METHODS IN OPTIMIZATION PROCESSES PROBLEMS	151
Kulik B.A., Kurbanov V.G., Fridman A.A. PARALLEL PROCESSING OF DATA AND KNOWLEDGE BY MEANS OF N-TUPLE ALGEBRA	168
Musaev A.A. NUMERICAL RESEARCH OF THE OPERATING STRATEGY BASED ON THE TRENDS ANALYSIS	180
Vikhman V.V., Pankov M.A. SECURITY INCREASING OF HASH FUNCTIONS IN INFORMATION SYSTEMS ON THE BASIS OF MULTI-ITERATIVE HASHING ALGORITHM WITH SEVERAL MODIFIERS	194
Shchekotov M.S. ANALYSIS OF INDOOR POSITIONING APPROACHES BASED ON WI-FI TRILATERATION	206
Smokty O.I. THE INFLUENCE OF TURBULENT ATMOSPHERE, MOLECULAR-AEROSOL SCATTERING AND BOARD OPTICAL DEVICE ON THE SATELLITE VIDEO-INFORMATION QUALITY AT THE EARTH'S REMOTE SENSING FROM SPACE	215

С.Н.БАРАНОВ, А.М.ТЕЛЕЖКИН  
**МЕТРИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ПРОГРАММНЫХ  
РАЗРАБОТОК**

---

*Баранов С.Н., Тележкин А.М. Метрическое обеспечение программных разработок.*

**Аннотация.** В статье систематизируются и сводятся воедино свыше 110 различных процессных, проектных и продуктовых метрик, установившихся в практике промышленной разработки программных продуктов, из которых в каждом отдельном программном проекте используется только их небольшая часть. Выбор, сбор, накопление и регулярный анализ этих метрических данных позволяет создавать качественное программное обеспечение в заданных ограничениях и определять направления для дальнейшего совершенствования процесса разработки в организации и проектной группе, принимая решения на основе фактов, а не предположений.

**Ключевые слова:** Технология программирования, проектные метрики, процесс разработки ПО, модель СММІ.

*Baranov S.N., Telezhkin A.M. Metrics for Software Development.*

**Abstract.** Over 110 different process, project, and product metrics used in the current practice of software development are assembled and classified in this paper, of which only a small number is used in each particular project. Selecting, collecting, accumulating and regular analysis of these metric data allows for creation of high-quality software under the given constraints and for identifying ways to further improve the software development process in the given organization or project team through decision making based on facts rather than assumptions or opinions.

**Keywords:** software engineering, project metrics, software development process, СММІ.

---

**1. Введение.** Создание высококачественного программного обеспечения (ПО) в современную эпоху информационных технологий давно стало одним из важных и существенных высокотехнологичных производств в развитых и развивающихся странах. Огромная сложность ПО как технического объекта, по сравнению с механизмами и техническими средствами предшествующих технологических эпох, является следствием огромного количества состояний, в которых может находиться программная система (под состоянием такой системы обычно понимается совокупность одновременных значений всех ее переменных, изменяющихся с течением времени).

Любое конкретное поведение программной системы рассматривается как некоторый путь в дискретном пространстве состояний; перебрать все такие пути при тестировании невозможно практически. Например, если в программе имеется всего  $N$  независимых 16-разрядных переменных, то нижняя оценка для общего числа  $S$  ее состояний выражается равенством  $S=2^{16 \times N}$ , а число путей длины  $L$  в этом пространстве из какого-либо начального состояния задается «экспоненциальной башней» высоты  $L$  из значений  $S$ , что обозначается ме-

тафорой «экспоненциальный взрыв» (exponential explosion).

Вместе с тем, неправильное поведение программы на каком-либо конкретном пути из множества всех допустимых путей (которое является подмножеством всех возможных путей в данном пространстве состояний), как правило, приводит к неправильному срабатыванию того устройства, в котором она установлена и может привести даже к катастрофическим последствиям.

Дисциплина программирования, сложившаяся в индустрии программного обеспечения за последние 40 лет, создала ряд формализованных моделей для процесса разработки высококачественных программных продуктов (см. рисунок 1), из которых в настоящее время наиболее полной является модель [1], разработанная Институтом технологии программирования при университете Карнеги-Меллон.

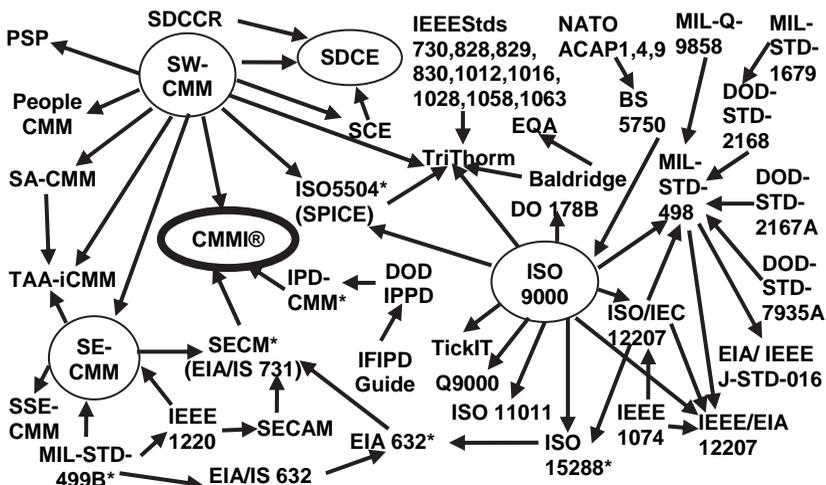


Рис. 1. Взаимосвязанные модели процесса разработки

Метрическая программа этой модели предполагает постоянный сбор и анализ различных метрик, характеризующих текущее состояние разрабатываемого продукта, ход процесса разработки и достигнутый уровень зрелости организации-разработчика. Выбор, сбор, накопление и регулярный анализ этих метрических данных позволяет создавать качественное программное обеспечение в заданных ограничениях и определять направления для дальнейшего совершенствования процесса разработки в организации и проектной группе, принимая решения на основе фактов, а не предположений. Вместе с тем, модель CMMI,

как и многие другие, не определяет ни конкретный состав этих метрик, ни способы их сбора, хранения и анализа. В данной работе систематизируется накопленный опыт по метрическому обеспечению программных проектов для его дальнейшего применения в практических разработках. Наряду с общемировым опытом, использован опыт авторов в научных исследованиях и в промышленных разработках в компаниях Моторола [2-3], Эксиджен [4] и других.

**2. Классификация метрик.** На сегодняшний день известно свыше 500 различных измеряемых показателей, так или иначе относящихся к разработке программных продуктов. Разумеется, в каждом отдельном проекте используется только небольшая часть этого множества. В данной работе собраны и приведены в систему 112 различных метрик, некоторые из которых, в свою очередь, обеспечивают целых куст отдельных метрических показателей (всего свыше 230). Выбор состава метрик, способов их сбора и анализа определяется целями и задачами проекта и рядом других сопутствующих условий, которые к тому же могут меняться в процессе работы. Огромную роль в метрической программе играет наличие достоверной исторической базы данных (ИБД) проектов, как уже выполненных, так и находящихся в процессе разработки в данной организации, что позволяет проводить объективный анализ этих данных и делать оценки для текущих и будущих разработок применительно к данным условиям. Создание и поддержание такой ИБД является важной задачей, а сама база – существенным информационным активом организации-разработчика.

По своему назначению метрики можно разделить на три основных класса: продуктовые, проектные и процессные; причем, различаются метрики базовые – основные, ни от чего не зависящие, общие для многих разработок, производные – выводимые из базовых и представляющие дополнительную информацию, и специфические – индивидуальные для каждого отдельного продукта, проекта или процесса. Общими для всех метрик характеристиками являются единица измерения, способ измерения и фаза (фазы) жизненного цикла, на которой эта метрика может быть измерена. Значения многих метрик изменяются в процессе разработки, так что отслеживание их динамики является важной задачей для оперативного управления проектом.

По ИБД для некоторых часто используемых метрик определяются их референсные значения: среднее, лучшее и, возможно, наихудшее по организации-разработчику, а также аналогичные данные по промышленности на определенный срок. Далее эти три типа метрик рассматриваются более подробно. Для каждой метрики указывается ее название с англоязычным эквивалентом, определение ее сущности,

если она не следует из ее названия, и единица измерения, если это не просто число или процент. Наряду с метриками, продукты, проекты и процессы характеризуются еще и атрибутами (свойствами), как правило, в описательном виде.

**2.1. Продуктовые метрики** характеризуют непосредственно измеряемые свойства самого программного продукта (изделия).

1. Исходные требования – *Source requirements*. Число исходных требований, общее и по типам (функциональные, интерфейсные, по производительности, системные и т.д.).

2. Размер онтологии – *Ontology size*. Число разных онтологических сущностей, задействованных в описании исходных требований.

3. Изменчивость требований – *Volatility of requirements*. Число изменений, добавлений и удалений в формулировках исходных требований, совершенных по инициативе заказчика, абсолютное и в отношении к общему числу требований.

4. Полнота требований – *Completeness of requirements*. Число требований, добавленных по инициативе разработчиков, абсолютное и в отношении к общему числу требований.

5. Противоречивость требований – *Inconsistency of requirements*. Число изменений и удалений требований, совершенных по инициативе разработчиков, абсолютное и в отношении к общему числу.

6. Завершенность требований – *Finalization of requirements*. Число незакрытых вопросов по уточнению исходных требований, общее и по типам требований.

7. Системные компоненты – *System components*. Число и наименования различных системных компонентов в продукте. Определяется по высокоуровневому проекту и структуре разбиения работ.

8. Программные модули – *Software modules*. Число и наименования различных программных модулей в продукте. Определяется по компонентам высокоуровневого проекта и структуре разбиения работ.

9. Используемые интерфейсы – *Interfaces used*. Число и наименования различных интерфейсов, используемых в продукте. Определяется по высокоуровневому проекту и структуре разбиения работ.

10. Входные и выходные каналы – *I/O channels*. Число и наименования различных входных и выходных каналов в продукте. Соотносится со сложностью по Холстеду [5] высокоуровневого проекта.

11. Используемые языки программирования – *Programming languages used*. Число и названия языков программирования.

12. Используемые базы данных – *Databases used*.

13. Используемые серверы приложений – *Application servers used*. Число и названия таких серверов.

14. Используемые операционные системы – Operating systems used. Число и названия таких операционных систем.

15. Используемые технологии – Technologies used.

16. Размер кода – Code size. Число строк кода на входном языке программирования, не считая пустые строки и строки, состоящие только из комментариев; единица измерения – KLOC или KAELOC.

17. Размер кода для модульного тестирования – Code size for unit testing. Число строк кода на входном языке программирования для модульного тестирования; единица измерения – KLOC или KAELOC.

18. Ветвления в коде – Code branches. Число всех операторов ветвления и цикла в исходном коде.

19. Параллелизм кода – Code parallelism. Число параллельно исполняемых ветвей в исходном коде.

20. Сложность кода по Холстеду – Halstead complexity. Четыре основных показателя: число различных/всех ( $\eta_1/N_1$ ) операторов и число различных/всех ( $\eta_2/N_2$ ) операндов в программе и пять производных, определяемых через них:  $N$  – длина программы,  $V$  – объем программы,  $L^*$  – оценка длины реализации программы,  $\lambda$  – уровень языка программирования,  $I$  – интеллектуальное содержание программы [5].

21. Цикломатическая сложность кода (сложность по МакКейбу) – Cyclomatic complexity (McCabe complexity). Цикломатическое число  $\lambda(G) = m - n + 2 \times p$ , где  $m$  – число вершин,  $n$  – число дуг и  $p$  – число компонентов связности в графе  $G$  передач управления в программе [6].

22. Качество кода – Code quality. Оценка числа остаточных дефектов в коде (как известных, так и еще не выявленных), абсолютное и в отношении к общему размеру кода; единица измерения – число/KLOC или число/KAELOC, либо в виде  $N_{\sigma}$ .

23. Повторное использование кода – Code reuse. Процент повторно использованного кода в общем коде поставки.

24. Размер кода для повторного использования – Code for reuse. Число строк вновь разрабатываемого кода с дополнительной целью его дальнейшего повторного использования; единица измерения – KLOC (KAELOC).

25. Покрытие тестами требований – Requirement coverage with tests. Процент числа требований, покрытых тестами, от общего числа требований, суммарный и по типам требований.

26. Пост-релизные дефекты – Post-release faults. Число дефектов, обнаруженных в продукте после поставки заказчику в течение определенного срока (обычно 6 или 12 месяцев), абсолютное и в отношении к размеру всего поставляемого кода; единица измерения – число/KLOC

или число/КАЕЛОС, либо в виде №.

27. Покрытие тестами кода – Code coverage with tests. Процент кода, покрытого тестами, от всего кода.

28. Покрытие тестами ветвлений в коде – Branch coverage with tests. Процент числа ветвлений, покрытых тестами, от числа ветвлений во всем поставляемом коде.

29. Размер документации – Documentation size. Число страниц в разработанной документации, поставляемой заказчику.

30. Повторное использование документации – Documentation reuse. Процент повторно использованной документации в общем объеме документации, поставляемой заказчику (по числу страниц).

31. Использованные патенты – Patents used. Число и наименования запатентованных решений, использованных в данном продукте.

32. Инновационность продукта – Product innovativeness. Число признанных инновационных решений, использованных в данном продукте, сделанных участниками проекта и реализованных в нем.

33. Пакет поставки – Delivery package. Метрические данные по видам составляющих компонентов пакета поставки, включая исходных код, выполняемый код, документацию, отчеты о дефектах, материалы тестирования, спецификации, планы, процедуры, технологии оценивания, предложения по инновациям и т.д.

34. Обратная связь по продукту – Product feedback. Число предложений, отзывов, рекламаций и т.д., полученных в течение определенного периода (обычно 6 или 12 месяцев) после поставки продукта.

**2.2. Проектные метрики** измеряемым образом характеризуют сам проект по созданию и сопровождению программного продукта.

1. Трудоемкость – Effort. Суммарные трудозатраты, общие и по фазам проекта, по видам работ и категориям исполнителей, по плану и по факту на данный момент; единица измерения – число человеко-месяцев, человеко-дней или человеко-часов.

2. Производительность труда – Performance. Отношение размера разработанного в данном проекте кода в законченном продукте к общей трудоемкости на его создание, плановое и по факту; единица измерения – KLOC (КАЕЛОС) на человеко-день (человеко-месяц).

3. Длительность проекта – Project duration. Число календарных дней или месяцев, отведенных на разработку, плановое и по факту.

4. Экономия от автоматизации разработки кода – Cost saving due to code development automation. Экономия затрат на разработку за счет автоматизации разработки кода, плановая и по факту; единица измерения – руб.

5. Автоматизация разработки кода – Code development automa-

tion. Процент автоматически созданного кода в общем размере кода, плановый и по факту.

6. Автоматизация разработки тестового кода – Test development automation. Процент автоматически созданного тестового кода в общем размере тестового кода (код тестов плюс код тестового окружения), плановый и по факту.

7. Автоматизация прогона тестов – Test run automation. Процент тестов, исполняемых без участия тестировщика, в общем числе тестов в эталонном тестовом наборе, плановый и по факту.

8. Экономия от автоматизации тестирования – Cost saving due to test automation. Экономия затрат на тестирование за счет автоматизации разработки и прогона тестов, плановая и по факту; единица измерения – руб.

9. Методы и инструменты – Methods and tools. Число и наименования различных методов и инструментов, используемых в разработке, включая лицензии на ПО.

10. Стоимость проекта – Project cost. Суммарная текущая стоимость проекта в денежном выражении, плановая и по факту, общая и по видам затрат, которые включают затраты на заработную плату, оборудование, ПО, специальные компоненты, обучение и переподготовку, книги и журналы, командировки, аренду, услуги сторонних организаций, расходы на юридическое сопровождение, накладные и представительские расходы, расходы на обеспечение секретности, стоимость исправления дефектов и т.д.; единица измерения – руб.

11. Стоимость разработчика – Cost of software developer. Стоимость программной разработки в суммарных затратах на одного участника в месяц, плановая и по факту, включая расходы на оборудование и накладные расходы; единица измерения – руб./месяц.

12. Предел стоимости – NTE (Not to Exceed). Суммарная величина всех затрат на проект, которую нельзя превысить ни при каких обстоятельствах; превышение означает немедленное прекращение проекта; единица измерения – руб.

13. Стоимость строки кода – Cost of a line of code. Отношение суммарных затрат на проект к размеру кода в LOC или AELOC в законченном программном продукте; единица измерения – руб.

14. Частота совершения ошибок – Error rate. Число совершаемых ошибок на единицу размера кода и документации; единица измерения – число/KLOC (KAELOC) для кода и число/стр. для документации.

15. Стоимость обеспечения качества – Cost of quality. Процент суммарных трудозатрат на деятельности по обеспечению качества

продукта, включая обзоры, тестирование, повышение квалификации и т.п., в суммарных трудозатратах на проект, плановый и по факту.

16. Стоимость переделок – Cost of poor quality. Процент суммарных трудозатрат на нахождение, исправление и повторную проверку исправления дефектов с учетом трудозатрат на регрессионное тестирование, вызванное исправлениями, в суммарных трудозатратах на проект, плановый и по факту.

17. Разработчики – Developers. Число лиц, создающих код продукта и сопроводительную документацию, плановое и по факту.

18. Распределенность разработки – Distributed (multi-site) development. Число отдельных групп (часто разнесенных по разным площадкам), работающих над проектом, если оно больше единицы.

19. Тестировщики – Testers. Число инженеров, тестирующих продукт, плановое и по факту.

20. Штат проекта – Staffing. Число ставок, включая руководство, разработчиков, тестировщиков и поддержку, плановое и по факту.

21. Женщины – Women. Число женщин, участвующих в проекте, абсолютное и как % от общего числа участников проекта.

22. Запланированные поставки – Planned deliverables. Число и наименования поставок заказчику, включенных в план разработки.

23. Точность поставок – On-time delivery. Процент поставок заказчику, совершенных вовремя, в общем числе всех совершенных поставок, плановый и по факту.

24. Точность планирования – Schedule accuracy. Процент отклонений фактических показателей хода проекта от запланированных в графике проекта в общем числе запланированных показателей.

25. Точность оценивания – Estimation accuracy. Процент отклонений фактических значений параметров проекта от их первоначальных оценок, по каждому параметру в отдельности и усредненный по всем учтенным параметрам.

26. Завершенность разработки – Development completion. Процент объема выполненных работ (по числу завершенных работ в структуре разбиения работ, по трудоемкости, по затратам и т.д.) от запланированного на данный момент, плановый и по факту.

27. Завершенность тестирования – Testing completion. Процент выполненных работ по тестированию (по числу созданных тестов, совершенных тестовых прогонов, успешно прошедших тестов и т.д.) от запланированного на данный момент, плановый и по факту.

28. Перепланирования проекта – Project replannings. Число утвержденных изменений плана в процессе работы, плановое и по факту.

29. Отчеты об ошибках – Defect reports. Число полученных отчете-

тов об ошибках за данный период времени.

30. Плотность выявленных дефектов – Defect density. Отношение числа выявленных дефектов в коде и документе к его размеру; единица измерения – число/KLOC или число/KAELLOC для кода и число/стр. для документов, либо в виде №.

31. Эффективность сдерживания дефектов – Defect containment effectiveness. Процент числа дефектов, выявленных и устраненных до сдачи продукта, от сделанной оценки общего числа дефектов.

32. Эффективность сдерживания серьезных (уровень 3 и выше по 5-бальной шкале) дефектов – MFC (Major Fault Containment). Процент серьезных ошибок, не перешедших в дефекты, от общего числа всех выявленных серьезных дефектов, плановое и по факту.

33. Эффективность сдерживания дефектов по фазам проекта – Defect containment effectiveness on project phases. Процент дефектов, выявленных и устраненных на данной фазе, и относящихся к ней, от оценки общего числа дефектов, относящихся к данной фазе.

34. Удовлетворенность заказчика – Customer satisfaction. Число баллов (обычно по 10-бальной шкале, где 10 – высшая оценка), поставленных заказчиком при анкетировании, плановое и по факту.

35. Размер тестового набора – Test suite size. Число тестов в эталонном тестовом наборе для проверки готовности продукта к выпуску.

36. Процент прохождения тестов – Tests passed. Процент прошедших тестов в общем числе тестов в эталонном тестовом наборе, плановый и по факту.

37. Циклы тестирования – Test cycles. Число полных циклов системного тестирования, совершенных до выпуска продукта, плановое и по факту.

38. Опыт совместной работы данного коллектива – Project team continuity. Среднее время работы одного человека в данном коллективе; единица измерения – число месяцев или лет.

39. Результативность (эффективность) взаимодействия внутри коллектива – Effectiveness of team interaction. Процент числа проблем, разрешенных внутри коллектива, в общем числе проблем, поднятых в ходе исполнения проекта.

40. Рациональность (экономичность) взаимодействия внутри коллектива – Efficiency of team interaction. Средние затраты на разрешение одной проблемы внутри коллектива разработчиков (число взаимодействий, длительность процесса разрешения, понесенные трудозатраты и т.д.); единица измерения – в зависимости от способа.

41. Опыт работы с данной платформой – Platform experience. Срок работы участников проекта с данной платформой, суммарный по

всем разработчикам и тестировщикам, и в пересчете на 1 человека; единица измерения – число месяцев или лет.

42. Опыт работы в данной предметной области – Subject domain experience. Срок работы участников проекта в данной предметной области, суммарный по всем разработчикам и тестировщикам, и в пересчете на 1 человека; единица измерения – число месяцев или лет.

43. Опыт работы в данной методологии и с данными инструментальными средствами – Method and tool experience. Срок работы участников проекта в данной методологии и с данными инструментальными средствами, суммарный по всем разработчикам и тестировщикам, и в пересчете на 1 человека; единица измерения – число месяцев или лет.

44. Время жизни дефекта – Defect longevity. Среднее и максимальное время от момента обнаружения дефекта до сообщения о его закрытии по всем выявленным и закрытым дефектам, плановое и по факту; единица измерения – число дней.

45. Процент незакрытых дефектов – Still open defects. Процент числа еще не закрытых дефектов в общем числе обнаруженных дефектов на данный момент времени.

46. Результативность (эффективность) обзоров – Review effectiveness. Процент числа дефектов, выявленных на обзорах, в общем числе выявленных дефектов на данный момент времени.

47. Рациональность (экономичность) обзоров – Review efficiency. Отношение суммарных трудозатрат на подготовку и проведение обзоров к общему числу найденных на обзорах дефектов – т.е., средняя стоимость нахождения одного дефекта путем обзора; единица измерения – человеко-час.

48. Инновационность проекта – Project innovativeness. Число рацпредложений и заявок на изобретение, поступивших от участников проекта по его тематике в ходе выполнения проекта и в течение некоторого срока (обычно до 6 месяцев) после его завершения.

49. Публикационность проекта – Project publicity. Число публикаций, включая защиты диссертаций, сделанных участниками проекта по его тематике в ходе выполнения проекта и в течение некоторого срока (обычно до 2 лет) после его завершения. Учитываются как открытые публикации, так и публикации для служебного пользования.

50. Загруженность оборудования – Equipment usage (load). Процент времени использования оборудования в суммарном времени, в течение которого оно находится в распоряжении участников проекта (по видам оборудования и лицензиям на ПО), плановый и по факту.

51. Использование уникального оборудования – Unique equipment usage. Суммарное время, потраченное на настройку и подготовку

уникального оборудования организации или заказчика и прогон на нем создаваемого продукта, плановое и по факту; единица измерения – число человеко-месяцев, человеко-дней или человеко-часов.

52. Проектные риски – Project risks. Общее число различных проектных рисков, включенных в план управления рисками.

53. Сработавшие риски – Triggered risks. Число актуализировавшихся проектных рисков, абсолютное и в отношении к общему числу проектных рисков, включенных в план управления рисками.

54. Переоценка рисков – Risk reassessment. Число пересмотров плана управления рисками (сколько раз производился пересмотр рисков по их перечню и характеристикам), плановое и по факту.

55. Смягчение последствий рисков – Risk mitigation. Трудозатраты на предотвращение и смягчение последствий рисков, плановые и по факту.

56. Воздействие рисков – Risk impact. Убытки в денежном выражении, которые проект может понести при срабатывании всех рисков и каждого в отдельности, плановые и по факту.

57. Декомпозиция проекта на продукты – Project decomposition to products. Число самостоятельных продуктов, появляющихся в результате данного проекта, если оно больше единицы.

58. Декомпозиция проекта на функциональные подгруппы – Project decomposition to functional subgroups. Число функциональных групп, на которое в итоге был разбит проект.

59. Компьютеры – Computers. Число компьютеров (настольных и др., включая тестовые и сервера), выделенных для данной группы.

60. Достаточность ресурсов – Sufficiency of resources. Процент необходимых по плану ресурсов в фактическом их объеме по видам ресурсов и усредненный по всем видам ресурсов.

61. Постоянство разработчиков – Staff continuity. Процент лиц, работавших от начала до конца проекта, в общем числе разработчиков.

62. Интенсивность взаимодействия внутри организации – Intensity of interaction within the organization. Число человеко-часов потраченных на общение внутри организации. Подсчитывается взаимодействие между разработчиками, между разработчиками и тестировщиками, между руководством проекта и командой.

63. Интенсивность взаимодействия с заказчиком – Intensity of interaction with customer. Число человеко-часов потраченных на общение с заказчиком.

64. Результативность (эффективность) тестирования – Effectiveness of testing. Число дефектов, выявленных путем тестирования, абсолютное и в отношении к общему числу выявленных дефектов.

65. Рациональность (экономичность) тестирования – Efficiency of testing. Трудозатраты на нахождение дефектов путем тестирования, абсолютные и в отношении к общему числу выявленных дефектов.

66. Метрики – Metrics. Число и наименования используемых в проекте метрик.

**2.3. Процессные метрики** характеризуют применяемый в создании продукта технологический процесс разработки и, соответственно, уровень зрелости организации, в которой этот процесс установлен.

1. Уровень зрелости разработчиков – Maturity level. Определяется путем официального оценивания или самооценивания; единица измерения – номер уровня по модели СММІ.

2. Повышение квалификации – Training. Среднее число часов для одного сотрудника на обучение и переподготовку в течение года, плановое и фактическое.

3. Выполненные разработки (проекты) – Completed projects. Число и наименования разработок (проектов), выполненных в организации за последние несколько лет. Определяется по ИБД организации.

4. База заказчиков – Customer base. Число и наименования разных заказчиков, для которых выполнялись разработки за последние несколько лет. Определяется по ИБД организации.

5. Портфель патентов – Patent portfolio. Число и наименования патентов, полученных данной организацией за пять несколько лет.

6. Текучесть кадров – Attrition rate. Процент сотрудников, покинувших организацию, в общем числе сотрудников в расчете на год.

7. Нарушения процесса – Process violations. Число зафиксированных нарушений процесса разработки за период по их категориям.

8. Улучшения процесса – Process improvements. Число поданных предложений по улучшению процесса за период.

9. Выработка ресурса – Burn-down chart. Число человеко-часов, остающихся на выполнение данной задачи; в методике Scrum наглядно отображает "сгорание" выделенного ресурса на выполнение задач по всему проекту в целом и по каждому спринту в отдельности.

10. Процессные активы – Project assets. Число шаблонов документов, моделей жизненного цикла, программных и тестовых наборов и других процессных активов, накопленных в организации.

11. Используемость процессных активов – Project assets usage. Процент повторно использованных процессных активов из базы процессных активов организации в общем числе процессных активов по отдельным активам и их группам.

12. Опытность команды – Staff experience. Процент ведущих опытных разработчиков в общем числе разработчиков.

**2.4. Референсные значения для некоторых метрик** по промышленности США за 2000 г. (см. [7]) приведены в таблице 1. Поскольку данные по отечественной промышленности программного обеспечения малодоступны, то организации-разработчики постепенно накапливают собственные данные по уже выполненным проектам и в дальнейшем используют их для сравнения и расчетов.

Таблица 1. Эталонные данные по промышленности США

Метрика	Единица измерения	Среднее	Лучшее
Производительность труда	КАЕЛОС на человеко-месяц	3.23	7.14
Стоимость строки кода	Доллары США на КАЕЛОС	4,334	1,962
Плотность дефектов	Число дефектов на КАЕЛОС	15.6	8.1
Эффективность сдерживания дефектов	%	95.0	99.5
Пост-релизные дефекты	Число дефектов на КАЕЛОС	0.780	0.041

В таблице 2 приведены усредненные данные отечественного предприятия Эксиджен [4], накапливавшиеся в течение 2008-2010 гг., по 69 выполненным проектам общим объемом 3 309 378 строк кода (LOC) на языке Си, что составляет 8 273 445 эквивалентных строк на языке ассемблера (АЕЛОС) — см. коэффициенты пересчета в таблице 4.

Таблица 2. Пример метрических данных из базы данных Эксиджен

Метрика	мин.	макс.	среднее
Размер проекта (КАЕЛОС)	3 578	1 535 498	135 630,2
Число найденных дефектов	9	4 484	362,0
Число разработанных тестов	4	10 890	987,7
% покрытия кода при модульном тестировании	11	93	62,4
Цикломатическая сложность кода	2	50	17,4

В модели SEER-SEM [8] интересны данные по типичному распределению трудоемкости в жизненном цикле разработки, которые можно использовать как начальное приближение при планировании проекта, а при планировании инспекций (обзоров) кода и документов можно опираться на следующие данные (таблица 3), пока в организации не накоплены собственные результаты по этим метрикам.

Таблица 3. Оптимальный темп инспекций

Шаг инспекции	Программный код	Текст
Запуск	15 мин	30 мин
Обзор	500 строк кода в час	560 строк текста в час
Подготовка	100 строк кода в час	140 строк текста в час
Инспекционное совещание	125 строк кода в час	140 строк текста в час
Выводы	30 мин	45 мин

Если разные компоненты программного продукта создаются на разных языках программирования, то размер этого продукта определяется в KAELOC – K Assembler Equivalent Lines Of Code, а для пересчета KLOC в KAELOC используются переходные коэффициенты, эмпирически установленные для разных языков программирования (таблица 4). С помощью этих же коэффициентов можно сравнивать размеры программных продуктов, написанных на разных языках.

Таблица 4. Пересчет KLOC в KAELOC для некоторых языков

Язык программирования	Коэффициент пересчета	Язык программирования	Коэффициент пересчета
Ada	4,5	LISP	1,5
Assembler	1,0	Macro-Assembler	1,0
C	2,5	Pascal	3,5
C++	11,0	Query languages	25,0
Forth	5,0	Unix shell scripts	1,5
FORTRAN	3,0	4-th Generation Languages	16,0

При определении качества программы (оценки числа остаточных дефектов) часто используется подход  $N\sigma$  (уровни сигма), где  $\sigma$  – среднеквадратическое отклонение при нормальном распределении вероятности наличия причины дефекта (ошибки) в исходном коде, а  $N$  – уровень сигма. Принятый в промышленности его пересчет в число дефектов (ошибок) в строках исходного кода приведен в таблице 5.

Таблица 5. Уровни сигма и оценка числа остаточных дефектов

Уровень сигма	Число дефектов (ошибок) на 1 миллион строк исходного кода	Уровень сигма	Число дефектов (ошибок) на 1 миллион строк исходного кода
1	691 462	4	6 210
2	308 538	5	233
3	66 807	6	3,4

В настоящее время в индустрии программного обеспечения уровень шесть сигма принят за эталон качества для надежного ПО.

**3. Сбор и хранение метрических данных.** Сбор и хранение метрических данных в организации обычно осуществляется группой обеспечения качества (SQA – Software Quality Assurance), которая ведет (курирует) каждый проект в организации и таким образом участвует в разработке. Собираемые для анализа метрические данные обычно хранятся в специальной базе данных, которая является частью процессных активов организации.

Член группы обеспечения качества, приписанный к данному проекту, ведет регулярный сбор метрических данных и проводит их анализ с помощью специальных инструментов. Результаты анализа сообщаются руководителю проекта и руководству организации, а также всем другим заинтересованным лицам. Задача группы обеспечения качества – отслеживать соответствие всех технологических процессов в разработке ПО принятым моделям и стандартам и заданным плановым показателям. При выявлении отклонений хода проекта от плана немедленно возбуждается анализ причин такого отклонения, и вырабатываются соответствующие поправочные действия.

**4. Представление и анализ метрик.** Собираемые метрики о ходе проекта регулярно анализируются его участниками для принятия решений о необходимости тех или иных поправочных действий для успешного завершения проекта. Регулярность такого анализа определяется выбранной моделью жизненного цикла и технологии разработки. Например, в технологии Scrum такой анализ выполняется ежедневно в начале каждого рабочего дня, тогда как в модели СММІ обычно приняты еженедельные отчеты всех участников разработки и ежемесячные операционные обзоры с участием руководства организации.

В конце разработки обычно проводится ретроспективный анализ всего хода работ, и делаются выводы на будущее по дальнейшим изменениям и усовершенствованиям типового процесса разработки в данной организации в соответствии с ее целями. В [10] приведены примеры ежемесячных одностраничных отчетов о ходе проекта, которые генерируются автоматически средствами MS Excel по данным из ИБД проекта на запрошенную дату. На таких ежемесячных операционных обзорах также сообщаются и анализируются и другие метрические показатели хода проекта: S-кривые по размеру кода и объему созданной документации, количеству найденных и закрытых дефектов, трудоемкости по разным типам работ и другие, определенные в проектном плане. Эти диаграммы, показывающие динамику изменения соответствующего показателя с течением времени, также строятся по табличным данным, которые допускают автоматическую обработку и более развитыми инструментами, например, такими как MathLab и др.

**5. Планирование проекта на основе метрик.** План разработки – это проигрывание будущих работ и инструмент для определения роли каждого участника. Он увязывает отдельные части разработки вместе, служит точкой отсчета для изменений и помогает понять, когда цель проекта достигнута, и проект может быть закончен. Этот план строится на основании структуры разбиения работ и оценок трудозатрат на исполнения выявленных блоков работы, которые устанавливаются по аналогичным проектам, выполненным ранее.

Общая схема предварительного планирования программного проекта представлена на рисунке 2. Она детализирует переход от исходных начальных требований к будущему программному продукту, полученных от заказчика, к начальной версии плана управления проектом в виде дорожной карты из ряда последовательных шагов.

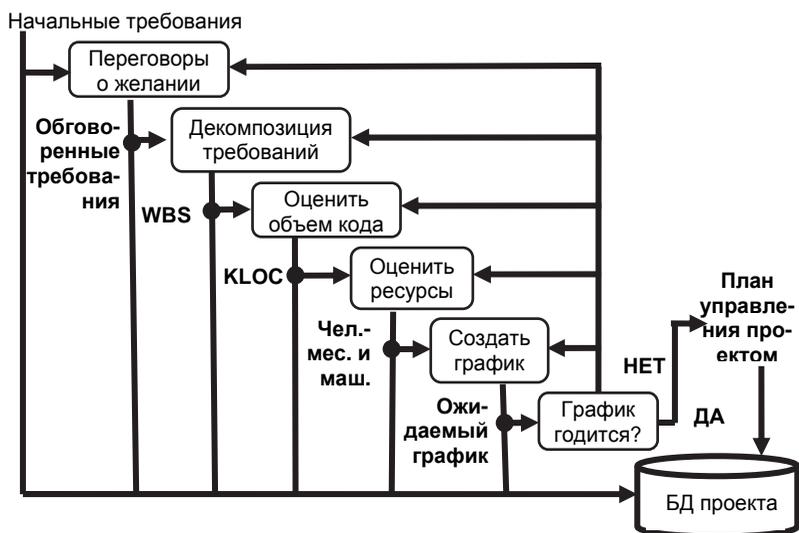


Рис. 2 Дорожная карта для планирования проекта

Исходя из начальных требований строится структура разбиения работ (Work Breakdown Structure – WBS). На следующем шаге эти требования уточняются и разносятся по подсистемам будущего продукта. Таким образом, на выходе этого шага появляется WBS, которая служит основой для дальнейшего планирования. По структуре разбиения работ и количеству требований, подлежащих реализации в каждом компоненте будущего продукта оценивается объем кода, подлежащий разработке в тысячах строк программного текста KLOC или в их ассемблерном эквиваленте KAELOC. Здесь существенным является при-

влечение уже известных данных по выполненным проектом для обоснования этих оценок. Затем, исходя из полученных оценок и трудоемкости разработки тысячи строк кода (с поправками на сложность, опыт исполнителей и другие факторы), оцениваются необходимые трудозатраты в человеко-месяцах и для некоторых проектов – необходимые дополнительные ресурсы по машинному времени и уникальному оборудованию. На основании этих данных строится ожидаемый график исполнения работ, учитывающих полученные оценки трудозатрат и взаимосвязи между компонентами в структуре разбиения работ. Заключительным шагом является принятие решения о том, устраивает ли полученный график работ как заказчика, так и исполнителя, и если это не так, то происходит возврат к одному из предыдущих шагов для внесения необходимых поправок.

Все получаемые исходные данные и промежуточные оценки вместе с их обоснованием накапливаются в базе данных проекта для последующего хранения и анализа при принятии решений.

Из целого ряда инструментов для составления предварительного плана работ в промышленности чаще всего используется модель издержки затрат COSOMO – COConstructive COSt MOdel, предлагающая три формулы для расчета важнейших плановых показателей:

$$\text{Трудоемкость} = a_b \times (KLOC)^{b_b} [\text{человеко – месяцев}]$$

$$\text{Срок}_\text{разработки} = c_b \times (\text{Трудоемкость})^{d_b} [\text{месяцев}]$$

$$\text{Число}_\text{разработчиков} = \frac{\text{Трудоемкость}}{\text{Срок}_\text{разработки}} [\text{человек}]$$

где коэффициенты  $a_b$ ,  $b_b$ ,  $c_b$  и  $d_b$ , разные для разных типов проектов и определенные усреднением по огромной и представительной выборке конкретных программных проектов, выполненных в разное время разными разработчиками, представлены в таблице 6:

Таблица 6. Коэффициенты базовой модели COSOMO

Тип проекта	$a_b$	$b_b$	$c_b$	$d_b$
Органический	2,4	1,05	2,5	0,38
Полуразделенный	3,0	1,12	2,5	0,35
Встроенный	4,6	1,20	2,5	0,32

Для органического (organic) типа характерна небольшая команда разработчиков с достаточным опытом разработки, и жесткими требованиями к создаваемому продукту. Полуразделенный (semi-detached) тип предполагает команду среднего размера со смешанным

опытом работы и различными по степени жесткости требованиями к продукту. Отличительной особенностью встроенного (embedded) типа является большое число жестких требований.

Таким образом, входные параметры для получения оценок трудоемкости, срока разработки и числа разработчиков в этой модели – это тип проекта и объем кода, который надо создать. Накапливая данные по фактическим трудозатратам в данной организации, можно совершенствовать эту модель, уточняя значения ее коэффициентов.

В начале 2000-х годов была предложена более развитая модель СОСОМО II [9], со значительно большим числом характеристик программного проекта и в силу этого дающая более точные оценки. Новой группой входных данных для оценки трудоемкости, по сравнению с базовой моделью СОСОМО, являются так называемые масштабируемые показатели программного обеспечения, подлежащего разработке. В так называемой промежуточной модели (intermediate model) СОСОМО II их значения выбираются по 6-ти бальной шкале: очень низкий, низкий, номинальный, высокий, очень высокий и сверхвысокий. В значении «номинальный» все эти коэффициенты равны 1 и отклоняются в ту или иную сторону для понижающих или повышающих оценок. В виде числовых множителей эти значения входят в формулу для трудоемкости как сомножители поправочного коэффициента EAF (effort adjustment factor), который изменяется в диапазоне от 0,9 до 1,4:

$$\text{Трудоемкость} = a_i \times (KLOC)^{b_i} \times EAF[\text{человеко – месяцев}]$$

где коэффициенты  $a_i$  и  $b_i$  те же, что и  $a_b$  и  $b_b$ , в базовой модели, за исключением органического:  $a_i = 3,2$  и встроенного:  $a_i = 2,8$  проектов. Формулы для оценки срока разработки и числа разработчиков и соответствующие коэффициенты  $c_i$  и  $d_i$  в промежуточной модели те же, что и  $c_b$  и  $d_b$  в базовой модели. В продвинутой модели (advanced model) СОСОМО II, дополнительно учитываются следующие показатели:

- *Атрибуты продукта – Product attributes*
  - Код для повторного использования – Developed for reusability;
  - Соответствие документации по продукту потребностям жизненного цикла – Documentation match to lifecycle needs;
- *Атрибуты команды разработчиков – Personnel attributes*
  - Преемственность коллектива – Personnel continuity;
  - Опыт работы с платформой – Platform experience;
- *Атрибуты проекта – Project attributes*

- Работа на нескольких площадках – Multi-site development.

В этой модели, наряду с показателями стоимости, учитываются еще и показатели масштаба ПО (Software Scale Drivers), значения которых также задаются по 6-ти бальной шкале, и к которым относятся:

- Наличие опыта аналогичных разработок – Precedentness;
- Гибкость самой разработки – Development flexibility;
- Решения по архитектуре/рискам – Architectural/risk resolution;
- Сплоченность команды разработчиков – Team cohesion;
- Уровень зрелости процесса разработки – Process maturity.

Разумеется, все оценки являются приблизительными и должны анализироваться и корректироваться в ходе проекта по мере того, как в ИБД проекта появляются и накапливаются фактические данные по трудозатратам на отдельных этапах и подэтапах разработки.

Наряду с представленными методами, планирование проекта можно осуществлять на основе проекта-аналога, который, по мнению эксперта, может служить дополнительным источником более точных данных, чем усредненные значения по ИБД организации, для оценивания соответствующих характеристик нового проекта. Система САМПО+ [11] поддержки создания ИБД позволяет на основе данных ИБД проводить объективную оценку выполнимости новых проектов при помощи автоматизированного поиска проекта-аналога.

Поиск аналога базируется на методологии рекомендуемого множества характеристик, которое формируется исходя из критерия функциональной пригодности информации, то есть ее актуальности, полноты и корректности. Только наличие всех трех критериев функциональной пригодности позволяет говорить о возможности классификации проектов в ИБД. Эта методология в общем случае состоит в последовательной проверке функциональной пригодности информации относительно: 1) *внесенных* в базу выполненных проектов; 2) *добавляемых* в базу выполненных проектов; и 3) *иницируемых* проектов.

Информация, хранимая в ИБД компании, функционально пригодна относительно *выполненных* проектов, если классификация проектов, полученная после кластеризации на всех предметно значимых поисковых наборах, согласуется с мнением экспертов.

Информация, хранимая в ИБД, функционально пригодна относительно *добавляемого* проекта, если данный проект может быть вписан в существующие классификаторы, либо путем включения в существующий класс, либо путем добавления нового класса в один из существующих классификаторов.

Информация, хранимая в ИБД, функционально пригодна отно-

сительно *иницируемого* проекта, если ему может быть сопоставлен какой-либо выполнявшийся ранее проект-аналог из ИБД.

В случае неудачи на каком-либо из перечисленных шагов состав характеристик ИБД расширяется и проводится ее повторное исследование на полноту, которое состоит из трех вышеуказанных этапов.

В рамках подхода, ориентированного на привлечение предметных знаний эксперта [12], с помощью классического алгоритма, основанного на вычислении оценок [13], модифицированного для решения поставленной задачи, формируется множество проектов (рекомендуемое множество), которые представляются эксперту в качестве возможных аналогов исследуемому проекту.

**6. Заключение.** Рассмотренный перечень измеряемых показателей является основой для объективного количественного управления ходом разработки ПО при заданных ограничениях. Не претендуя на исчерпывающую полноту, он, тем не менее, задает направление для постоянного совершенствования процесса разработки ПО в соответствии с моделью зрелости способностей СММИ для создания высококачественных и экономичных программных продуктов.

Специально рассмотрены вопросы планирования проекта с использованием методологии моделирования и принятия решений на основе алгоритмических сетей [14, 15] и исторической базы метрических данных по уже выполненным проектам, реализованной в системе САМПО+.

Задачей дальнейших исследований является расширение и уточнение этого списка, а также сбор и анализ фактических данных по выполненным проектам с целью установления текущей "базовой линии" как точки отчета для дальнейших улучшений процесса.

## Литература

1. CMMI® for Development. Version 1.3 // Carnegie-Mellon University. 2010. 450 p.
2. Babkin A.V. Moving to CMMI: Approach, Results and Lessons Learned // Proceedings of St. Petersburg IEEE Chapters. 2005. vol. 2. pp. 183–188.
3. Baranov S. An Industrial Technology of Test Automation Based on Verified Behavioral Models of Requirement Specifications for Telecommunication Applications // IEEE EUROCON. St. Petersburg 2009. pp. 122–129.
4. Тележкин А.М. Создание исторических баз данных при помощи системы САМПО+ // Региональная информатика (РИ-2012). Труды конференции. Санкт-Петербург, 24-26 октября 2012 г. СПб. 2013. pp 84–90.
5. Холстед М.Х. Начала науки о программах // М.: Фин. и статистика. 1981. 128 с.
6. Watson A.H., McCabe Th.J., Dolores R. Structured Testing: a Testing Methodology Using the Cyclomatic Complexity Metric // National Institute of Standards and Technology Special Publication 500-235. 1996. 123 p.
7. Capers, J. Software Assessments, Benchmarks, and Best Practice // Addison-Wesley.

2000. 339 p.
8. Сайт продукта SEER® by Galorath. URL: <http://www.galorath.com>. (Дата обращения: 18.10.2014).
  9. COCOMO® II - Constructive Cost Model. URL: <http://csse.usc.edu/tools/COCOMOII.php>. (Дата обращения: 18.10.2014).
  10. Баранов С.Н. и др. Процесс разработки программных изделий // М.: Наука. 2000. 176 с.
  11. Тележкин А.М. Система САМПО+ для создания и анализа исторической базы данных проектов // Известия ВУЗов. Приборостроение. 2014. Т.57. №11. С. 58–62.
  12. Морозов В.П. Поддержка принятия решений, ориентированная на знания эксперта // Региональная информатика (РИ-2010). Труды конференции. Санкт-Петербург, 20-22 октября 2010 г. СПОИСУ. СПб. 2011. С. 69–73.
  13. Журавлев Ю.И., Никифоров В.В. Алгоритмы распознавания, основанные на вычислении оценок // Кибернетика. 1971. №3. С. 1–11.
  14. Иванищев В.В., Марлей В.Е. Введение в теорию алгоритмических сетей // СПб. Изд-во СПбГТУ. 2000. 179 с.
  15. Соколов Б.В., Михайлов В.В., Морозов В.П. Методология и технология автоматизации аналитико-имитационного моделирования на основе алгоритмических сетей // Шестая всероссийская научно-практическая конференция «Имитационное моделирование. Теория и практика» ИММОД-2013. Казань. 2013. С. 256–262.

## References

1. CMMI® for Development. Version 1.3. Carnegie-Mellon University. 2010. 450 p.
2. Babkin A.V. Moving to CMMI: Approach, Results and Lessons Learned. Proceedings of St. Petersburg IEEE Chapters. 2005. vol.2. 2005. pp. 183–188.
3. Baranov S. An Industrial Technology of Test Automation Based on Verified Behavioral Models of Requirement Specifications for Telecommunication Applications. IEEE EUROCON, St. Petersburg. 2009. pp. 122–129.
4. Telezhkin A.M. [Creating Historical Databases with the SAMPO+ System]. *Regional'naja informatika (RI-2012). Trudy konferencii* [Regional Informatics (RI-2012)]. St.Petersburg, 24-26 October 2012. Conference Proceedings]. St.Petersburg, 2013. pp. 84–90. (In Russ.)
5. Halstead M.H. *Nachala nauki o programmah* [Elements of Software Science] М.: Finansy i statistika. 1981. 128 p. (In Russ.)
6. Watson A.H., McCabe Th.J., Dolores R. Structured Testing: a Testing Methodology Using the Cyclomatic Complexity Metric. National Institute of Standards and Technology Special Publication 500-235. 1996. 123 p.
7. Capers J. Software Assessments, Benchmarks, and Best Practice. Addison-Wesley. 2000. 339 p.
8. Site SEER® by Galorath. Available at: <http://www.galorath.com>. (Accessed: 18.10/2014).
9. COCOMO® II - Constructive Cost Model. Available at: <http://csse.usc.edu/tools/COCOMOII.php>. (Accessed: 18.10/2014).
10. Baranov S.N. et al. *Process razrabotki programmyh izdelij* [Software Development Process]. М: Nauka 2000. 176 p. (In Russ.)
11. Telezhkin A.M. [The SAMPO+ System for Creation and Analysis of Project Historical Database]. *Izvestiya vuzov. Priborostroyenie – Transactions of Universities. In-*

- strumentation. 2014. vol. 57. no. 11. pp. 58–62. (In Russ.)
12. Morozov V.P. [Decision-Making Support Oriented to Experts' Knowledge]. *Regional'naja informatika (RI-2010). Trudy konferencii* [Regional Informatics (RI-2010). Conference Proceedings]. SPOISU. SPb. 2011. pp. 69–73. (In Russ.)
  13. Zhuravlev Yu.I., Nikiforov V.V. [Recognition Algorithms base of Estimate Calculations]. *Kibernetika – Cybernetics*. 1971. vol. 3. pp. 1–11. (In Russ.)
  14. Ivanishchev V.V., Marley V.Ye. *Vvedenie v teoriju algoritmicheskikh setej* [Introduction to the Theory of Algorithmic Networks]. SPb. Izd-vo SPbGTU. 2000. 179 p. (In Russ.)
  15. Sokolov B.V., Mikhailov V.V., Morozov V.P. [Methodology and Technology for Automation of Analytical and Imitational Modeling Based on Algorithmic Networks]. *Shestaja vserossijskaja nauchno-prakticheskaja konferencija «Imitacionnoe modelirovanie. Teorija i praktika» IMMOD-2013* [Sixth All-Russia Scientific and Practical Conference "Imitational Modeling. Theory and Practice" IMMOD-2013]. Kazan. 2013. pp. 256–262. (In Russ.)

**Баранов Сергей Николаевич** — д-р физ.-мат. наук, главный научный сотрудник, Санкт-Петербургский институт информатики и автоматизации Российской академии наук (СПИИРАН), сотрудник международной научной лаборатории университета ИТМО. Область научных интересов: технология программирования, формальные методы. Число научных публикаций — 100. snbaranov@googlemail.com; 199178, Санкт-Петербург, 14 линия, д. 39; р.т. +7-812-328-0887.

**Baranov Sergey Nikolaevich** — Ph.D., Dr. Sci., chief researcher, St. Petersburg Institute for Informatics and Automation of Russian Academy of Sciences (SPIIRAS), researcher of international scientific laboratory of the university ITMO. Scientific interests: software engineering, formal methods in software development. The number of publications — 100. snbaranov@googlemail.com; 39, 14-th Line, St. Petersburg, 199178, Russia; office phone +7-812-328-0887.

**Тележкин Александр Михайлович** — аспирант, Санкт-Петербургский институт информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: технология программирования, формальные методы в разработке программного обеспечения. Число научных публикаций — 6. telezhkin@gmail.com; 192102, Санкт-Петербург, ул. Фучика 4к; р.т. +7-812-325-05-64.

**Telezhkin Alexander Mikhailovich** - Ph.D. student, St. Petersburg Institute for Informatics and Automation of Russian Academy of Sciences (SPIIRAS). Scientific interests: software engineering; formal methods in software development. The number of publications — 6. telezhkin@gmail.com; 4K, Fuchika st., St. Petersburg, 192102, Russia; office phone +7-812-325-05-64.

**Поддержка исследований.** Работа выполнена при государственной финансовой поддержке ведущих университетов Российской Федерации (субсидия 074-U01), университет ИТМО.

**Acknowledgements.** This research is partially supported by the leading universities of the Russian Federation (grant 074-U01), university ITMO.

## РЕФЕРАТ

### *Баранов С.Н., Тележкин А.М. Метрическое обеспечение программных разработок.*

Рассмотренный в статье перечень из 112 измеряемых показателей является основой для объективного количественного управления ходом разработки ПО при заданных ограничениях. Не претендуя на исчерпывающую полноту, он, тем не менее, задает направление для постоянного совершенствования процесса разработки ПО в соответствии с моделью зрелости способностей СМММ для создания высококачественных и экономичных программных продуктов.

Специально рассмотрены вопросы планирования проекта с использованием методологии моделирования и принятия решений на основе алгоритмических сетей и исторической базы метрических данных по уже выполненным проектам, реализованной в системе САМПО+.

Сформулирована задача дальнейших исследований – расширение и уточнение этого списка, а также сбор и анализ фактических данных по выполненным проектам с целью установления текущей "базовой линии" как точки отчета для дальнейших улучшений процесса.

## SUMMARY

### *Baranov S.N., Telezhkin A.M. Metrics for Software Development.*

The considered list of 112 measurable indications is a basis for objective quantitative managing a software development project under given constraints. Being incomplete, it nevertheless determines the direction for continuous improvement of the software process in accordance with the CMMI model for creating high-quality and efficient software products.

Planning issues under the methodology of modeling and decision-making based on algorithmic networks and historical database of metrics of completed projects within the system SAMPO+ are considered as well.

Directions for further study are defined: extending and refining the list of metrics as well as accumulating and analyzing actual data on completed projects in order to establish project and process baselines for further process improvements.

М.В. ЦВЕТКОВ, О.В. СМИРНОВА, Ф.Р. ГАЛЬЯНО  
**СИСТЕМА МОНИТОРИНГА ЛЕДОВОЙ ОБСТАНОВКИ И  
ОБЕСПЕЧЕНИЯ БЕЗОПАСНОГО СУДОХОДСТВА ПО  
СЕВЕРНОМУ МОРСКОМУ ПУТИ НА БАЗЕ  
ИНТЕЛЛЕКТУАЛЬНОЙ ГИС**

---

*Цветков М.В., Смирнова О.В., Гальяно Ф.Р. Система мониторинга ледовой обстановки и обеспечения безопасного судоходства по Северному морскому пути на базе интеллектуальной ГИС.*

**Аннотация.** В статье обсуждается проблема разработки системы мониторинга ледовой обстановки в Арктике и обеспечения безопасности судоходства по Северному морскому пути с использованием интеллектуальных ГИС. Демонстрируется пример корректировки маршрута движения судов по участку Северного морского пути с учетом ледовой обстановки в районе на основе разработанной системы мониторинга.

**Ключевые слова:** анализ данных, интеллектуальные ГИС, оценка ледовой обстановки, безопасность судоходства.

*Tsvetkov M.V., Smirnova O.V., Galiano P.R. Ice Situation and Safe Navigation Monitoring System on the Northern Sea Route Based on Intelligent GIS.*

**Abstract.** In this paper we discuss the problem of ice situation in the Arctic and safe navigation monitoring system development for the Northern Sea Route with application of intelligent geoinformation system. Case study illustrates the adjustment of the ship route on the part of the Northern Sea Route in accordance with the ice situation using the developed monitoring system.

**Keywords:** data analysis, intelligent GIS, ice situation assessment, safe navigation.

---

**1. Введение.** Арктический регион является одной из наиболее перспективных территорий в составе РФ. В последние годы в Арктике происходит интенсивное освоение нефтяных и газовых месторождений, развивается судоходство в акваториях арктических морей, решаются различные задачи, связанные с обороноспособностью страны. Северный морской путь является одной из крупнейших морских транспортных магистралей Арктики и активно используется для перевозки полезных ископаемых, добываемых российскими компаниями за полярным кругом, снабжения удаленных населенных пунктов Крайнего Севера, а также решения других важных задач [1].

В то же время арктический регион характеризуется суровыми климатическими условиями и экстремальной ледовой обстановкой. Во многих районах Арктики устойчивый ледовый покров держится в течение 7 месяцев, большая часть территории покрыта вечной мерзлотой. Все эти факторы могут значительно влиять на решение многих важных задач, связанных с освоением Арктики.

Сложная ледовая обстановка во многих районах прохождения Северного морского пути значительно осложняет судоходство по этому перспективному во многих отношениях маршруту. Одним из очевидных путей решения данной проблемы является использование сил Ледокольного флота РФ. Однако использование атомных ледоколов для проводки судов по пути не может быть качественно спланировано без подробной оценки ледовой обстановки в регионе и решения сопутствующих навигационных задач.

Обеспечение режима безопасного судоходства на Северном морском пути невозможно без современных средств мониторинга ледовой обстановки и поддержки принятия решений. Например, анализ ледовой обстановки предусматривает периодическую разработку и обновление ледовых карт, а маршруты проводки судов по участкам пути могут достаточно часто корректироваться в связи с изменением различных условий судоходства. Пространственная специфика таких задач приводит к необходимости использования географических информационных систем (ГИС) в качестве информационной базы для построения перспективных систем мониторинга морской обстановки на Северном морском пути.

Существующие ГИС отечественного и иностранного производства находят широкое применение, в первую очередь, в задачах анализа данных по ледовой обстановке и составления ледовых карт. В качестве основного источника данных о ледовой обстановке в Арктике выступают различные космические системы дистанционного зондирования Земли (ДЗЗ) [1, 2]. Результатом обработки соответствующих данных ДЗЗ является бумажная или электронная карта ледовой обстановки в интересующем регионе [3, 4], а основная цель такой деятельности – выдача карты конечному потребителю (лицу, принимающему решения (ЛПР)); например, начальнику порта, капитану и судна и т. д.) [1]. Однако, кроме данных о ледовой обстановке, ЛПР получает другую важную информацию от доступных источников, которая также должна быть учтена при принятии ответственных решений по обеспечению безопасного судоходства.

Целью данной статьи является анализ возможностей использования интеллектуальной ГИС для мониторинга ледовой обстановки в Арктике и обеспечения безопасности судоходства по Северному морскому пути. По мнению авторов, такая система способна обеспечить интеллектуальный анализ данных (в том числе, данных ДЗЗ), визуализировать в необходимом для понимания виде результаты анализа, и, что более важно, осуществить поддержку принятия сложных решений,

связанных с возникновением опасных ситуаций на Северном морском пути.

**2. Интеллектуальная ГИС как основа для построения системы мониторинга ледовой обстановки и обеспечения безопасного судоходства.** Архитектура системы мониторинга ледовой обстановки и обеспечения безопасного судоходства в Арктическом регионе представлена на рисунке 1. Основой построения данной системы служит интеллектуальная геоинформационная система (ИГИС) под которой понимается геоинформационная система, которая включает в свой состав интегрированные подсистемы искусственного интеллекта, а также прикладные компоненты, позволяющие обеспечить всестороннюю поддержку принятия решений оператором системы мониторинга [5, 6].

Ядром системы мониторинга ледовой обстановки и обеспечения безопасного судоходства является база знаний, включающая онтологию. Онтология представляет собой «каркас» для представления основных понятий и связей между ними предметной области. Основной задачей базы знаний является хранение и накопление данных, информации и знаний о реальных объектах предметной области.

Экспертная система используется для разработки, наглядного представления и реализации сценариев обработки данных в целях распознавания и прогнозирования ледовой обстановки в регионе, а также интеллектуального сопровождения решения задач безопасного судоходства. Результаты работы экспертной системы предоставляются пользователю в виде интерактивных подсказок или в виде диалогов (рисунок 2).

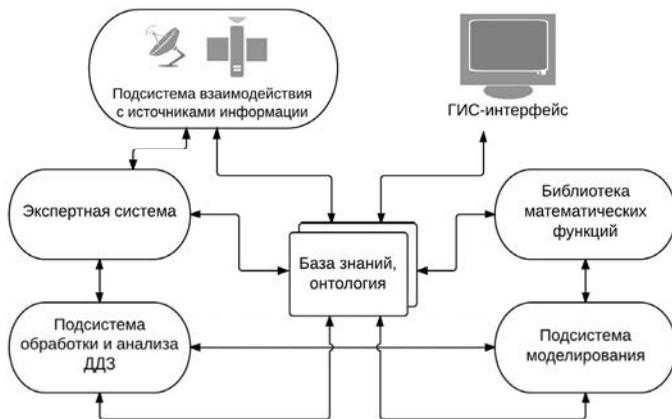


Рис. 1. Архитектура системы мониторинга ледовой обстановки

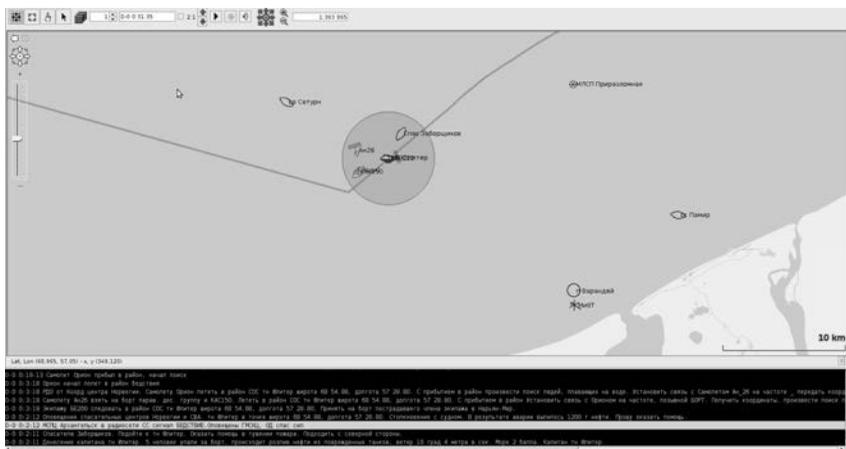


Рис. 2. Пример работы экспертной системы мониторинга безопасности судоходства в морском районе

Основными источниками информации системы мониторинга ледовой обстановки и обеспечения безопасного судоходства об обстановке в районе являются специальные системы наблюдения, использующие различные принципы сбора и первичной обработки данных. К таким системам можно отнести: радиолокационные и гидроакустические станции, оптико-электронные системы наблюдения, различные системы дистанционного зондирования и т. д.

В последние десятилетия важнейшим источником информации о ледовой обстановке являются космические системы дистанционного зондирования Земли.

Космические системы ДЗЗ предназначены для получения информации о состоянии объектов естественного и искусственного происхождения на поверхности Земли и в акватории Мирового океана, природных ресурсов Земли и происходящих в атмосфере метеорологических процессах. Глобальность и высокая информативность наблюдений из космоса позволяют решать широкий круг специальных задач, проводить картографирование, исследование природных ресурсов Земли, метеорологические и океанографические наблюдения. В интересах решения этих задач может осуществляться плановое, периодическое и оперативное наблюдение из космоса за различными объектами и районами.

В настоящее время в мире многими странами эксплуатируются десятки космических аппаратов (КА) ДЗЗ. В различных стадиях разработки находятся сотни новых проектов по реализации перспективных

возможностей наблюдения и съемки из космоса, одновременно наблюдается значительный прогресс в области повышения уровня технической оснащенности космических аппаратов и сокращения затрат на их создание и эксплуатацию. Все это обеспечивается не только за счет применения качественно новых конструкционных материалов и методов проектирования, минимизации массогабаритных характеристик, разработки унифицированных орбитальных платформ, а также благодаря интеллектуализации бортовых и наземных средств на основе современных компьютерных технологий [7].

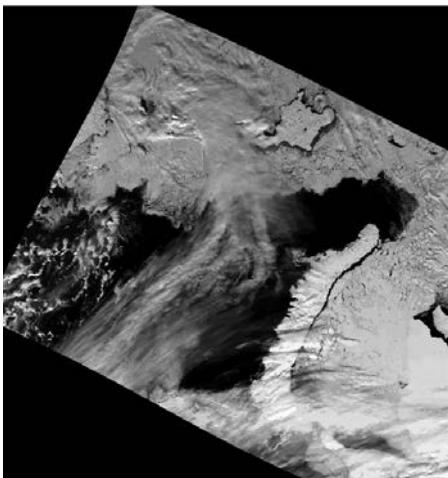
Получаемые от космической системы мониторинга ледовой обстановки данные представляют собой набор снимков (изображений) района определенного разрешения и размера, обработка которых позволяет составлять динамические (например, с интервалом 1–2 дня) ледовые карты. Для обработки снимков используются современные алгоритмы обработки данных ДЗЗ, реализованные в составе программных средств системы мониторинга.

Подсистема обработки и анализа данных ДЗЗ позволяет в интерактивном режиме получать, выбирать варианты расчета, задавать исходные данные с помощью пользовательского интерфейса, просматривать результаты обработки, повторять расчеты и накапливать результаты в выходных файлах.

Алгоритмы обработки данных ДЗЗ представляют собой комбинацию алгоритмов обработки изображений (нормализация, подавление шумов и т. п.), выделения информативных элементов изображения (сегментов, контуров), расчета их признаков (геометрических, яркостных, текстурных) и последующей классификации. Различие целей анализа и изменчивость входных данных приводит к необходимости использования широкого набора различных алгоритмов на всех этапах анализа. С другой стороны, спецификой ДЗЗ являются значительные объемы данных и необходимость взаимосвязанного анализа различных спектральных диапазонов, что налагает ограничения на выбор набора алгоритмов, пригодных для решения задачи [8].

Исходными данными служат данные ДЗЗ (изображения) и соответствующая им метаинформация (рисунок 3).

Традиционная часть ГИС – это ГИС-интерфейс, предназначенный для визуального представления пространственных данных в различных географических цифровых форматах и объектов, хранящихся в базе знаний, адаптированный к представлению данных о ледовой обстановке в арктическом регионе.



```
GeoTIFF_information:
Version: 1
Key_Revision: 1.0
Tagged_Information:
ModelTiepointTag (2, 3):
0 0 0
-1092018.06121026-574639.0225332630
ModelPixelScaleTag (1, 3):
250 250 0
End_Of_Tags.
Keyword_Information:
GTModelTypeGeoKey (Short,1): ModelTypeProjected
GTRasterTypeGeoKey (Short,1): RasterPixelFormatArea
GTCitationGeoKey (Ascii,8): "unnamed"
GeographicTypeGeoKey (Short,1): GCS_WGS_84
ProjectedCSTypeGeoKey (Short,1): User-Defined
ProjectionGeoKey (Short,1): User-Defined
ProjCoordTransGeoKey (Short,1): CT_PolarStereographic
ProjLinearUnitsGeoKey (Short,1): Linear_Meter
ProjNatOriginLatGeoKey (Double,1): 90
ProjFalseRastingGeoKey (Double,1): 0
ProjFalseNorthGeoKey (Double,1): 0
ProjScaleNatOriginGeoKey (Double,1): 1
ProjStraightVertPoleLongGeoKey (Double,1): 55
End_Of_Keys.
End_Of_GeoTIFF.

Projection Method: CT_PolarStereographic
ProjNatOriginLatGeoKey: 90.000000 ° 90d 0' 0.00"N
ProjStraightVertPoleLongGeoKey: 55.000000 ( 55d 0' 0.00"E)
ProjScaleNatOriginGeoKey: 1.000000
ProjFalseRastingGeoKey: 0.000000 m
ProjFalseNorthGeoKey: 0.000000 m
GCS: 4326/WGS 84
Datum: 6326/World Geodetic System 1984
Ellipsoid: 7030/WGS 84 (6378137.00,6356752.31)
```

Рис. 3. Исходные данные ДЗЗ и метainформация [2]

Подсистема моделирования предназначена для моделирования различных геопространственных процессов. Она позволяет осуществлять визуальную разработку соответствующих сценариев, выдачу рекомендаций для принятия решений, оценку принятых решений, контроль за развитием процессов и действий их участников на основе технологии экспертных систем и представления знаний на основе онтологий. Сценарии описываются в виде двумерных ориентированных графов, узлами которых являются отдельные задачи сценария, и точки принятия решений, в которых происходит разветвление сценария на различные ветви в зависимости от выполнения или невыполнения определенных условий [9].

Существенно облегчить сложность вычислительного процесса, разработку отдельных задач при создании сценариев помогает библиотека математических функций. Она содержит специальный математический аппарат для решения задач моделирования и анализа ледовой обстановки, а также задачи обеспечения безопасного судоходства в Арктическом регионе.

**3. Применение системы мониторинга для оценки ледовой обстановки и корректировки маршрутов движения судов по Северному морскому пути.** Оценка ледовой обстановки начинается с разработки отдельных задач и компоновки их в единый сценарий моделирования геопространственных процессов. Сценарий обработки снимков разрабатывается визуально с помощью технологии Drag-and-Drop (рисунок 4). В левой панели окна разработки сценариев находится

ся палетка со значками элементов сценария, которые в ходе разработки перетаскиваются на канву центральной панели и затем соединяются линиями с помощью мыши. Прямоугольники представляют собой отдельные задачи, включающие функции из библиотеки математических функций, решаемые в ходе выполнения сценария. В поле Initial\_Tasks указываются исходные задачи и/или решения, с которых начинается вычислительный процесс. В поле Final\_Tasks указываются задачи, после выполнения которых, заканчивается вычислительный процесс. В поле Finish\_Type указывается тип окончания сценария: или при окончании всех конечных задач (AND), или при окончании хотя бы одной из них (OR) [9, 10].

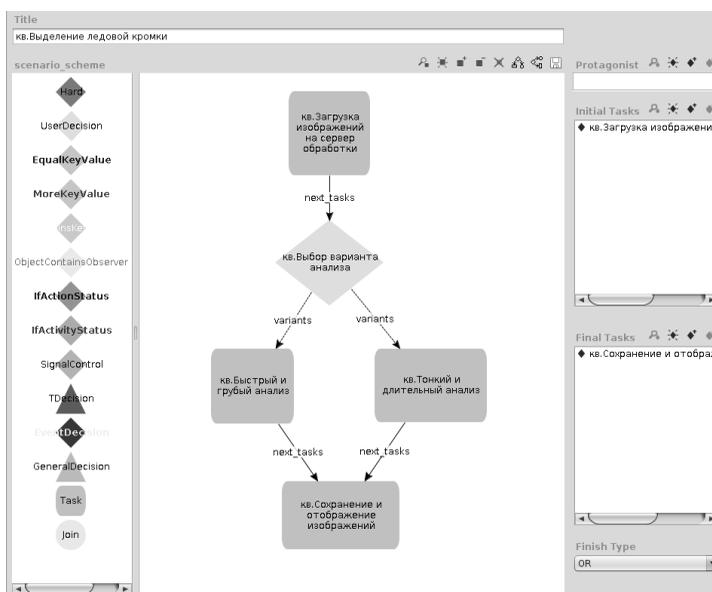


Рис. 4. Окно разработки сценария обработки данных ДЗЗ

Отдельные задачи сценария разрабатываются также визуально, как показано на рисунке 4. В окне разработки задач в палетке слева находятся значки типовых действий, из которых конструируется схема решения задачи на канве в центральной части окна. В правой части окна находятся поля начальных и конечных действий, а также следующих задач, которые будут выполняться после окончания данной задачи. Поле имеет тот же смысл, что и в окне разработки моделирующей программы, но в отношении конечных действий.

После запуска сценария появляются исходное окно выполнения и пользовательский интерфейс. Исходное окно сценария (рисунок 5) снабжено подсказками для пользователя системы. В частности, пользователю предлагается загрузить с помощью пользовательского интерфейса анализируемое изображение и экспертное изображение (в скобках указываются значения по умолчанию).

В ходе выполнения сценария обработки возможен выбор двух вариантов проигрывания:

– быстрый и грубый анализ; используется только один заданный уровень сегментации изображения;

– тонкий и длительный анализ; в данном случае используется иерархия уровней сегментации изображения.

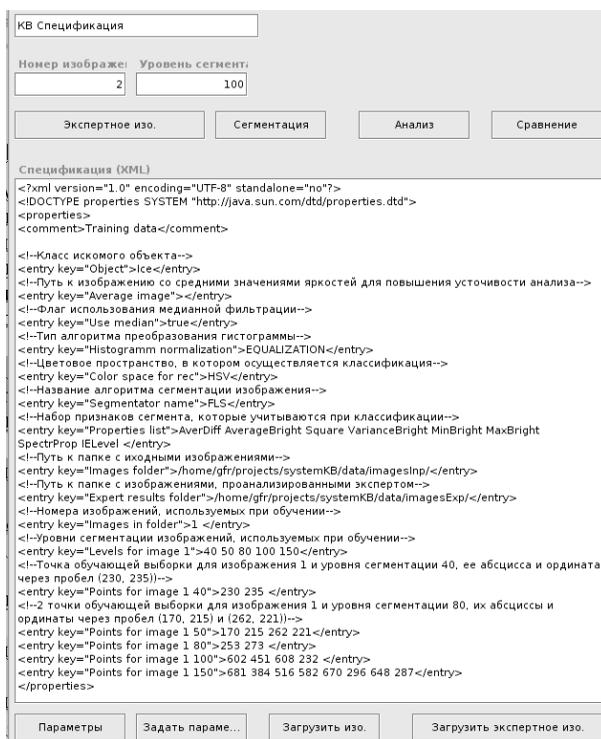


Рис. 5. Пользовательский интерфейс для запуска сценария

Пользовательский интерфейс позволяет пользователю задавать исходные параметры для дальнейшей обработки исходных данных, в

том числе и данных ДЗЗ. К исходным параметрам относятся номер исходного изображения, уровень сегментации изображения, экспертное изображение, алгоритм сегментации изображения. Все исходные параметры задаются в спецификации XML (рисунок 6).

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<!DOCTYPE properties SYSTEM "http://java.sun.com/dtd/properties.dtd">
<properties>
  <comment>Training data</comment>
  <entry key="Object">Ice</entry>
  <entry key="Average image"></entry>
  <entry key="Use median">true</entry>
  <entry key="Segmentator name">FLS</entry>
  <entry key="Histogram normalization">EQUALIZATION</entry>
  <entry key="Color space for rec">HSV</entry>
  <entry key="Properties list">AverDiff AverageBright Square VarianceBright MinBright MaxBright SpectrProp IELevel </entry>
  <entry key="Images in folder">1 </entry>
  <entry key="Levels for image 1">40 50 80 100 150</entry>
  <entry key="Points for image 1 40">230 235 </entry>
  <entry key="Points for image 1 50">170 215 262 221</entry>
  <entry key="Points for image 1 80">253 273 </entry>
  <entry key="Points for image 1 100">602 451 608 232 </entry>
  <entry key="Points for image 1 150">681 384 516 582 670 296 648 287</entry>
  <entry key="Images folder">/home/gfr/projects/systemKB/data/imagesInp</entry>
  <entry key="Expert results folder">/home/gfr/projects/systemKB/data/imagesExp</entry> </properties>
```

Рис. 6. Структура спецификации XML

Для просмотра информации о параметрах, заданных пользователем, необходимо нажать кнопку Параметры пользовательского интерфейса (рисунок 7).

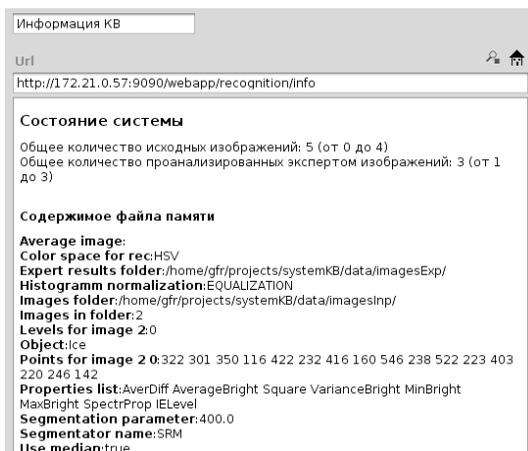


Рис. 7. Окно параметров сценария обработки данных ДЗЗ

Результат обработки данных ДЗЗ доступен после нажатия кнопки Сегментация пользовательского интерфейса (рисунок 8). Кроме того, пользователь может сравнить визуально и с помощью расчетных параметров (рисунок 9) полученные результаты с экспертными значениями. Окончательный результат анализа данных ДЗЗ приведен на рисунке 10.

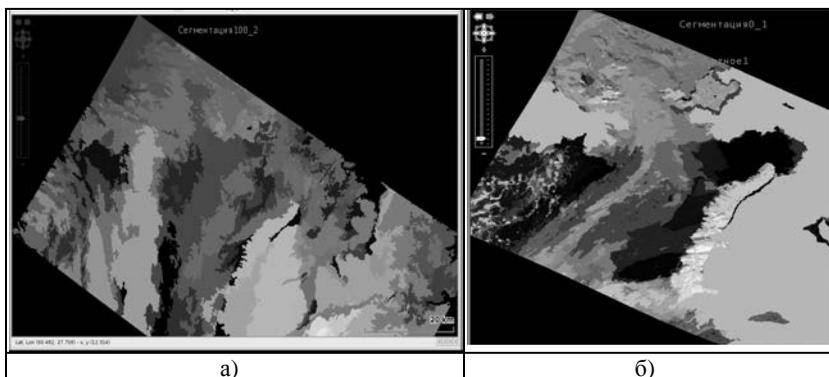


Рис. 8. Окна просмотра результаты сегментации данных ДЗЗ во время выполнения сценария: а) результат сегментации данных ДЗЗ; б) сравнение полученного результата сегментации с экспертным изображением

Результаты сравнения

Url <http://172.21.0.57:9090/webapp/recognition/compare>

**Сравнение с результатами экспертного анализа**

Общее количество изображений 3  
 Общее количество пикселей 2126490

---

Абсолютное количество правильно классифицированных пикселей 1826083  
 Абсолютное количество ложных срабатываний 54481  
 Абсолютное количество пропусков объекта 245926

---

Относительное количество правильно классифицированных пикселей 0.858731054460637  
 Относительное количество ложных срабатываний 0.025620153398323058  
 Относительное количество пропусков объекта 0.11564879214103993

Рис. 9. Результирующие параметры обработки данных ДЗЗ

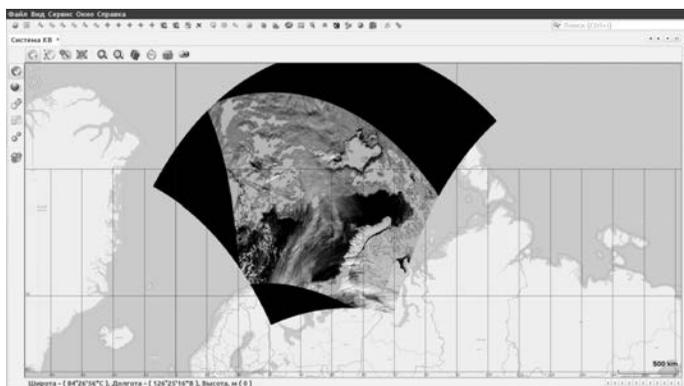


Рис. 10. Результат анализа данных ДЗЗ (ледовая кромка)

Результаты выделения ледовой кромки являются основой для составления оперативной электронной ледовой карты (рисунок 11). Оператор системы мониторинга может использовать ледовую карту региона для поддержки принятия решений, связанных с обеспечением безопасного судоходства на Северном морском пути.

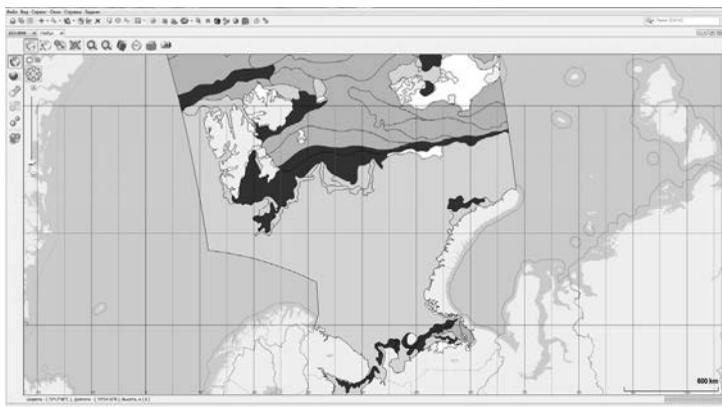


Рис. 11. Ледовая карта участка акватории Северного морского пути в системе мониторинга

С этой целью данные ледовой карты используются как факты для экспертной системы, которая посредством применения специальных правил уведомляет оператора о возникновении потенциально опасных ситуаций, необходимости закрытия морских районов и смены маршрутов движения судов. Также ЭС может вырабатывать рекомендации по применению судов ледокольного флота РФ для проводки судов по участкам Северного морского пути. На рисунке 12 показан пример проведения транспортного судна по участку Северного морского пути с учетом ледовой обстановки, полученной в результате обработки данных ДЗЗ, с использованием системы мониторинга ледовой обстановки и обеспечения безопасного судоходства. Локализация льда в южной части акватории пролива (рисунок 12) приводит к необходимости использования его северного участка для прокладки маршрутов движения судов.

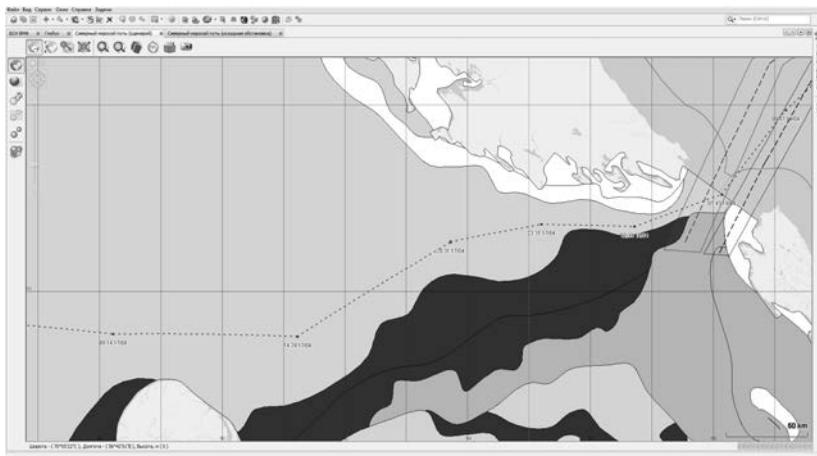


Рис. 12. Поддержка принятия решения по корректировке маршрутов движения судов по участку Северного морского пути

**6. Заключение.** Данные ДЗЗ являются одним из наиболее перспективных способов получения информации о ледовой обстановке в районах прохождения Северного морского пути. Интеллектуальная обработка данных ДЗЗ невозможна без привлечения современных методов пространственного анализа и средств искусственного интеллекта (ИИ).

Интеллектуальные ГИС, включающие в свой состав интегрированные средства ИИ, а также прикладные компоненты количественного обоснования вырабатываемых рекомендаций, могут использоваться не только как информационная база для составления оперативных карт ледовой обстановки, но как основа для построения системы мониторинга ледовой обстановки и обеспечения безопасного судоходства на Северном морском пути.

### Литература

1. Официальный сайт Администрации Северного морского пути. URL: <http://www.nsga.ru> (дата обращения 24.06.2014)
2. Официальный сайт Арктического и Антарктического научно-исследовательского института. URL: <http://www.aari.nw.ru> (дата обращения 01.07.2014)
3. Мониторинг ледовой обстановки. URL: [http://www.scanex.ru/ru/data/Applications\\_ScanEx\\_p49-71.pdf](http://www.scanex.ru/ru/data/Applications_ScanEx_p49-71.pdf) (дата обращения 01.07.2014)
4. Shields B. Ships Use Sea-Ice and Iceberg Maps to Navigate in Polar Regions. URL: <http://www.esri.com/esri-news/arcwatch/0213/ships-use-sea-ice-and-iceberg-maps-to-navigate-in-polar-regions> (дата обращения 01.07.2014)

5. Интеллектуальные географические информационные системы для мониторинга морской обстановки / под общ. ред. чл.-кор. РАН Юсупова Р. М. и д-ра техн. наук Поповича В.В. СПб.: Наука. 2013. 284 с.
6. *Popovich V. Intelligent GIS Conceptualization // Proceedings of the Sixth International Workshop "Information Fusion and Geographic Information Systems" (IF&GIS 2013). 2014. pp. 17–44.*
7. Концепция развития российской космической системы дистанционного зондирования Земли на период до 2025 года. URL: [www.gisa.ru/file/file766.doc](http://www.gisa.ru/file/file766.doc) (дата обращения 01.07.2014)
8. *Galiano F., Kharinov M., Vanurin S. Application of Sleator-Tarjan Dynamic Trees in a Monitoring System for the Arctic Region Based on Remote Sensing Data // Proceedings of the Sixth International Workshop "Information Fusion and Geographic Information Systems" (IF&GIS 2013). 2014. pp. 137–147.*
9. *Sorokin R. Advantages of Intelligent Geographic Information System Research Prototype // Proceedings of the Fourth International Workshop "Information Fusion and Geographic Information Systems: Towards the Digital Ocean". 2011. pp 33–44.*

## References

1. Official'nyj sajt Administracii Severnogo morskogo puti [Official web site of the Northern sea route administration]. Available at: <http://www.nsra.ru> (accessed 24.06.2014) (In Russ.).
2. Official'nyj sajt Arkticheskogo i Antarkticheskogo nauchno-issledovatel'skogo instituta [Official web site of the Arctic and Antarctic Research Institute]. Available at: <http://www.aari.nw.ru> (accessed 01.07.2014) (In Russ.).
3. *Monitoring ledovoj obstanovki* [Ice situation monitoring]. Available at: [http://www.scanex.ru/ru/data/Applications\\_ScanEx\\_p49-71.pdf](http://www.scanex.ru/ru/data/Applications_ScanEx_p49-71.pdf) (accessed 01.07.2014) (In Russ.).
4. Shields B. Ships Use Sea-Ice and Iceberg Maps to Navigate in Polar Regions. Available at: <http://www.esri.com/esri-news/arcwatch/0213/ships-use-sea-ice-and-iceberg-maps-to-navigate-in-polar-regions> (accessed 01.07.2014).
5. *Intellektual'nye geograficheskie informacionnye sistemy dlja monitoringa morskoy obstanovki. pod obshh. red. chl.-kor. RAN Jusupova R. M. i d-ra tehn. nauk Popovicha V.V.* [Intelligent geographic information system for maritime monitoring]. edited by Jusupov R. M., Popovich V.V. St.Petersburg: Nauka. 2013. 284 p. (In Russ.).
6. Popovich V. Intelligent GIS Conceptualization. Proceedings of the Sixth International Workshop "Information Fusion and Geographic Information Systems" (IF&GIS 2013). 2014. pp. 17–44.
7. *Koncepcija razvitija rossijskoj kosmicheskoy sistemy distancionnogo zondirovanija Zemli na period do 2025 goda* [The concept of the Russian space system of the Earth's remote sensing development for the period till 2025]. Available at: [www.gisa.ru/file/file766.doc](http://www.gisa.ru/file/file766.doc) (accessed 01.07.2014) (In Russ.).
8. Galiano F., Kharinov M., Vanurin S. Application of Sleator-Tarjan Dynamic Trees in a Monitoring System for the Arctic Region Based on Remote Sensing Data. Proceedings of the Sixth International Workshop "Information Fusion and Geographic Information Systems" (IF&GIS 2013). 2014. pp. 137–147.
9. Sorokin R. Advantages of Intelligent Geographic Information System Research Prototype. Proceedings of the Fourth International Workshop "Information Fusion and Geographic Information Systems: Towards the Digital Ocean". 2011. pp 33–44.

**Цветков Михаил Владимирович** — к-т техн. наук, старший научный сотрудник лаборатории объектно-ориентированных географических информационных систем (НИЛ ООГИС) Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: географические информационные системы, методы и технологии искусственного интеллекта, обработка геопространственных данных. Число научных публикаций — 37. [tmv@oogis.ru](mailto:tmv@oogis.ru), [www.oogis.ru](http://www.oogis.ru); СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)355-9682, факс +7(812)355-9674.

**Tsvetkov Mikhail Vladimirovich** — Ph.D., senior researcher of Research laboratory of object-oriented geoinformation systems (OOGIS RL), St. Petersburg Institute for Informatics and Automation of Russian Academy of Sciences (SPIIRAS), Russia. Research interests: geographic information systems, artificial intelligence, geospatial data processing. The number of publications — 37. [tmv@oogis.ru](mailto:tmv@oogis.ru), [www.oogis.ru](http://www.oogis.ru); SPIIRAS, 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)355-9682, fax +7(812)355-9674.

**Смирнова Оксана Вячеславовна** — к-т техн. наук, старший научный сотрудник базовой кафедры КТЭА Санкт-Петербургского государственного электротехнического университета «ЛЭТИ» им. В.И. Ульянова (Ленина) (СПбГЭТУ). Область научных интересов: математическое моделирование сложных процессов, геоинформационные технологии. Число научных публикаций — 42. [sov@oogis.ru](mailto:sov@oogis.ru); ул. Профессора Попова, д.5, 197376, Санкт-Петербург, РФ; р.т. +7(812)355-9682, факс +7(812)355-9674

**Smirnova Oksana Vjacheslavovna** — Ph.D., senior researcher of St.Petersburg Electrotechnical University "LETI". Research interests: mathematical modeling of complex process, geoinformation technology. The number of publications — 42. [sov@oogis.ru](mailto:sov@oogis.ru); 5, str. prof. Popova, St.Petersburg, Russia; office phone +7(812)355-968, fax +7(812)355-9674.

**Гальяно Филипп Робертович** — к-т техн. наук, сотрудник ЗАО СПИИРАН-НТБВТ. Область научных интересов: обработка изображений и распознавание образов. Число научных публикаций — 14. [galiano@oogis.ru](mailto:galiano@oogis.ru); СПИИРАН-НТБВТ, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р. тел.: +7 911 835-26-59.

**Galiano Philip Robertovich** — Ph.D., specialist, "Hi Tech Research and Development Office Ltd." (HTR&DO). Research interests: image processing and pattern recognition. The number of publications — 14. [galiano@oogis.ru](mailto:galiano@oogis.ru); SPIIRAS, 39, 14-th Line V.O., St.Petersburg, 199178, Russia; office phone +7 911 835-26-59.

**Поддержка исследований.** Работа выполнена в СПбГЭТУ при финансовой поддержке Министерства образования и науки Российской Федерации в рамках договора № 02.G25.31.0058 от 12.02.2013 г.

**Acknowledgements.** This research is supported by Ministry of Education and Science of the Russian Federation (№ 02.G25.31.0058 from 12.02.2013).

## РЕФЕРАТ

### *Цветков М.В., Смирнова О.В., Гальяно Ф.Р.* Система мониторинга ледовой обстановки и обеспечения безопасного судоходства по Северному морскому пути на базе интеллектуальной ГИС.

Арктический регион является одной из наиболее перспективных территорий составе РФ. В последние годы в Арктике происходит интенсивное освоение нефтяных и газовых месторождений, развивается судоходство в акваториях арктических морей, решаются различные задачи, связанные с обороноспособностью страны. Северный морской путь является одной из крупнейших морских транспортных магистралей Арктики и активно используется для перевозки полезных ископаемых, добываемых российскими компаниями за полярным кругом, снабжения удаленных населенных пунктов Крайнего Севера, а также решения других важных задач.

Целью данной статьи является анализ возможностей использования интеллектуальной ГИС для мониторинга ледовой обстановки в Арктике и обеспечения безопасности судоходства по Северному морскому пути. Такая система способна обеспечить интеллектуальный анализ данных (в том числе, данных дистанционного зондирования Земли), визуализировать в удобном для пользователя виде результаты анализа, и, что более важно, осуществить поддержку принятия сложных решений, связанных с возникновением опасных ситуаций на Северном морском пути.

Основой построения системы мониторинга ледовой обстановки и обеспечения безопасного судоходства служит интеллектуальная геоинформационная система. Ядром этой системы является база знаний, включающая онтологию. Экспертная система в составе системы мониторинга применяется для разработки, наглядного представления и реализации сценариев обработки данных в целях распознавания и прогнозирования ледовой обстановки в регионе, а также интеллектуального сопровождения решения задач безопасного судоходства. Основными источниками информации системы мониторинга являются специальные системы наблюдения, в том числе и космические системы дистанционного зондирования Земли, использующие различные принципы сбора и первичной обработки данных.

Возможности разработанной системы мониторинга демонстрируются на примере корректировки маршрута движения судов по участку Северного морского пути с учетом ледовой обстановки в районе.

## SUMMARY

### *Tsvetkov M.V., Smirnova O.V., Galiano P.R.* **Ice situation and safe navigation monitoring system on the Northern Sea Route based on intelligent GIS.**

The Arctic is one of the most perspective territories in Russian Federation. In recent years there has been an intensive development of oil and gas fields in the Arctic, navigation in water areas of the Arctic seas, various problems connected with defense capability of the country. The Northern Sea Route is one of the largest Arctic sea traffic arteries. It is actively used for transportation of mineral resources extracted by Russian companies above the Arctic Circle, supply support of remote settlements in Far North, and also for solution of other important problems.

The issue of this paper is to analyze the possible application of intelligent geographic information system for ice situation monitoring in the Arctic and safe navigation on the Northern Sea Route. Such system may be able to analyze intelligent data (including remote sensing data), visualize results in a user-friendly style, and what is more important, and provide decision-making support for potentially hazardous conditions on the Northern Sea Route.

The intelligent GIS is the basis for construction of the ice situation and safe navigation monitoring system. The core of this system is knowledge base including ontology. Expert system is used for design, visualization and realization of data processing scenarios in accordance with ice situation recognition and prediction, and also for solving problems of safe navigation. The main information sources of the monitoring system are special surveillance systems including space system that is using different principles of the collecting and primary processing data.

Abilities of the developed monitoring system are shown in the example of the ship route adjustment on the part of the Northern Sea Route in accordance with the ice situation.

## РАЗВИТИЕ МЕТОДОЛОГИИ И ТЕХНОЛОГИИ ИМИТАЦИОННЫХ ИССЛЕДОВАНИЙ СЛОЖНЫХ СИСТЕМ

---

*Девятков В.В. Развитие методологии и технологии имитационных исследований сложных систем.*

**Аннотация.** В данной статье предлагается усовершенствованная методология и технология имитационных исследований сложных систем, как результат развития и совершенствования традиционной методологии. Основным отличием усовершенствованной методологии является последовательная автоматизация процесса исследования и интеграция всех программ в единый комплекс. Программные комплексы, создаваемые на основе данной методологии, позволили в среднем сократить в несколько раз время проведения исследования сложных систем и значительно увеличить круг потенциальных пользователей имитационного моделирования.

**Ключевые слова:** имитационное моделирование, методология имитационных исследований, концепция состояний и этапов, концепция программ.

*Devyatkov V.V. The Development of Methodology and Technology of Simulation Studies of Complex Systems.*

**Abstract.** In this paper we propose an improved methodology and technology of simulation studies of complex systems as a result of development and improvement of the traditional methodology. The main difference is the improved methodology consistent automation of the process of research and the integration of all programs in a single complex. Software systems that are created on the basis of this methodology allowed an average cut several times during the study of complex systems and significantly increase the number of potential users of simulation.

**Keywords:** simulation, methodology for simulation studies, the concept of states and stages, the concept of program.

---

**1. Введение.** Традиционная методология проведения процесса имитационного исследования сложных систем (далее ИИСС) представляет собой последовательную реализацию множества этапов, которые должен выполнить исследователь. Методология ИИСС детально описана классиками имитационного моделирования (далее ИМ) Н.П. Бусленко, Т. Нейлором и Р. Шенноном [1-3]. На рисунке 1 показана структура проведения имитационного исследования, предлагаемая Р. Шенноном, это последовательное выполнение 11 этапов. Современные исследователи, как правило, используют именно эти этапы [4, 5], меняются в основном только их названия. Например, первый этап чаще всего называется «Постановка задачи», второй этап - «Сбор данных» и т.д.

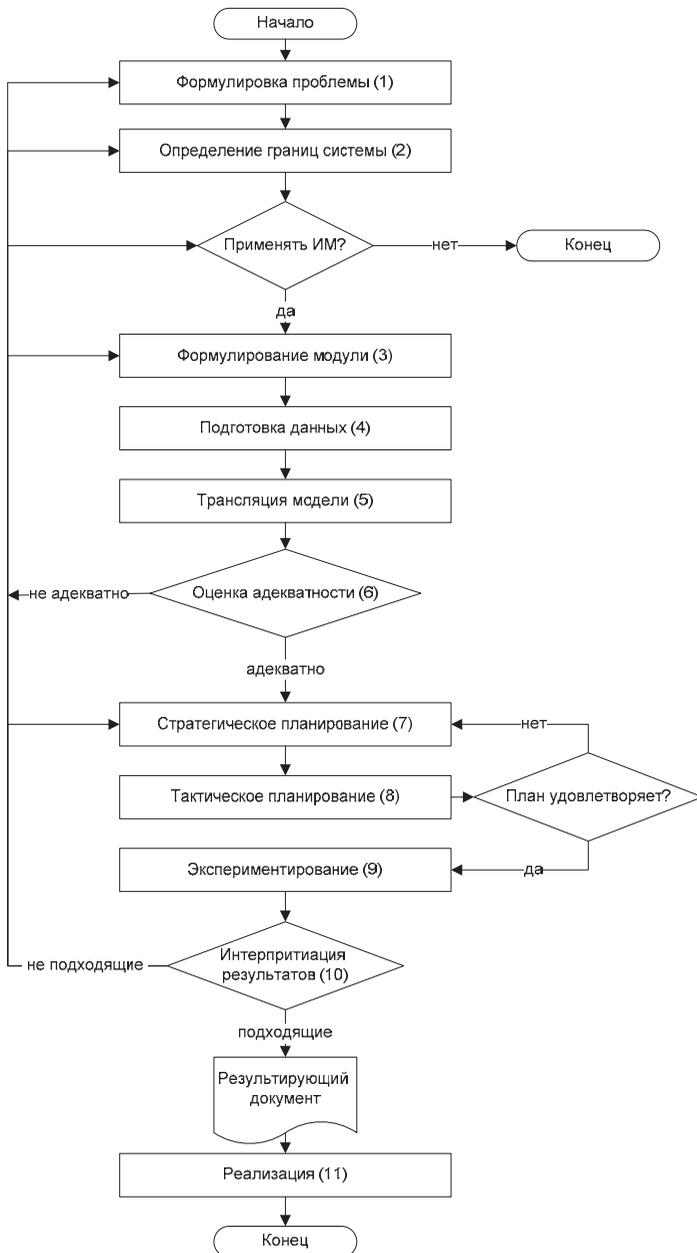


Рис. 1 Структура процесса проведения имитационного моделирования

На момент создания традиционной методологии, более 40 лет назад, большинство действий и операций выполнялись исследователем вручную. И только небольшая их часть (этапы 3, 4, 5 и 9) была частично автоматизирована в рамках первых языков ИМ. В связи с последующей объективной необходимостью автоматизации процесса ИИСС, появлением новых теоретических методов и алгоритмов сам процесс ИИСС становился все более сложной системой и требовал детального изучения и совершенствования [6, 7]. Наряду с очевидной структурой процесса ИИСС, показанной на рисунке 1, существует еще целый ряд элементов, связей и воздействий, которые не влияли на эффективность ручного проведения исследования, но стали определяющими факторами в условиях автоматизации процесса ИИСС. Эти факторы необходимо описать, понять их причинно-следственные связи с процессом ИИСС и найти способы повышения эффективности всего исследования в целом.

Представим традиционный процесс ИИСС ( $S$ ) формально в виде совокупности функционально зависящих и взаимодействующих друг с другом множеств – целей, этапов, состояний этапов, их взаимосвязей, ручных и автоматизированных управляющих воздействий и показателей функционирования моделируемой системы.

$$S = ((C_i), (S_i), (S_i^j), (Q_i), (Q_i^j), (U_{pi}^j), (U_{npi}^j), (П_i), (П_i^j)), \quad (1)$$

где,  $i=(1, \dots, n)$  – количество этапов процесса ИИСС,  $j=(1, \dots, m_i)$  количество состояний  $i$ -го этапа.

Все указанные в формуле (1) множества являются разнородными и сложно формализуемыми, но исключительно важными для представления и понимания сути исследования. Процесс ИИСС всегда имеет определенные цели и задачи этапов ( $C_i$ ), иначе проведение исследования бессмысленно. Структурно, любое имитационное исследование состоит из обязательных для исполнения этапов ( $S_i$ ), показанных на рисунке 1 и входящих в них более детальных действий и операций – состояний этапов ( $S_i^j$ ). Связь этапов и состояний этапов между собой отображена во множествах ( $Q_i$ ) и ( $Q_i^j$ ), которые при конкретной реализации описываются матрицами смежности или иерархическими структурами. Процесс ИИСС итеративен и является системой с обратной связью. Непосредственным участником процесса является исследователь, который управляет посредством ручных ( $U_{pi}^j$ ) или автоматизированных и автоматических ( $U_{npi}^j$ ) воздействий. Ручное воздействие – это вмешательство исследователя в процесс ИИСС без применения программных средств. Автоматизированное воздействие это

вмешательство исследователя в процесс ИИСС интерактивными средствами той или иной программной среды. Автоматическое воздействие - это вмешательство в процесс ИИСС без участия исследователя, возникающее в соответствии с заложенной разработчиком логикой работы программных средств. Для анализа качества и эффективности поведения исследуемой системы всегда формулируются показатели функционирования процесса ИИСС, его этапов и состояний - множества  $(P_i)$  и  $(P_i^j)$ . Характеристики указанных выше множеств могут изменяться от исследуемой системы, условий проведения исследования, знаний и предпочтений исследователя, его квалификации и ряда других факторов.

С момента возникновения и до настоящего времени, ввиду сложности применения, ИМ было доступно только профессионалам, в основном это были ученые. До сих пор актуально меткое определение Р. Шеннона, вынесенное в название его книги - «Имитационное моделирование систем - искусство и наука».

В связи с непрерывным усложнением окружающих нас систем практический интерес к имитационным исследованиям неуклонно возрастает [8, 9]. По нашим оценкам ежегодно, только в России, требуют исследования сотни тысяч сложных систем, и выполнить их только силами профессионалов их невозможно. У государства и бизнеса есть необходимость и желание воспользоваться ИМ, но нет практических инструментальных средств, чтобы сделать это быстро, просто и за разумные средства.

Имеющийся у автора опыт построения моделей многих сложных систем и проведения с их использованием исследований показывает, что необходимы, и самое главное возможны, изменения в организации и технологиях проведения процесса ИИСС. В результате должны быть созданы более совершенные инструментальные программные средства ИМ, с которыми мог бы работать более широкий круг исследователей. Например, любой специалист с инженерным уровнем образования.

**2. Эволюция и совершенствование традиционного процесса ИИСС.** Для проведения ИИСС в мире разработано и успешно используется большое количество самых разнообразных программных средств, которые можно использовать при автоматизации моделирования (рисунок 2). Например, это различные системы: для структурированного описания систем с использованием стандартов - IDF, UML и BPMN представлений; для статистической обработки исходных данных и результатов – ППП СТАТИСТИКА, система SAS, Stat::FIT, StatGraphics и т.д.; десятки языков имитационного

моделирования для создания модели и проведения экспериментов – GPSS World, AnyLogic, Extend, SIMIO и т.д.; ряд систем оптимизации параметров исследуемых систем - OptQuest, IOSO и другие программные средства реализации этапов ИИСС. Часть этих программ создана специально для моделирования, но большинство разрабатывалась для других применений и в процессе ИИСС используется только некоторые функции программы. Наличие этих программ большой плюс для исследования, но все эти программы необходимо применять в рамках единого процесса исследования, реализующего последовательно один этап за другим. Наиболее сложным и длительным для исследователя является передача, в соответствии с технологией, результатов выполнения одной программы на вход другой программы. Это требует введение множества ручных подготовительных операций.

В настоящее время одним из основных и определяющих факторов успешного проведения процесса ИИСС является выбор исследователем совокупности программных средств, реализующих цели и задачи исследования.

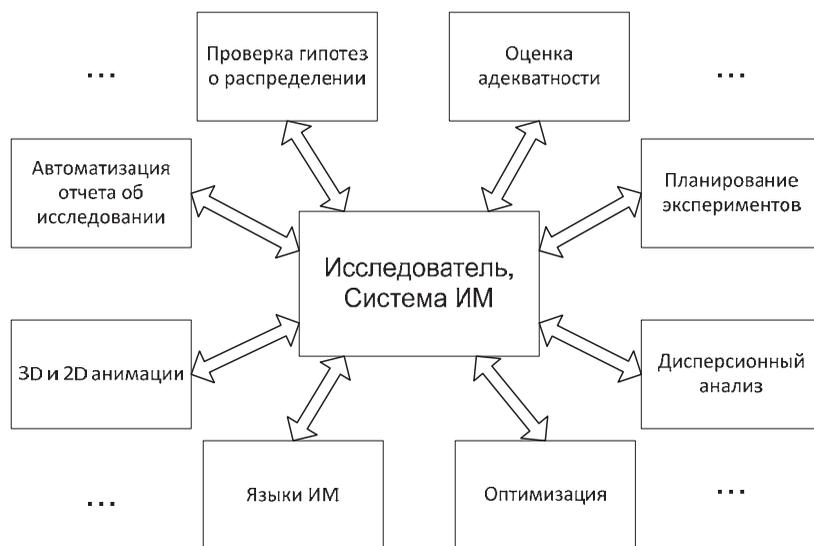


Рис. 2. Программы автоматизации действий исследователя при проведении ИИСС

Каждый язык моделирования или дополнительная программа, используемая в комплексном или автономном режиме, содержат свой набор программных компонент, связей, диалогов и формируемых в

модели показателей функционирования системы. Обозначим через  $Tech_s^i$ , совокупность множеств, представляющих конкретные программные технологии, которые используются для реализации состояний процесса ИИСС на  $i$ -ом этапе:  $Tech_s^i = ((S_i^j), (Q_i^j), (U_i^j), (U_{pi}^j), (Pi_i^j))$ . Тогда, с учетом оставшихся ручных операций (задание целей, формирование показателей, передача результатов одного из этапов в качестве исходных данных на другой этап), формула (1) примет следующий вид:

$$S = ((C_i), (S_i), (Q_i), (U_{pi}), (Tech_s^i), (Pi_i)). \quad (2)$$

Далее автором было проведено детальное исследование изменений, которые произошли в традиционном процессе проведения ИИСС с момента его возникновения и по настоящее время [10, 11].

В результате были выделены следующие основные направления и тенденции эволюции традиционного подхода к проведению ИИСС:

1. Идет последовательная автоматизация ручных операций и действий исследователя на всех этапах ИИСС посредством создания новых программ.

2. Непрерывно совершенствуются ранее созданные программные средства с целью улучшения их функциональных свойств, повышения скорости вычислений и упрощения интерфейса пользователя.

3. Происходят значительные структурные изменения процесса ИИСС – некоторые этапы исчезают, появляются новые этапы, совершенствуются или объединяются существующие этапы и т.д.

4. Современные программные технологии позволяют при построении очень сложных моделей применять методы и технологии коллективной работы над одним исследованием. Например, за ввод данных отвечают одна группа исследователей, за эксперименты – вторая, за анализ и интерпретацию результатов – третья.

5. Во многих программных системах активно используются современные методы распределения вычислений (например, облачные технологии) в процессе проведения ИИСС между компьютерами различной мощности. В частности, этап «Экспериментирование» переводится на высокопроизводительные сервера (и даже СУПЕР ЭВМ), а действия на всех этапах по диалоговому взаимодействию с пользователем – на персональные и мобильные устройства (планшеты, смартфоны и др.).

Несмотря на такие глубинные изменения, методологическая основа проведения имитационного исследования осталась прежней и базируется на традиционном подходе и концепции автономных

«Этапов и состояний». При этом не в полной мере учитываются возможности современных информационных технологий по комплексированию программ и данных. В результате при проведении ИИСС приходится искусственно водить множество ручных операций для «стыковки» этапов, отсутствует единая база данных моделей и результатов. При этом каждый исследователь часть операций по интеграции исследования выполняет самостоятельно.

Все это приводит к неэффективному проведению процесса ИИСС:

Во-первых, это длительные сроки исследования ( $T_{ин}$ ). Время исследования в процессе эволюции удалось в среднем снизить с шести до трех месяцев, тем не менее, оно остается недопустимо большим. Это не устраивает большинство потенциальных пользователей. Как показал анализ, значительные резервы для уменьшения  $T_{ин}$  есть. В основном  $T_{ин}$  увеличивается при выполнении ручных операций, связывающих соседние в процессе ИИСС этапы.

Во-вторых, сложность использования метода ИМ из-за завышенных требований к квалификации исследователя. Исследователь должен знать методологию исследования, обладать компетенцией в области статистики, планирования экспериментов, оптимизации и ряде других наук, он должен уметь использовать различные программы для проведения процессе ИИСС и т.д. В результате от создания модели и проведения исследования оторваны самые заинтересованные лица – руководители, системные аналитики и инженеры Заказчика.

Таким образом по результатам детального анализа был сделан вывод, что проводить исследование существующими программными средствами в рамках традиционного подхода к ИМ становится все сложнее и очевидный потенциал увеличения практических применений моделирования не может быть в полной мере реализован. Поэтому, наряду с полной автоматизацией и совершенствованием отдельных программ, необходимы концептуальные изменения при проведении ИИСС. При этом должны использоваться самые современные информационные технологии [12, 13]. Итоговым результатом должна стать более совершенная методология проведения ИИСС и возможность создания на ее основе удобной и простой программной технологии ИМ для массового применения.

### **3. Развитие традиционной методологии проведения ИИСС.**

В качестве основного направления совершенствования процесса ИИСС был выбран последовательный уход от ручных операций и объединение всех используемых программ в единый программный

комплекс. При объединении, наряду с уже имеющимся функциями программ, необходимо было предусмотреть возможность создания общих для всех этапов баз данных моделей, исходных данных и результатов. Кроме этого требовалось создать средства для проектирования простого и удобного языка взаимодействия исследователя с моделью, настраиваемого на предметную область каждой создаваемой модели.

В результате в традиционную методологию проведения процесса ИИСС были внесены существенные изменения [14], направленные на систематизацию и объединение всех эволюционных изменений в рамках единой концепции и создание программных спецификаций для массового создания имитационных приложений. Основным постулатом при разработке методологии был переход от концепции «Этапов и состояний» к концепции «Программ и технологий вычислений».

В результате процесс ИИСС был представлен в виде парадигмы, изображенной на рисунке 3.

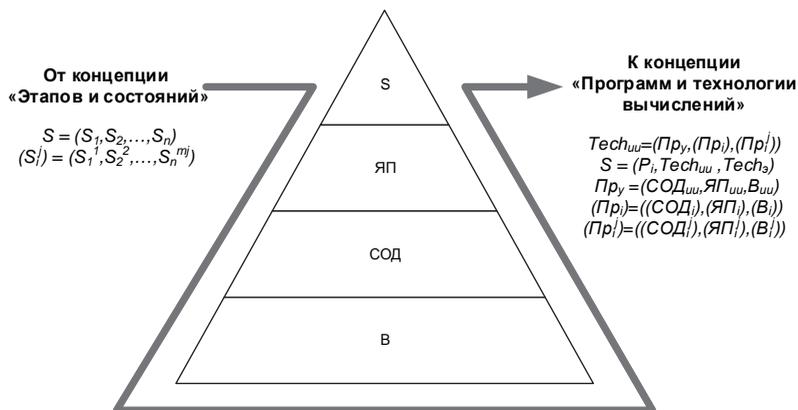


Рис. 3. Парадигма имитационного исследования

Парадигма состоит из четырех, взаимодополняющих абстрактных уровней:

- структурно-функциональный уровень – S;
- язык взаимодействия исследователя с программами – ЯП (язык пользователя);
- уровень данных – СОД (система обработки данных процесса ИИСС);
- архитектура организации и проведения вычислений – В.

Как показано на рисунке 3 в концепции «Программ и технологии вычислений» были введены формальные обозначения множеств программ, реализующих общее управление процессом ИИСС ( $Pr_y$ ), функции этапов ( $Pr_i$ ) и их состояний ( $Pr_i^j$ ).

В результате было сформулировано понятие общей программной технологии проведения исследования  $Tech_{ин}=(Pr_y,(Pr_i),(Pr_i^j))$  и выделено множество оставшихся ручных операций ( $P_i$ ).

В итоге процесс проведения ИИСС в рамках усовершенствованной методологии был представлен в следующем виде.

$$S = (P_i, Tech_{ин}, Tech_э), \quad (3)$$

где,  $i=(1, \dots, n)$  – новые видоизмененные этапы.

Введение программных технологий  $Tech_{ин}$  и  $Tech_э$  позволило осуществить переход от ряда неструктурированных, разнородных множеств и соответствующего им вербального описания - к программам, которые могут быть структурированы, алгоритмически описаны и технологически реализованы (формула 3). Для формального описания программ, структур данных имеются различные методы и средства теории программирования и вычислений. Поэтому, при создании конкретного инструментального средства ИМ в рамках усовершенствованной методологии, переход от концептуального описания к программным спецификациям является чисто технологическим процессом.

Основным связующим звеном процесса проведения и управления ИИСС при новом подходе является интегрированная программная технология.

**Определение.** Интегрированная программная технология  $Tech_{ин}$  это совокупность правил и принципов объединения программ, обеспечивающие исполнение следующих возможных действий и операций при проведении процесса ИИСС:

1) Формирование общей цели, задач и показателей для всего исследования и локальных задач и показателей для отдельных этапов.

2) Автоматизированный ввод и обработка управляющих воздействий исследователя (команд и данных) интегрирующих каждый отдельный этап в процесс ИИСС.

3) Централизованное накопление, сохранение, и обработка исходных данных и результатов, необходимых для реализации отдельных этапов и всего процесса ИИСС.

4) Оперативный интерактивный анализ экспериментов и серий экспериментов с моделью.

5) Обобщение результатов и выработка рекомендаций на каждом этапе и по итогам проведения всего процесса ИИСС.

6) Автоматизированное оформление отчета об исследовании.

Необходимо отметить, что *Tech<sub>ини</sub>* при реализации конкретного инструментального средства ИМ может включать не все указанные действия и операции. Это зависит от текущего уровня автоматизации и интеграции процесса ИИСС в используемой системе ИМ. Кроме выбранной программной технологии на характеристики процесса ИИСС влияют архитектура вычислений, уровень используемых информационных технологий и качество программной и методической реализации этих операций.

**4. Практическое применение усовершенствованной методологии.** Конкретную программную реализацию с использованием данной методологии будем называть системой автоматизации имитационных исследований (САИИ).

В приведенных ниже примерах создания САИИ при реализации в качестве базовой системы ИМ использовался язык GPSS World. Хотя в целом это непринципиально, так как методология не привязана ни к одной из существующих систем моделирования. Выбор GPSS World обусловлен предпочтениями автора, возможностями языка и практикой моделирования на нем по всему миру.

Сначала был создан САИИ для профессионалов - расширенный редактор GPSS World.

В самом общем виде основные преимущества реализованных в редакторе решений сформулируем следующим образом. Исследователь в процессе ИИСС может:

1. Осуществлять общее управление исследованием (моделями, экспериментами и результатами), используя специально созданную структуру данных, которая ему привычна, понятна и удобна. Это библиотека с иерархией «проект – модели – эксперименты – отчеты».

2. Начинать разработку модели с формулирования проблемы, посредством создания иерархической графической структурной схемы модели, которая конструируется исследователем из библиотеки ранее созданных типовых элементарных блоков.

3. Осуществлять автоматическую генерацию и сборку текста модели по созданной структурной схеме и введенным исходным данным.

4. Создавать в новом текстовом редакторе модель, используя автовыворачивание, подсветку синтаксиса, контекстные подсказки «на лету», дополнительный синтаксический анализ, группировку и т.д.

5. Более эффективно и быстро, чем стандартными средствами GPSS World, отлаживать модели при помощи мощнейшего интерактивного отладчика, позволяющего визуально отслеживать не только продвижение транзактов, но и изменение при этом состояния и значений других объектов в модели.

6. Проектировать формы ввода исходных данных и подготовки серий экспериментов на основе созданной модели. Каждая форма – это совокупность простых и удобных программных диалогов по заданию параметров моделируемой системы с использованием языка предметной области разрабатываемой модели.

7. Конструировать интерактивные формы анализа результатов моделирования. В результате профессионал получает мощнейший инструмент, позволяющий отобразить все необходимые показатели модели.

8. Выбрать факторы и показатели для данной модели, которые будут участвовать в серии экспериментов. Задать один из возможных методов планирования экспериментов и автоматически сформировать план экспериментов.

9. Выполнить одиночный эксперимент с моделью, либо провести серию экспериментов по сформированному плану.

10. Визуализировать ход выполнения любого эксперимента в динамике, в том числе и в анимированном графическом виде.

11. Осуществить сравнительный интерактивный анализ результатов экспериментов и серий экспериментов, сохранять полученные результаты в общем хранилище данных.

12. Автоматически формировать, дополнять вручную и распечатывать результирующий исследование документ (отчет), включающий любые полученные результаты моделирования.

13. Обеспечить удаленную (сетевую) работу с одной или несколькими копиями системы GPSS World.

Необходимо отметить, что новые возможности расширенного редактора, предоставляемые пользователю, никоим образом не изменяют спецификации языка моделирования. Тексты моделей на языке GPSS World, написанные ранее, будут полностью совместимы с новым редактором. Пример формы вывода результатов, созданной с использованием технологии массовой разработки приложений, приведен на рисунке 4.

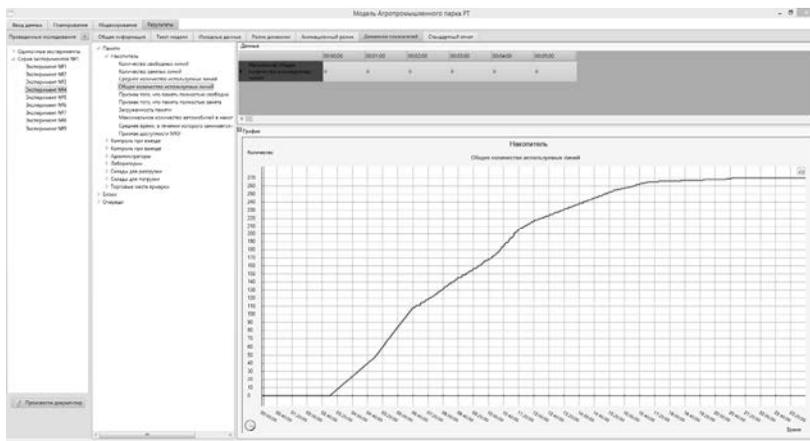


Рис. 4. Форма вывода динамических результатов в САИИ

Расширенный редактор стал базисом для создания технологии массового создания САИИ в различных предметных областях. Суть технологии заключается в создании предметно-ориентированных библиотек типовых элементарных блоков, базисных конструкций языка пользователя в виде шаблонов форм ввода данных и анализа результатов. В итоге исследователь создает имитационные модели, подобно детскому конструктору «лего» из имеющихся блоков.

**5. Заключение.** Усовершенствованная методология была апробирована автором при проведении целого ряда практических имитационных исследований: выбор наилучших транспортно-логистических решений при проектировании Агрпромышленного парка Республики Татарстан, оценка и анализ пропускной и провозной способности железнодорожных участков и направлений для ОАО РЖД, выработка практических рекомендаций при проектировании технологии судосборочных производств в сухих доках и т.д.

Одним из наших наиболее значимых примеров практического внедрения методологии и технологии массового создания САИИ стал проект исследования транспортной системы Универсиады 2013 года в Казани. За несколько месяцев было разработано более 60 моделей для исследования транспортной инфраструктуры различных спортивных объектов и движения клиентских маршрутов Универсиады по городу. По результатам имитационного исследования были предоставлены количественно рассчитанные и обоснованные рекомендации по транспортным схемам и управлению движением для ГИБДД и список особо загруженных участков дорог, въездов на объекты, парковки, на которых необходимо было особо тщательно вести постоянный

мониторинг и контроль над потоками транспорта Универсиады. Все рекомендации были учтены при проведении самой Универсиады. В итоге транспортная система во время Универсиады работала без сбоев, очередей и задержек. Необходимость, своевременность и качество проведенного нами системного анализа признано федерацией студенческого спорта (FISU), правительствами России и Татарстана.

Таким образом принципиальным преимуществом усовершенствованной методологии и новой программной технологии является возможность создание интегрированного и эффективного программного комплекса для реализации полного цикла процесса проведения ИИСС. Внедрение усовершенствованной технологии показало, что происходит сокращение времени проведения исследования, в среднем с трех месяцев до нескольких недель. При этом исследование по существу превращается в имитационную экспертизу, которая гарантированно дает количественные оценки и рекомендации по совершенствованию системы за допустимое для заказчика время.

### Литература

1. *Бусленко Н.П.* Моделирование сложных систем // М.: Наука. 1978. 400 с.
2. *Шеннон Р.* Имитационное моделирование систем – искусство и наука / пер. с английского // М.: Мир. 1978. 418 с.
3. *Нейлор Т.* Машинные имитационные эксперименты с моделями экономических систем / пер. с английского // М.: Мир. 1975. 500 с.
4. *Balci O.* A Life Cycle for Modeling and Simulation // *Simulation*. 2012. Vol. 88(7). pp. 870–883.
5. *Толуев Ю.И.* Стандартные этапы создания имитационных моделей производственных и логистических систем // Материалы шестой всероссийской научно-практической конференции «Имитационное моделирование. Теория и практика» ИММОД-2013. Казань: Изд-во ФЭН Академии наук РТ. 2013. Том 1. С. 80–89
6. *Tolk A., Turnitsa C.* Conceptual modeling with processes // *Proceedings of the Winter Simulation Conference. WSC*. 2012. pp. 2641–2653.
7. *Wainer G. A.* Discrete-Event Modeling and Simulation: A Practitioner's Approach // CRC Press. 2009. 520 p.
8. *Ross S. M.* Simulation (Fifth edition) // Academic Press. 2012. 328 p.
9. *Tolk A., Diallo S.Y., Padilla J.J.* Semiotics, Entropy, and Interoperability of Simulation Systems – Mathematical Foundations of M&S Standardization // *Proceedings of the Winter Simulation Conference. WSC*. 2012. pp. 2751–2762.
10. *Heath B.L., Ciarallo Hill R.R.* Validation in the Agent-based Modeling Paradigm: Problems and a Solution // *Simulation and Process Modeling*. 2012. 7 (4). pp. 229–239.
11. *Bauer K., Schneider K.* Predicting Events for the Simulation of Hybrid Systems // *Proceedings of the 10th IEEE International Conference on Computer and Information (CIT 2010)*. pp. 1833–1840.
12. *Guizzardi G., Wagner G.* Towards an Ontological Foundation of Discrete Event Simulation // *Proceedings of the Winter Simulation Conference. WSC*. 2010. pp. 652–664.
13. *Furht B., Escalante A.* Handbook of Cloud Computing // Springer. 2010. 634 p.

14. *Десятков В.В. Методология и технология имитационных исследований сложных систем: современное состояние и перспективы развития: монография // М.: Вузовский учебник.: ИНФРА-М. 2013. 448 с.*

## References

1. Buslenko N.P. *Modelirovanie slozhnyh sistem* [Simulation complex systems]. М.: Nauka. 1978. 400 p. (In Russ.).
2. Shannon R. *Imitacionnoe modelirovanie sistem – iskusstvo i nauka* [System simulation – art and science]. М.: Mir. 1978. 418 p. (In Russ.).
3. Naylor T. *Mashinnye imitacionnye jeksperimenty s modeljami jekonomicheskikh sistem* [Computer simulation experiments with models of economic systems]. М.: Mir. 1975. 500 p. (In Russ.).
4. Balci O. A Life Cycle for Modeling and Simulation. *Simulation*. 2012. vol. 88(7). pp. 870–883.
5. Toluev U.I. [Standard steps for creating simulation models of production and logistic systems]. *Materialy shestoj vserosijskoj nauchno-prakticheskoy konferencii «Imitacionnoe modelirovanie. Teorija i praktika» IMMOD-2013* [Proceedings of the Sixth Russian scientific and practical conference «Simulation. Theory and Practice IMMOD 2013»]. Kazan: FEN Tatarstan Academy of Sciences. 2013. vol. 1. pp. 80–89. (In Russ.).
6. Tolk A., Turnitsa C. Conceptual modeling with processes // *Proceedings of the Winter Simulation Conference. WSC. 2012. pp. 2641–2653.*
7. Wainer G. A. *Discrete-Event Modeling and Simulation: A Practitioner's Approach*. CRC Press. 2009. 520 p.
8. Ross S. M. *Simulation* (Fifth edition). Academic Press. 2012. 328 p.
9. Tolk A., Diallo S.Y., Padilla J.J. Semiotics, Entropy, and Interoperability of Simulation Systems – Mathematical Foundations of M&S Standardization. *Proceedings of the Winter Simulation Conference. WSC. 2012. pp. 2751–2762.*
10. Heath B.L., Ciarallo F.W., Hill R.R. Validation in the Agent-based Modeling Paradigm: Problems and a Solution. *Simulation and Process Modeling*. 2012. vol. 7(4). pp. 229–239.
11. Bauer K., Schneider K. Predicting Events for the Simulation of Hybrid Systems. *Proceedings of the 10th IEEE International Conference on Computer and Information (CIT 2010)*. pp. 1833–1840.
12. Guizzardi G., Wagner G. Towards an Ontological Foundation of Discrete Event Simulation. *Proceedings of the Winter Simulation Conference. WSC. 2010. pp. 652–664.*
13. Furht B., Escalante A. *Handbook of Cloud Computing*. Springer. 2010. 634 p.
14. Devyatkov V.V. *Metodologija i tehnologija imitacionnyh issledovanij slozhnyh sistem: sovremennoe sostojanie i perspektivy razvitija: monografija* [Methodology and technology of simulation studies of economic systems: current state and prospects: monograph]. М.: Vuzovsky uchebnic.: INFRA-M. 2013. 448 p. (In Russ.).

**Десятков Владимир Васильевич** — к-т техн. наук, заместитель директора по науке, Институт информатики Академии наук Республики Татарстан. Область научных интересов: информатика, системный анализ, имитационное моделирование. Число научных публикаций — 73. vladimir@elina-computer.ru; 420111 Казань, ул. Левобулачная, д. 36а; р.т. +7 843 292-38-67, факс +7 843 292-39-05.

**Devyatkov Vladimir Vasil'evich** — Ph.D., Deputy director for science, Institute of Computer Science Tatarstan Academy of Science. Scientific interests: informatics, system analysis, simulation modeling. The number of publications — 73. vladimir@elina-computer.ru; Levobulachnay str., 36a, Kazan, Russian Federation 420015; office phone +7 843 292-38-67, fax +7 843 292-39-05.

## РЕФЕРАТ

### *Десятков В.В.* **Развитие методологии и технологии имитационных исследований сложных систем.**

В статье рассматриваются вопросы методологического обеспечения имитационных исследований сложных систем. Анализируется традиционный подход к имитационным исследованиям. Выделен основной недостаток традиционного подхода – ориентация на устаревшие программные технологии. Это сдерживает создание инструментальных средств имитационных исследований для массового применения.

Показано, что требуется совершенствование традиционной методологии с учетом возможностей современных информационных технологий. В связи с этим введена новая парадигма, описывающая процесс имитационных исследований и состоящая из четырех абстрактных уровней (структуры, языка пользователя, структуры данных и организации вычислений).

Это позволило сформулировать концепции и принципы новой методологии имитационных исследований, а также разработать концептуальные модели систем автоматизации имитационных исследований для различных применений.

В статье приведены примеры реализации имитационного приложения, разработанного с использованием новой методологии.

## SUMMARY

### *Devyatkov V.V.* **The development of methodology and technology of simulation studies of complex systems.**

This article discusses the methodological support of simulation studies of complex systems. Analyzes the traditional approach to simulation studies. Highlighted the main drawback of the traditional approach - focus on the latest software technology. It retards the development of simulation tools for the study of mass application.

It is shown that is necessary to improve the traditional methodology, taking into account the capabilities of modern information technology. In this regard, a new paradigm was introduced describing the process of simulation studies and consists of four abstract levels (structure, the user's language, data structures and organization of computing).

It is possible to formulate the concepts and principles of a new methodology for simulation studies, as well as to create a conceptual model of automation systems simulation studies for various applications.

The article gives examples of simulation applications developed using the new methodology.

## РАЗРАБОТКА И ИССЛЕДОВАНИЕ ПРЕДМЕТНО НЕЗАВИСИМОГО КЛАССИФИКАТОРА ТЕКСТОВ ПО ТОНАЛЬНОСТИ

---

*Рубцова Ю.В. Разработка и исследование предметно независимого классификатора текстов по тональности.*

**Аннотация.** В статье представляется метод построения классификатора для классификации текстов по тональности на два и на три класса (положительные и негативные; положительные, нейтральные и негативные тексты). Представляются результаты экспериментов, показывающие высокую точность работы метода при классификации на два класса независимо от предметной области к которой принадлежит текст. Эффективность представленного метода подтверждается экспериментами на текстовой коллекции блогов с разметкой по оценочной тональности семинара РОМИП-2012. Для оценки используются метрики: precision, recall, accuracy и F-меры. Значение F-меры для предлагаемого метода при классификации на 2 класса составляет 93%. Помимо блоговой коллекции РОМИП-2012, используются коллекция новостей и коллекция текстов социальных сетей.

**Ключевые слова:** анализ тональности текстов, машинное обучение, классификация текстов, автоматическая классификация, извлечение классификационных признаков.

*Rubtsova Y.V. Research and Development of Domain Independent Sentiment Classifier.*

**Abstract.** The paper presents the method of constructing a sentiment classifier on two and three classes (positive and negative, positive, neutral and negative texts). It is also presents the results of experiments. Results show the high accuracy of the proposed method on texts, which are not belong to any pre specified domains. The effectiveness of the presented method is confirmed by experiments' results on the text collection of blogs from ROMIP 2012 seminar. It was used following metrics for classifier evaluation: precision, recall, accuracy and F-measure. The value of F-measure of the proposed method for classification into 2 classes is up to 93%. In addition to blog collection ROMIP 2012 for experiments were used a collection of news and a collection of short-texts from social networks.

**Keywords:** sentiment analysis, machine learning, text classification and categorization, feature extraction.

---

**1. Введение.** Практически каждый второй россиянин старше 18 лет является активным пользователем Интернета, а более 80% из них имеют хотя бы один аккаунт в социальных сетях. Социальные сети становятся объектом пристального внимания социологов, психологов, маркетологов и PR-специалистов. Огромный объем информации, появляющийся в социальных сетях, приходится регулярно обрабатывать и классифицировать для того, чтобы иметь возможность решать поставленные перед специалистами задачи. Одна из таких задач – это поиск отзывов и упоминаний и классификация найденных текстов по тональности. Практическая ценность анализа тональности включает, но не ограничивается следующими примерами:

– при запуске нового продукта, компании могут быстро узнать, как покупатели оценивают этот продукт, нужно ли что-то исправить в продукте или рекламных материалах;

– правительство сможет отследить реакцию населения на новый закон, уточнения, заявление;

– организаторы мероприятий (напр. конференций) могут собрать отзывы участников в социальных сетях и оценить как прошло мероприятие, понравилось оно или нет;

– IT компании могут разработать эффективную систему поддержки пользователей, учитывая вопросы и отзывы пользователей;

– проведение маркетинговых исследований: изучение потребительских предпочтений, измерение степени удовлетворения потребностей потребителей, определение эффективности распространения продуктов или услуг;

– финансовые рынки. В работе [1] говорится, что существует множество новостей, статей в блогах и сообщений в твиттере о каждом акционерном обществе. Система автоматического анализа тональности может использовать эти источники для извлечения отзывов, что может стать основой для автоматической торговой системы.

В задаче классификации текстов по тональности можно выделить несколько подзадач:

– Классификация на 2 класса: положительный и отрицательный;

– Классификация на 3 класса: положительный, нейтральный и отрицательный;

– Классификация на 5 классов;

– Классификация на более, чем 5 классов, например, классификация по 10-бальной шкале.

Задачу классификации отзывов на два класса достаточно успешно решают как с помощью словарей и правил, так и с помощью машинного обучения. При классификации узкотематических текстов на два класса, точность классификаторов, основанных на униграммах, превышает 82% [2]. При решении задачи классификации более, чем на два класса, точность классификации существенно снижается – это связано с субъективным восприятием информации: то, что один человек считает «позитивным», другой может отнести к «нейтральному» или даже «склонному к негативному». В статье [3] авторы показывают, что при разделении текстов на большое количество классов, даже человек показывает низкую точность классификации – точность снижается до 55%.

В данной работе предлагается новый метод автоматического извлечения и взвешивания признаков на основе размеченной коллекции коротких сообщений платформы Твиттер. Коллекция была подготовлена

автором статьи для решения задачи тренировки классификатора текстов по тональности на 2 и на 3 класса. Особое внимание в статье уделяется описанию коллекции, из которой извлекаются списки слов – униграммы. Для каждого слова проводится морфологический анализ и рассчитывается набор статистических характеристик.

## **2. Подготовка коллекций для извлечения униграмм.**

Машинное обучение с учителем показывает наиболее точные результаты при классификации текстов по сравнению с другими методами, такими как машинное обучение без учителя и машинное обучение, основанное на правилах и словарях. Для обучения классификатора в качестве вектора признаков, на которые раскладывается документ, используют униграммы, биграммы или  $n$ -граммы. В задачах обработки текста на естественном языке популярно [2-5] представление документов в виде  $N$ -грамм, где  $N$ -граммы — последовательности слов или символов длины  $n$ . Для  $n=1$  такая последовательность состоит из одного слова и называется униграммой, для  $n=2$  такая последовательность называется биграммой, и т.д. В большинстве предыдущих работ [2-5] для извлечения  $N$ -грамм использовались тренировочные коллекции текстов, принадлежащих к определенной предметной области: рецензии на фильмы; отзывы на товары; коллекция новостей, собранная с одного сайта. В этих случаях, на тестовых коллекциях, классификаторы показывали высокую точность как для моделей, основанных на униграммах [2], так и на  $N$ -граммах [6]. Однако, подобные классификаторы не пригодны для классификации текстов по тональности, принадлежащих к другим предметным областям. Качественные характеристики оценки одной предметной области могут не выражать тонального отношения или вовсе не встречаться в другой предметной области. Так, например, негативная характеристика «скучный» (сюжет) выражающая отношение к фильму, может не встречаться в предметной области «цифровые фотоаппараты». Поэтому, несмотря на наличие размеченных по тональности узкотематических коллекций текстов на русском языке, они не подходят для построения универсального словаря униграмм для задачи классификации текстов по тональности, который будет показывать удовлетворительные результаты независимо от предметной области.

В статье выдвигается и успешно подтверждается гипотеза, что для извлечения  $n$ -грамм можно использовать достаточно представительный, однородный, сбалансированный корпус коротких сообщений, собранных на основе платформы Твиттер. Термин «достаточно представительный корпус» означает, что добавление новых сообщений к коллекции повлечет за собой добавление сравнительно небольшого числа новых терминов – это позволит

извлечь достаточно полный набор -грамм для классификации текстов, принадлежащих разным предметным областям.

В качестве платформы для сбора корпуса был выбран Твиттер потому что:

1. Во-первых, пользователи твитера часто выражают субъективное и эмоционально окрашенное мнение о чем-либо;

2. Для выражения эмоций, пользователи используют живой, разговорный язык, который может содержать сленг и ненормативную лексику, усиливающие тональность сообщений;

3. При написании сообщений, пользователи могут допускать широко распространённые ошибки, которые исправляются редакторами новостных изданий, но которые нужно учитывать при классификации по тональности текстов из Интернета (например, блогов или сайтов отзывов на товары).

С помощью Streaming API twitter [7] была собрана коллекция текстов, состоящая из около 15 миллионов коротких сообщений, на основе которой с помощью метода [8] и предложенной автором фильтрации [9] был сформирован сбалансированный корпус, состоящий из следующих коллекций:

- коллекция положительных сообщений 114 991 записей;
- коллекция негативных сообщений 111 923 записей;
- коллекция нейтральных сообщений 107 990 записей.

Чтобы удостовериться в однородности исследуемых коллекций воспользуемся законом Ципфа [10], который описывает распределение частот слов в естественном языке, если все слова достаточно длинного текста упорядочить по убыванию частоты их использования, то частота  $n$ -го слова в таком списке окажется приблизительно обратно пропорциональной его порядковому номеру  $n$ . Например второе по используемости слово встречается примерно в два раза реже, чем первое, третье — в три раза реже, чем первое, и т. д. Проверим предположение, что высокая частота встречаемости слов, таких как “и”, “я”, “в” в разных коллекциях (позитивной и негативной) должна быть сопоставимо одинакова. В то же время более редкие слова вроде “ассоциация”, “фактор” должны встречаться сопоставимо реже независимо от принадлежности коллекции к положительному или негативному классу [8]. На рисунке 1 для каждой из коллекции (позитивной и негативной) показано распределение нормированной частоты встречаемости слов в зависимости от количества твитов, в котором это слово встречается. Количество твитов, в которых встречается слово, отложено на оси  $x$ . Видно, что графики близки друг к другу на однородных наборах данных.

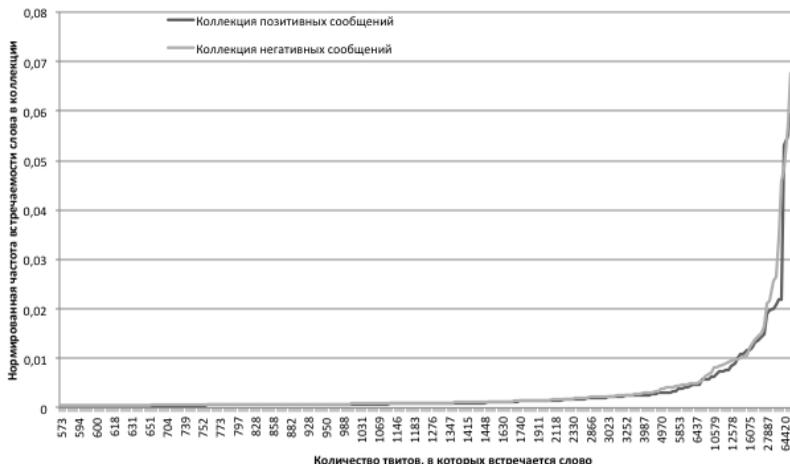


Рис 1. График распределения плотности слов в зависимости от количества твитов, в котором слово встречается

Для того чтобы удостовериться в полноте корпуса объединяем все три коллекции в одну, после чего производим вычисление количества уникальных терминов в зависимости от размера коллекции. На рисунке 2 показано, что при небольшом количестве твитов, добавление к коллекции новых сообщений влечет за собой увеличение числа уникальных терминов. Но, после достижения 340 000 уникальных терминов, добавление новых твитов к коллекции не влечет за собой значительного увеличения уникальных терминов.

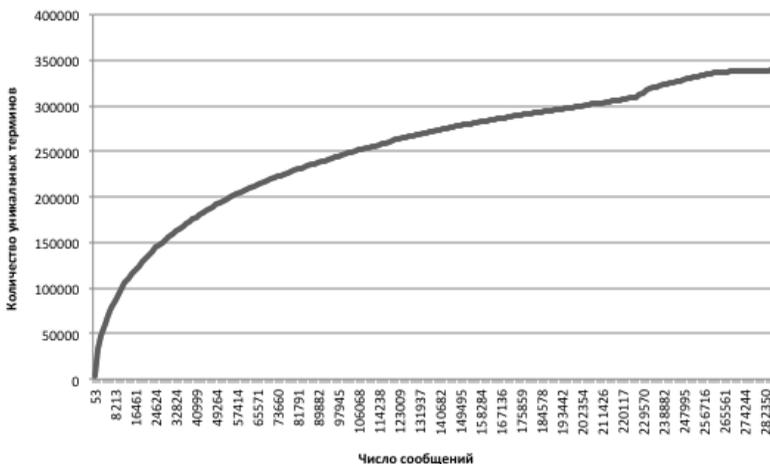


Рис 2. График числа уникальных терминов в зависимости от числа текстов в коллекции

Количество уникальных словоформ в коллекциях распределено следующим образом:

- Положительная коллекция – 150 720 уникальных словоформ;
- Нейтральная коллекция – 105 239 уникальных словоформ;
- Негативная коллекция – 191 677 уникальных словоформ.

Следующим шагом работы является выделение признаков для сокращения размерности векторного пространства и построение взвешенного словаря униграмм на основе размеченных коллекций.

**3. Выделение признаков для задачи текстовой классификации по тональности.** В формальном виде задача текстовой классификации представляется следующим образом [11].

Пусть существует описание документа  $d \in X$ , где  $X$  — векторное пространство документов, и фиксированный набор классов  $C = \{c_1, c_2, \dots, c_m\}$ . Из обучающей выборки (множества документов с заранее известными классами)  $D = \{ \langle d, c \rangle \mid \langle d, c \rangle \in X \times C \}$  с помощью метода обучения  $F$  необходимо получить классифицирующую функцию  $F(D) = \gamma$ , которая отображает документы в классы  $\gamma : X \rightarrow C$ . В решаемой задаче определения тональности множество  $C$  состоит из двух или трех элементов (положительный, отрицательный или положительный, нейтральный, отрицательный).

Все документы из обучающей и тестовой выборки представляют собой  $k$ -мерные векторы признаков. Таким образом документ определяется в виде вектора  $d = (w_1, w_2, \dots, w_{|V|})$ , где  $V$  — множество всех уникальных униграмм из обучающей выборки,  $w_i$  — вес  $i$ -й униграммы.

Для взвешивания униграмм в данной работе используются следующие пять весовых схем:

1. TF-IDF, которая вычисляется по формуле 1:

$$tf \cdot idf = tf \times \log \frac{T}{T(t_i)}, \quad (1)$$

здесь и далее,  $tf$  — это частота встречаемости термина в коллекции (положительных или отрицательных твитов).  $T$  — общее число сообщений в коллекциях положительных и отрицательных, а  $T(t_i)$  — число сообщений в положительной и отрицательной коллекциях, содержащих термин.

2. TF-RF, которая рассчитывается по следующей формуле:

$$tf \cdot rf = tf \times \log \left( 2 + \frac{a}{\max(1, c)} \right), \quad (2)$$

где  $a$  – количество сообщений (положительной) коллекции, содержащие термин,  $c$  – количество сообщений (отрицательной) коллекции, содержащие взвешиваемый термин. В работе [12] показано, что методы, основанные на мере RF, показывают лучшие результаты при вычислении веса слова с учетом принадлежности слова к разным классам, чем методы, основанные на мере TF-IDF.

3. Prob-Based [18], которая вычисляется по формуле 3:

$$prob - based = tf \times \log\left(1 + \frac{a}{c} \times \frac{a}{b}\right), \quad (3)$$

где  $a$  и  $c$  аналогично формуле 2,  $b$  – число сообщений (положительной) коллекции, которые не содержат взвешиваемый термин.

4. Формула 4 предназначена для вычисления TF-ICF [19]:

$$tf.icf = tf \times \log\left(1 + \frac{|C|}{cf(t_i)}\right), \quad (4)$$

где  $C$  – это число категорий,  $cf$  – число категорий, в которых встречается взвешиваемый термин.

5. ICF-Based [19] вычисляется по формуле 5:

$$icf - based = tf \times \log\left(2 + \frac{a}{\max(1, c)} \times \frac{|C|}{cf(t_i)}\right), \quad (5)$$

где  $C$  и  $cf$  аналогично формуле 4,  $a$  и  $c$  аналогично формуле 2.

В сообщениях, полученных из Твиттера, часто присутствует «шум» (очень редкие или наоборот очень частые слова, которые встречаются одинаково часто во всех классах), поэтому, перед взвешиванием униграмм, для уменьшения размерности вектора признаков и улучшения качества и скорости работы классификатора, произведена предобработка коллекций:

- Предлоги были отфильтрованы;
- Удалены спецсимволы (стикеры, изображение эмоций), которые пользователи используют при публикации записей с мобильных телефонов;
- Так же были отфильтрованы знаки препинания, такие как: запятая, точка с запятой, двоеточие, тире, точка. Восклицательные и вопросительные знаки были оставлены;
- Удалены имена собственные;
- Удалены значимые события (например, олимпиада);
- Все гиперссылки на другие документы были заменены паттерном «Link»;

- Все заглавные буквы приведены к строчным;
- Большое количество точек (более трех подряд идущих точек) были заменены паттерном «...».

В результате, получен словарь, который состоит из 17 143 взвешенных пятью способами униграмм.

С целью сокращения размерности вектора признаков, был проведен морфологический анализ слов коллекции и составлен второй словарь. Морфологический анализ был произведен с помощью TreeTagger для русского языка [13]. TreeTagger – это вероятностный инструмент для разметки текстов, разрешающий морфосинтаксические неоднозначности русского языка. В качестве униграмм во втором словаре выступают части речи. Как и в случае первого словаря, все униграммы были взвешены согласно пяти весовым схемам. Количество униграмм в морфологическом словаре всего 810. Чтобы не возникало путаницы, использование термина «словарь униграмм» далее по тексту относится к словарю униграмм построенному на извлеченных словоформах, использование термина «словарь морфологических униграмм» – к словарю униграмм построенному на информации о частях речи. На рисунке 3 схематично представлен процесс создания словарей униграмм.



Рис 3. Схема создания словарей униграмм

**4. Описание тестовых коллекций.** Автоматическая классификация отзывов на два и три класса проводилась на трех коллекциях.

**4.1. Коллекция коротких сообщений.** Коллекция подробно описана в разделе 2.

**4.2. Коллекция новостей.** Новостные коллекции были собраны на новостных web-сайтах. Разметка корпуса на положительные, нейтральные и негативные коллекции производилась ассессорами вручную. Отличие новостной коллекции от коллекции коротких текстов в том, что новости менее эмоциональны, лексика новостей более нейтральна и не изобилует жаргонизмами, сокращениями и ненормативной лексикой. Как правило, новостные тексты не содержат орфографических ошибок и символов, обозначающих эмоции на письме (смайликов) – тексты более строгие. Новостные тексты существенно длиннее 140 символов. Корпус новостных текстов состоит из следующих коллекций:

– коллекция положительных документов состоит из 46 339 новостей;

– коллекция негативных документов состоит из 46 337 новостей;

– коллекция нейтральных документов состоит из 46 340 новости.

**4.3. Коллекция текстов из блогов с разметкой по оценочной тональности и объектам ROMP 2012.** Число текстов коллекции 874. Предлагаемый набор текстов из блогов, участвовал в тестировании дорожек в рамках РОМИП-2011 [14]. В случае классификации на три класса имеем 534 положительных, 236 нейтральных и 103 отрицательных сообщений (одно сообщение исключено из-за ошибок кодировки). В случае классификации на два класса: 749 положительных текстов и 124 негативных сообщений. Разметка коллекции на два и три класса производилась ассессорами вручную. В коллекции содержатся тексты блогов пользователей с отзывами на цифровые фотоаппараты, рецензии на книги и на фильмы. Так же, как и в коллекции новостных текстов, длина текстов из блогов не ограничена. Следует отметить сильный дисбаланс тестовой коллекции в сторону положительных текстов.

Три выбранные для эксперимента коллекции принадлежат разным предметным областям, они достаточно разнообразны по содержанию и стилистике, имеют разную длину сообщения. Причем, если в коллекции коротких текстов есть лимит по длине, то в других коллекциях, текст может быть как очень длинным, так и сопоставимым по длине с текстом коротких Твиттер сообщений.

**5. Алгоритм классификации.** Согласно исследованиям [15, 16], наилучшие результаты в решении задачи автоматической текстовой классификации в целом и отзывов в частности, демонстрирует метод опорных векторов (Support Vector Machines – SVM). Так как вектор признаков имеет достаточно большую размерность, поэтому, в данной работе, для реализации алгоритма SVM была использована библиотека LIBLINEAR [17]. Библиотека LIBLINEAR – это реализация алгоритма SVM с линейным ядром. Все параметры алгоритма были оставлены в соответствии со значениями по умолчанию.

**6. Оценка качества.** В качестве оценки качества результатов обучения и работы классификатора выбраны четыре общепринятые характеристики: *assurasy*, *precision*, *recall* и *F-мера* [20].

По результату работы обученного классификатора на тестовой выборке для каждого класса вычисляются следующие значения:

- *TP* – количество истинно-положительных результатов.
- *TN* – количество истинно-отрицательных результатов.
- *FP* – количество ложно-положительных результатов.
- *FN* – количество ложно-отрицательных результатов.

*Precision* (точность) – это доля объектов классифицированы как *X*, которые действительно принадлежат классу *X* (формула 6).

$$Precision = \frac{TP}{TP + FP}, \quad (6)$$

*Recall* (полнота) – это доля всех объектов класса *X*, что классифицируется по алгоритму как принадлежащие классу *X*, вычисляется по формуле 7.

$$Recall = \frac{TP}{TP + FN}, \quad (7)$$

*F-мера* – это гармоническое среднее между *Precision* и *Recall*, формула 8.

$$F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall}, \quad (8)$$

*Assurasy* (аккуратность) – это доля правильно классифицированных объектов среди всех объектов, обработанных с помощью алгоритма классификации (формула 9):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \quad (9)$$

Для того, чтобы избежать проблемы переобучения, когда модель слишком хорошо работает на тестовых примерах, но плохо на реальных данных, для проверки модели используется метод перекрестной проверки (cross-validation). Вся обучающая выборка делится на  $k$ -частей, затем  $\forall i \in \{1, 2, \dots, k\}$  алгоритм классификации обучается на всей обучающей выборке, кроме части  $i$ , после чего, тестируется на  $i$ -й части. Результатом работы считается среднее арифметическое по всем проходам. Для данной работы значение  $k$  принимаем равным 5.

**7. Эксперименты и результаты.** Задача работы состоит в исследовании возможности и оценке качества применения словаря униграмм, извлеченных из эмоционально-окрашенных текстов платформы Твиттер для классификации текстов по тональности, извлеченных их других площадок и принадлежащих к заранее не известным предметным областям. С целью сравнения точности классификатора и выбора лучшего подхода, в зависимости от выбранного метода определения веса термина в коллекции для извлеченного набора униграмм, было проведено несколько экспериментов на трех разных по структуре наборах данных.

**7.1. Корпус коротких текстов.** Первый эксперимент был проведен для корпуса коротких текстов. Результаты работы классификатора на двух классах: положительный и отрицательный, представлены в таблице 1.

Таблица 1. Результаты классификации коротких текстов на два класса в зависимости от выбранной весовой схемы

	TF-IDF	TF-ICF	TF-RF	ICF-Based	Prob-Based
Accuracy	97.3805%	96.5346%	98.4741%	98.4953%	98.3026%
Precision	97.3402%	97.2926%	98.0017%	98.0157%	97.3489%
Recall	97.4928%	95.8256%	99.0067%	99.0345%	99.3549%
F-Measure	0.9742	0.9655	0.9850	0.9852	0.9834

Результат работы классификатора на трех классах: положительный, нейтральный, отрицательный, представлен в таблице 2. Меры TF-RF, ICF-Based и Prob-based предназначены для бинарной классификации, поэтому они не участвуют в классификации на три класса. Лучший результат выделен курсивом.

Таблица 2. Результаты классификации коротких текстов на три класса в зависимости от выбранной весовой схемы

	TF-IDF	TF-ICF
Accuracy	95.5981%	95.06645%
Precision	95.8093%	95.3112%
Recall	95.5205%	94.9847%
F-Measure	0.9566	0.9515

Так как эксперимент ставился на тех же коллекциях, которые использовались для извлечения униграмм, был получен ожидаемо хороший результат как при классификации на два класса, так и при классификации на три класса.

**7.2 Корпус новостей.** Следующий эксперимент был проведен для коллекции новостей, размеченных в ручную на три класса: положительные, нейтральные и отрицательные. Для представления текстов в векторном виде использовался вышеописанный словарь униграмм с пятью весовыми схемами. Работа классификатора, как и в первом эксперименте, проводилась перекрестной проверкой с шагом 5. Результаты работы классификатора на двух классах: положительный и отрицательный, представлены в таблице 3. Результат работы классификатора на трех классах: положительный, нейтральный, отрицательный, представлен в таблице 4.

Таблица 3. Результаты классификации текстов новостей на два класса в зависимости от выбранной весовой схемы

	TF-IDF	TF-ICF	TF-RF	ICF-Based	Prob-Based
Accuracy	89.4892%	79.2082%	95.0095%	95.0084%	91.1153%
Precision	92.3608%	80.0591%	94.5777%	94.589%	89.7698%
Recall	86.1003%	77.7941%	95.4941%	95.479%	92.8074%
F-Measure	0.8912	0.7891	0.9503	0.9503	0.9126

Таблица 4. Результаты классификации текстов новостей на три класса в зависимости от выбранной весовой схемы

	TF-IDF	TF-ICF
Accuracy	69.8619%	58.1397%
Precision	70.9246%	61.2780%
Recall	69.8624%	58.1403%
F-Measure	0.7039	0.5967

Эксперимент показал хороший результат при классификации новостей на два класса с использованием схемы взвешивания униграмм TF-RF. Более того, алгоритм достаточно точно классифицировал новостные тексты на три класса, показатель F-measure превышает 0.7 при взвешивании униграмм с помощью схемы TF-IDF. Следовательно, можно сделать вывод, что использование взвешенного словаря униграмм, выделенного из корпуса коротких эмоциональных текстов можно использовать для классификации текстов новостей по тональности.

**7.3. Коллекция текстов из блогов ROMIP 2012.** Третий эксперимент был поставлен на коллекции текстов из блогов. Результаты работы классификатора на двух классах: положительный и отрицательный, представлены в таблице 5.

Таблица 5. Результаты классификации текстов блогов на два класса в зависимости от выбранной весовой схемы

	TF-IDF	TF-ICF	TF-RF	ICF-Based	Prob-Based
Accuracy	79.3814%	78.1214%	87.6289%	87.6289%	86.0252%
Precision	89.5688%	89.0756%	90.0125%	90.0125%	90.9804%
Recall	85.9813%	84.9132%	96.2617%	96.2617%	92.9239%
F-Measure	0.8774	0.8694	0.9303	0.9303	0.9194

При классификации текстов блогов были получены одинаковые значения для весовых схем TF-RF и ICF-Based, это означает, что для коллекции блогов, в словаре униграмм нет слов, которые встречаются только в положительном классе или только в отрицательном. Следующий эксперимент был проведен для той же самой коллекции, но классификация проводилась на три класса, результаты представлены в таблице 6.

Таблица 6. Результаты классификации текстов блогов на три класса в зависимости от выбранной весовой схемы

	TF-IDF	TF-ICF
Accuracy	53.9773%	57.9545%
Precision	56.1341%	55.8903%
Recall	53.1164%	53.579%
F-Measure	0.5458	0.5471

При классификации длинных текстов на три класса, обе меры показали неудовлетворительные результаты согласно F-мере.

**7.4. Классификация текстов с использованием морфологического словаря униграмм.** Следующий эксперимент был поставлен с использованием морфологического словаря униграмм. Эксперимент показал, что сокращение размерности вектора признаков сильно сказывается на качестве классификации. Это связано с тем, что практически для 20% слов часть речи не определяется автоматически с помощью инструмента [13], эти слова, как правило, написаны с ошибками, использован сленг или повторение гласных букв внутри слова.

Результаты классификации коротких сообщений с использованием морфологического словаря униграмм на два и три класса представлены в таблицах 7 и 8 соответственно. Результаты

классификации новостных текстов на два и три класса представлены в таблицах 9 и 10 соответственно. Для коллекции блогов не проводилось экспериментов по классификации текстов по тональности с использованием морфологического словаря униграмм из-за сравнительно низких результатов, полученных для первых двух коллекций.

Таблица 7. Результаты классификации текстов коротких сообщений на два класса с использованием морфологического словаря униграмм

	TF-IDF	TF-ICF	TF-RF	ICF-Based	Prob-Based
Accuracy	54.3832%	50.5929%	61.588%	61.5814%	61.5395%
Precision	53.8769%	51.5864%	61.574%	61.5673%	61.487%
Recall	69.153%	40.1662%	64.3068%	64.3025%	64.4461%
F-Measure	0.605666	0.451656	0.6291	0.629	0.6293

Таблица 8. Результаты классификации текстов коротких сообщений на три класса с использованием морфологического словаря униграмм

	TF-IDF	TF-ICF
Accuracy	50.7952%	54.0306%
Precision	51.819%	49.0239%
Recall	52.0946%	52.5921%
F-Measure	0.5196	0.5074

Таблица 9. Результаты классификации новостных текстов на два класса с использованием морфологического словаря униграмм

	TF-IDF	TF-ICF	TF-RF	ICF-Based	Prob-Based
Accuracy	62.2532%	80.4198%	81.4557%	81.4449%	80.9593%
Precision	57.0629%	84.657%	77.7725%	77.7682%	77.2874%
Recall	98.9964%	74.3067%	88.0868%	88.0652%	87.6875%
F-Measure	0.724	0.7914	0.826	0.826	0.8216

Таблица 10. Результаты классификации новостных текстов на три класса с использованием морфологического словаря униграмм

	TF-IDF	TF-ICF
Accuracy	48.3684%	46.0746%
Precision	53.4827%	60.1676%
Recall	48.641%	46.56%
F-Measure	0.50947	0.52496

**Заключение.** Как показали эксперименты, извлечение однословных униграмм из достаточно представительной однородной коллекции коротких эмоционально окрашенных сообщений площадки Твиттер может успешно применяться для создания вектора признаков для SVM-классификатора при решении задачи автоматической

классификации текстов по тональности, независимо от того, к какой предметной области принадлежит текст. Полученные результаты классификации на два класса сопоставимы, а в некоторых случаях даже превосходят результаты классификаторов, спроектированных для классификации одной заранее определенной предметной области. Однако, при классификации на три класса, точность классификатора заметно ухудшается на длинных текстах (новости и тексты блогов).

В результате эксперимента определены весовые схемы, дающие наибольшую точность при классификации текстов по тональности на два и на три класса. Для классификации коротких текста на два класса лучший результат, согласно метрике F-measure, показывает схема ICF-Based, а для неограниченных по длине текстов – схема TF-RF.

Эксперименты показали, что использование морфологической разметки не дает сопоставимые результаты точности классификации в сравнении с использованием словаря униграмм. В дальнейшем планируется объединить словари и поставить эксперимент на объединенном словаре униграмм и морфологических униграмм. Также к перспективам исследования относится поиск возможностей для сокращения размерности вектора признаков.

### Литература

1. *Feldman R.* Techniques and Applications for Sentiment Analysis // Communications of the ACM. 2013. vol. 56. no. 4. pp. 82–89.
2. *Pang B., Lee L.* Thumbs up? Sentiment classification using machine learning techniques // Proc. of the Conference on Empirical Methods in Natural Language Processing (EMNLP). Philadelphia: ACL. 2002. pp. 79–86.
3. *Pang B., Lee L.* Seeing stars: exploiting class relationships for sentiment categorization with respect of rating scales // Proc. of ACL, 43rd Meeting of the Association for Computational Linguistics. Ann Arbor: ACM. 2005. pp. 115–124.
4. *Bespalov D., Bai B., Qi Y., Shokoufandeh A.* Sentiment classification based on supervised latent n-gram analysis // In Proceedings of the 20th ACM international conference on Information and knowledge management (CIKM '11). ACM. New York, NY, USA. 2011. pp. 375–382.
5. *Nguyen D.Q., Nguyen D.Q., Vu T., Pham S.B.* Sentiment classification on polarity reviews: an empirical study using rating-based features // In: 5th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis. Baltimore, Md. 2014. pp. 128–135
6. *Cui H., Mittal V., Datar M.* Comparative experiments on sentiment classification for online product reviews // Proceedings of the 21st national conference on Artificial intelligence. AAAI Press. 2006. vol. 2. pp. 1265–1270.
7. The Streaming APIs. URL: <https://dev.twitter.com/docs/streaming-apis> (дата обращения: 28.10.2014).
8. *Reed J.W., Jiao Y., Potok T.E., Klump B.A., Elmore M.T., Hurson A.R.* TF-ICF: A New Term Weighting Scheme for Clustering Dynamic Data Streams // In: Proc. Machine Learning and Applications (ICMLA '06). 2006. pp. 258–263.
9. *Рубцова Ю.В.* Метод построения и анализа корпуса коротких текстов для задачи классификации отзывов // Электронные библиотеки: перспективные методы и

- технологии, электронные коллекции: Труды XV Всероссийской научной конференции RCDL'2013. Ярославль. 2013. С. 269–275.
10. *Kechedzhy K. E., Usatenko O.V., Yampol'skii V. A.* Rank distributions of words in additive many-step Markov chains and the Zipf law Arxiv LANL // Phys. Rev. E. 2005. vol. 72. pp. 046138(1)–046138(6).
  11. *Manning D., Raghavan P., Schutze H.* Introduction to Information Retrieval // Cambridge University Press. 2008.
  12. *Lan M., Tan C.L., Su J., Lu Y.* Supervised and Traditional Term Weighting Methods for Automatic Text Categorization // IEEE Transactions on Pattern Analysis and Machine Intelligence. vol. 31. no. 4. 2009. pp. 721–735.
  13. *Schmid H.* Probabilistic part-of-speech tagging using decision trees // In Proc. of the International Conference on New Methods in Language Processing. 1994. pp. 44–49.
  14. Коллекция текстов из блогов с разметкой по оценочной тональности и объектам. URL: <http://romip.ru/ru/collections/sentiment-blog-collection-2012.html>. (дата обращения: 28.10.2014).
  15. *Joachims T.* Text categorization with support vector machines: Learning with many relevant features // In Proc. of the European Conference on Machine Learning (ECML 1998). 1998. pp. 137–142.
  16. *Sebastiani F.* Machine Learning in Automated Text Categorization. ACM Computing Surveys. vol. 34. no. 1. March 2002. pp. 1–47.
  17. *Fan R.-E., Chang K.-W., Hsieh C.-J., Wang X.-R., Lin C.-J.* LIBLINEAR: a Library for Large Linear Classification // J. of Machine Learning Research. 2008. vol. 9. pp. 1871–1874.
  18. *Liu Y., Loh H.T., Sun A.* Imbalanced text classification: A term weighting approach // Expert systems with Applications. 2009. vol. 36. no. 1. pp. 690-701.
  19. *Wang D., Zhang H.* Inverse-category-frequency based supervised term weighting scheme for text categorization // arXiv preprint arXiv: 1012.2609. 2010.
  20. *Olson D. L., Dursun D.* Advanced Data Mining Techniques (1st edition) // Springer. 2008. 138 p.

## References

1. Feldman R. Techniques and Applications for Sentiment Analysis. Communications of the ACM. 2013. vol. 56. no. 4. pp. 82–89.
  2. Pang B., Lee L. Thumbs up? Sentiment classification using machine learning techniques. Proc. of the Conference on Empirical Methods in Natural Language Processing (EMNLP). Philadelphia: ACL. 2002. pp. 79–86.
  3. Pang B., Lee L. Seeing stars: exploiting class relationships for sentiment categorization with respect of rating scales. Proc. of ACL, 43rd Meeting of the Association for Computational Linguistics. Ann Arbor: ACM. 2005. pp. 115–124.
  4. Bespalov D., Bai B., Qi Y., Shokoufandeh A. Sentiment classification based on supervised latent n-gram analysis. Proceedings of the 20th ACM international conference on Information and knowledge management (CIKM '11). ACM. New York, NY, USA. 2011. pp. 375–382
  5. Nguyen D.Q., Nguyen D.Q., Vu T., Pham S.B. Sentiment classification on polarity reviews: an empirical study using rating-based features. In: 5th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis. Baltimore, Md. 2014. pp. 128–135
  6. Cui H., Mittal V., Datar M. Comparative experiments on sentiment classification for online product reviews. Proceedings of the 21st national conference on Artificial intelligence. AAAI Press. 2006. vol. 2. pp. 1265–1270
  7. The Streaming APIs. Available at: <https://dev.twitter.com/docs/streaming-apis> (accessed 28.10.2014).
- 74 SPIIRAS Proceedings. 2014. Issue 5(36). ISSN 2078-9181 (print), ISSN 2078-9599 (online)  
[www.proceedings.spiiras.nw.ru](http://www.proceedings.spiiras.nw.ru)

8. Reed J.W., Jiao Y., Potok T.E., Klump B.A., Elmore M.T., Hurson A.R. TF-ICF: A New Term Weighting Scheme for Clustering Dynamic Data Streams. In: Proc. Machine Learning and Applications (ICMLA '06). 2006. pp. 258–263.
9. Rubtsova Y.V. [A Method for development and analysis of short text corpus for the review classification task]. *Trudy XV Vserossiiskoy nauchnoy konferencii RCDL'2013* [In proc. of The XVth All-Russian Scientific Conference RCDL'2013]. Jaroslavl'. 2013. pp. 269–275. (In Russ.)
10. Kechedzhy K. E., Usatenko O.V., Yampol'skii V. A. Rank distributions of words in additive many-step Markov chains and the Zipf law Arxiv LANL. Phys. Rev. E. 2005. vol. 72. pp. 046138(1)–046138(6).
11. Manning D., Raghavan P., Schütze H. Introduction to Information Retrieval. Cambridge University Press. 2008.
12. Lan M., Tan C.L., Su J., Lu Y. Supervised and Traditional Term Weighting Methods for Automatic Text Categorization. IEEE Transactions on Pattern Analysis and Machine Intelligence. vol. 31. no. 4. 2009. pp. 721–735.
13. Schmid H. Probabilistic part-of-speech tagging using decision trees. In Proc. of the International Conference on New Methods in Language Processing. 1994. pp. 44–49.
14. Kollekcija tekstov iz blogov s razmetkoj po ocenочноj tonal'nosti i ob'ektam [A collection of marked by sentiment texts from blogs]. Available at: <http://romip.ru/ru/collections/sentiment-blog-collection-2012.html>. (accessed 28.10.2014). (In Russ.)
15. Joachims T. Text categorization with support vector machines: Learning with many relevant features. In Proc. of the European Conference on Machine Learning (ECML 1998). 1998. pp. 137–142.
16. Sebastiani F. Machine Learning in Automated Text Categorization. ACM Computing Surveys. vol. 34. no. 1. March 2002. pp. 1–47.
17. Fan R.-E., Chang K.-W., Hsieh C.-J., Wang X.-R., Lin C.-J. LIBLINEAR: a Library for Large Linear Classification. J. of Machine Learning Research. 2008. vol. 9. pp. 1871–1874.
18. Liu Y., Loh H.T., Sun A. Imbalanced text classification: A term weighting approach. Expert systems with Applications. 2009. vol. 36. no. 1. pp. 690–701.
19. Wang D., Zhang H. Inverse-category-frequency based supervised term weighting scheme for text categorization. arXiv preprint arXiv: 1012.2609. 2010.
21. Olson D. L., Dursun D. Advanced Data Mining Techniques (1st edition). Springer. 2008. 138 p.

**Рубцова Юлия Владимировна** — аспирант, Институт систем информатики им. А.П. Ершова СО РАН. Область научных интересов: мат. лингвистика, корпусная лингвистика, классификация текстов по тональности. Число научных публикаций — 5. [yu.rubtsova@gmail.com](mailto:yu.rubtsova@gmail.com); 630105 г. Новосибирск, ул. Линейная 47/2; р.т. +79059516757.

**Rubtsova Yuliya Vladimirovna** — Ph.D. student, A.P. Ershov Institute of Informatics Systems. Scientific interests: sentiment analysis, mathematical linguistics, machine learning. The number of publications — 5. [yu.rubtsova@gmail.com](mailto:yu.rubtsova@gmail.com); 630105 Novosibirsk, Lineynaya street 47/2; office phone +79059516757.

## РЕФЕРАТ

### *Рубцова Ю.В.* **Разработка и исследование предметно независимого классификатора текстов по тональности.**

В статье представляется метод построения классификатора для классификации текстов по тональности на два и на три класса (положительные и негативные; положительные, нейтральные и негативные тексты). Во введении обосновывается постановка задачи, приводятся примеры практического применения.

Второй раздел статьи посвящен описанию коллекций из которых будут выделяться признаки для классификатора. Показывается, что коллекции достаточно представительные и однородные и они могут быть использованы для извлечения униграмм.

В третьем разделе описывается алгоритм извлечения признаков для задачи текстовой классификации по тональности. Описана процедура фильтрация коллекций и введены 5 весовых схем для взвешивания униграмм.

В следующем разделе приведены характеристики текстовых коллекций, которые были использованы для экспериментов. Эксперименты ставились на трех типах коллекций:

- короткие сообщения микроблогов;
- новостная коллекция;
- коллекция отзывов, тексты собраны из блогов.

В качестве алгоритма классификации был использован метод опорных векторов, этому посвящен шестой раздел.

Далее были введены и описаны оценки качества классификатора. В качестве оценки качества результатов обучения и работы классификатора выбраны четыре общепринятые характеристики: accuracy, precision, recall и F-мера.

В разделе 7 представлены результаты экспериментов. Всего было поставлено 10 экспериментов на 3-х тестовых коллекциях.

В результате экспериментов определены весовые схемы, дающие наибольшую точность при классификации текстов по тональности на два и на три класса. Для классификации коротких текста на два класса лучший результат, согласно метрике F-measure, показывает схема ICF-Based, а для неограниченных по длине текстов – схема TF-RF.

## SUMMARY

### *Rubtsova Y.V.* **Research and Development of Domain Independent Sentiment Classifier.**

The paper presents a method of constructing a classifier for text classification in tone on the two and three classes (both positive and negative, positive, neutral and negative texts). In the introduction the problem statement, and examples of practical application.

The second section is devoted to describe collections which will be used for feature extraction for the classifier. It was shown that the collections are representative and homogeneous enough to be used for extracting unigrams.

The third section describes the feature extraction algorithm for the task of text sentiment classification. The procedure of filtering collections and description of 5 weight schemes for weighting unigram also explained.

The following section describes the characteristics of text collections, which were used for the experiments. Experiments were performed on three types of collections:

- Short microblogging messages;
- A collection of news;
- A collection of reviews, the texts collected from blogs.

As a classification that algorithm was used is support vector machine, it is dedicated to the sixth section.

Next was introduced and described quality assessment for the classifier. As the assessment of the quality four common measures were selected: accuracy, precision, recall and F-measure.

Section 7 presents the results of experiments. It was delivered 10 experiments on 3 test collection in total.

The experiments determined weighting schemes which are the most accurate in the text sentiment classification into two or three classes. For the short text classification into two classes the best result, according to the F-measure, shows scheme ICF-Based, and for the long texts - scheme TF-RF.

И.С. СОЗИНОВА, А.С. РОМАНОВ, Р.В. МЕЩЕРЯКОВ

## ОПРЕДЕЛЕНИЕ ПОИСКОВОГО СПАМА С ИСПОЛЬЗОВАНИЕМ МЕТОДА ОПОРНЫХ ВЕКТОРОВ

---

*Созинова И.С., Романов А.С., Мещеряков Р.В.* **Определение поискового спама с использованием метода опорных векторов.**

**Аннотация.** В работе рассматривается классификация поискового спама. Выделяются признаки контента веб-страниц, характерные для спама. Предлагается использование метода опорных векторов для определения принадлежности веб-страницы к спаму. Приводятся результаты проведенных экспериментов.

**Ключевые слова:** поисковый спам, поисковый запрос, машина опорных векторов.

*Sozinova I.S., Romanov A.S., Meshcheryakov R.V.* **Search Spam Identification Using Support Vector Machine.**

**Abstract.** The paper deals with the classification of web spam. It is marked out the features of web pages context, which are typical for spam. It is offered to use the method of support vector machines to figure out whether the web page is the spam. The results of the experiments are given here.

**Keywords:** search spam, search query, support vector machine.

---

**1. Введение.** В настоящее время все большую актуальность приобретают методы и средства обработки текстовой информации в окружающем интеллектуальном пространстве [1] и прикладные аспекты генерации различных искусственных текстов. Одним из таких направлений, активно используемым в сети Интернет, стали методы генерации поискового спама.

Поисковый спам (далее спам) - это попытки обмана поисковой системы и манипулирования результатами поиска в целях изменения позиции того или иного веб-сайта [2].

В правилах поисковых систем оговорены пункты, согласно которым использование спама запрещено, а веб-сайты, использующие спам могут быть понижены при ранжировании или вовсе исключены из базы поисковой системы.

Негативное последствие спама заключается в существенном ухудшении качества результатов поисковых запросов. Например, по результатам такого запроса в числе первых выдается веб-сайт, в котором содержится минимум полезной для пользователя информации (в лучшем случае плагиат с оригинального веб-сайта) с максимальным количеством страниц, повторяющих друг друга. Такие веб-сайты могут использоваться для переадресации пользователя на другой ресурс, в продвижении которого заинтересован заказчик, или обычного

мошенничества. Кроме того зачастую такие веб-сайты занимают более высокие позиции в поисковой выдаче, чем более информативные для пользователя веб-сайты, продвижение которых в рейтинге осуществляется их владельцами легальными методами поисковой оптимизации без нарушения правил и влияния непосредственно на алгоритм поиска. Последнее оборачивается для владельцев веб-сайтов дополнительными затратами на их продвижение и возможными финансовыми потерями от упущенных клиентов.

Решение данной проблемы лежит на стыке нескольких областей знаний и носит междисциплинарный характер, при этом имеет большое значение в сфере информационной безопасности, защищая интересы как конечных пользователей поисковой системы, так и владельцев веб-сайтов, не относящихся к поисковому спаму.

**2. Классификация и методы выявления поискового спама.** Общепринятая классификация подразумевает деление спама на два вида: контентный и ссылочный [3]. Ссылочный спам связан с манипулированием внутренними и внешними ссылками на веб-странице и использованием их для перенаправления пользователя на другие веб-сайты. Контентный спам связан с манипулированием содержимым веб-сайта с целью привлечения пользователей.

В соответствии с этой классификацией можно выделить два базовых подхода для проверки результатов поисковой выдачи. Методы выявления ссылочного спама [4] лежат на поверхности, являются хорошо изученными и устоявшимися, в то время как исследование контентного спама требует более глубокого анализа текста веб-страницы и включает несколько аспектов, затрагивающих такие области знания, как статистика, теория вероятностей, лингвистика, морфология, семиотика и пр., а также предполагают наличие знаний об инструментах и методах создания веб-сайтов, их продвижения в сети Интернет [3,5-6]. С точки зрения повышения качества обнаружения спама контентные методы являются более перспективным направлением для исследований.

Существует несколько способов выявления поискового спама, которые подразделяются на его автоматическое детектирование, ручной анализ, а также совместное использование данных методов. Добиться наилучшего результата можно при изучении экспертом веб-страницы на предмет ее принадлежности к поисковому спаму. Но поисковые машины используют автоматическую проверку результатов поисковой выдачи.

Важными направлениями в борьбе с поисковым спамом являются методы обнаружения дубликатов текстов, автоматически сгенерированных и неестественных текстов.

Обзор методов обнаружения дубликатов приведен в работе [7]. В их основе лежит эффективное обнаружение фрагментов скопированных текстов на основе алгоритмов шинглирования.

В основе многих методов обнаружения неестественных текстов лежит подход, предложенный в работе [8]. Этот подход основывается на анализе статистических характеристик текстов и применении машинного обучения для построения автоматического классификатора поискового спама. Развитием данного подхода является работа [9]. В ней предлагается использовать метод скрытого распределения Дирихле для определения спам-текстов.

В работе [10] предлагается подход, основанный на анализе сочетаемости пар слов для обнаружения неестественных текстов. В основе подхода лежит предположение, что неестественные тексты с большей вероятностью содержат редкие пары слов. Авторы предлагают алгоритм для подсчета доли редких пар слов и показывают, что эта характеристика улучшает качество определения поискового спама.

В работе [11] предлагается подход к определению неестественных текстов, в основе которого лежит гипотеза, что такие тексты не могут одновременно удовлетворять всем ограничениям, свойственным естественным текстам. При обучении алгоритма выделяется большое количество статистических признаков, связанных с читаемостью, единством стиля и жанровыми особенностями, которые впоследствии используются в автоматическом классификаторе.

Подход, описанный в работе [12], существенно развивает исследование [11] за счет учета свойств рассматриваемой модели тематической структуры текста для определения неестественных текстов.

Несмотря на существование большого числа методов противодействия спаму, поисковые системы продолжают стабильно выдавать спам-страницы на выходе в определенном проценте случаев (от 1% до 3,5%).

**3. Методика идентификации контентного поискового спама с помощью машины опорных векторов.** В данной работе разработан подход, связанный с анализом характеристик спам-контента, отличающих его от «легальных» веб-страниц. На основе данных характеристик, аналогично работам [11, 12], связанным с машинным обучением, будет производиться классификация с использованием метода опорных векторов (SVM). Стоит отметить, что данный классификатор показывал отличные результаты при решении авторами смежных задач, связанных с обработкой текстов [13-16].

На рисунках 1-2 представлена методика идентификации контентного поискового спама в виде IDEF0 диаграммы (2 уровня).

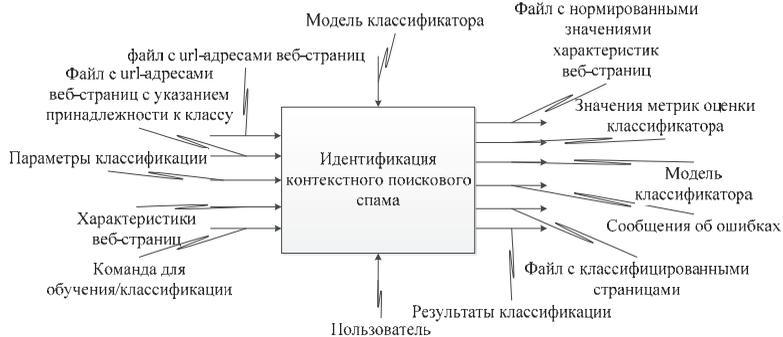


Рис. 1. Схема IDEF0 методики идентификации поискового спама (1 уровень)

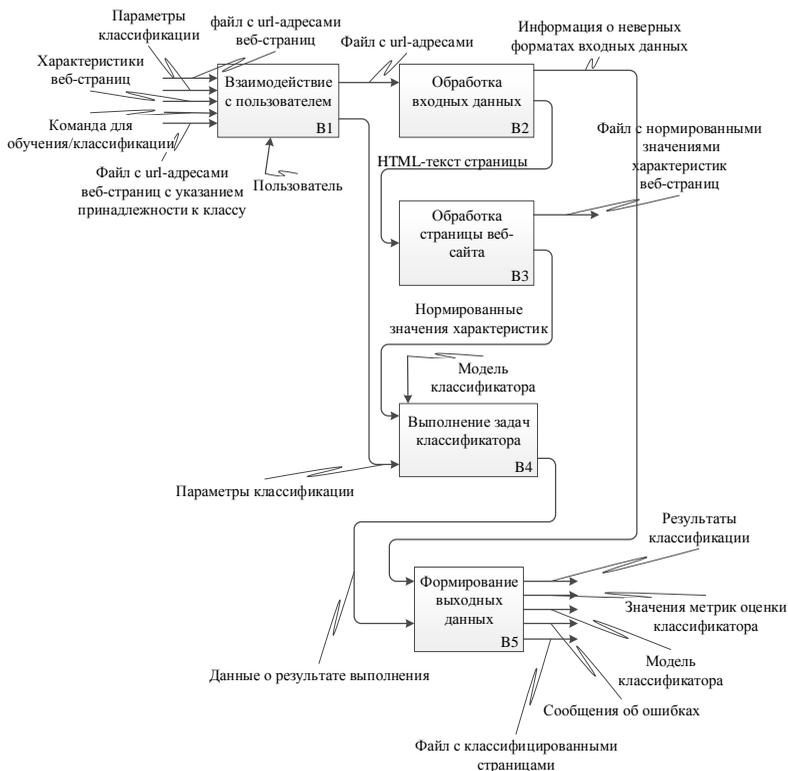


Рис. 2. Схема IDEF0 методики идентификации поискового спама (2 уровень)

Отобранные характеристики с их описаниями приведены в таблице 1.

Таблица 1. Отличительные характеристики контента веб-страницы

№ пп	Характеристика	Описание
	Специфика веб-страницы	
1	Доля текста на странице	Заметно снижается для определенного процента спам-сайтов (берем все, что между тегов)
2	Мета - тег «keywords» (50-80 символов) Количество слов, Плотность слов	Одна из наиболее распространенных характеристик, не показательна для части классического спама.
3	Тег «title» (50-80 символов) Количество слов, Плотность слов	Перенасыщение ключевыми словами заголовка является устойчивой характеристикой для спама
4	Мета-тег «description»: Количество слов, Плотность слов.	Может быть сгенерирован автоматически и перенасыщен ключевыми словами.
5	Анкор: Число ссылок, Без анкоров, Количество внешних, Внешние без анкоров.	Текст анкера может быть перенасыщен ключевыми словами, вследствие чего являться отличительной особенностью для спама
6	Доля анкорного текста	Увеличивается для спама
7	Доля видимого текста	Увеличивается для спама в 50% случаев
	Текстовые характеристики (автоматические/искусственные тексты)	
8	Среднее количество знаков пунктуации на предложение	Требуются дополнительные эксперименты
9	Число слов	Заметно увеличивается для спама
10	Средняя длина слова	Увеличивается для спама
11	Количество длинных слов	Увеличивается для спама
12	Количество знаков экспрессивной пунктуации («!», «?», «...»)	Требуются дополнительные эксперименты
13	Количество слов, начинающихся с заглавной буквы	Требуются дополнительные эксперименты

Метод опорных векторов сводится к задаче оптимизации следующего вида:

$$\begin{cases} \arg \min_{w,b} \|w\|^2, \\ y_i(\langle w, x_i \rangle + b) \geq 1, i = 1, \dots, m' \end{cases}$$

которая является стандартной задачей квадратичного программирования и решается с помощью множителей Лагранжа.

Классифицирующая функция  $F$  принимает вид:

$$F(x) = \text{sign}(\langle w, \varphi(x) \rangle + b).$$

Выражение:

$$k(x, x') = \langle \varphi(x), \varphi(x') \rangle$$

называется ядром классификатора.

Выбор подходящей ядерной функции – ключевая задача при получении качественной модели регрессии.

Классический SVM классификатор не всегда на практике справляется с задачей классификации:

- когда невозможно провести плоскость, разделяющую классы;
- когда выборка из другого класса оказывается слишком мала.

Для того чтобы избежать данных проблем вводится коэффициент регуляризации  $C$ :

$$\min_{\omega, \xi, b} \frac{1}{2} \frac{\|\omega\|^2}{\omega} + C \cdot \sum_{i=1}^n \xi_i.$$

**4. Результаты экспериментов.** Последовательность шагов проведения экспериментов для оценки точности классификации приведена ниже:

– выбор характеристик веб-страницы на основе ее особенностей как HTML-объекта, а также наиболее удачных текстовых характеристик;

– создание тестового, а также обучающего наборов url адресов релевантных сайтов и не релевантных веб-страниц;

– генерация спам-страниц для составления обучающей и тестовой спам-базы;

– подсчет интересующих параметров в выборках;

– нормирование параметров выборок в диапазон  $[0...1]$ ;

– обучение модели SVM на данных обучающей выборки;

– подача на вход обученной модели SVM данных тестовых выборок, работа классификатора, считывание результатов;

– подбор параметров  $C$  и  $\gamma$  методом перебора возможных значений из трех областей:  $\{C \leq 1, 1 < C \leq 100, C > 100\}$ ,  $\{\gamma \leq 1, 1 < \gamma \leq 100, \gamma > 100\}$ . Повтор с шага 5 для каждого выбранного значения. Выбор значений параметров с наилучшим показателем F-меры. Повтор с шага 5 для каждого;

– переопределение набора характеристик веб-страницы; повтор с шага 4 для нового набора характеристик;

– анализ полученных результатов.

На рисунке 3 представлен график зависимости величины f-меры от объема выборки текстов. Вычисление значений f-меры проводилось на основе результатов скользящей проверки с разбиением массива данных на 10 частей. Как видно из графика, классификация для заданных тестовой и обучающей выборки имеет высокие показатели. Полученный результат можно объяснить схожестью групп файлов спама между собой. Небольшие отклонения кривой f-меры укладываются в построенные для точек кривой доверительных интервалов.

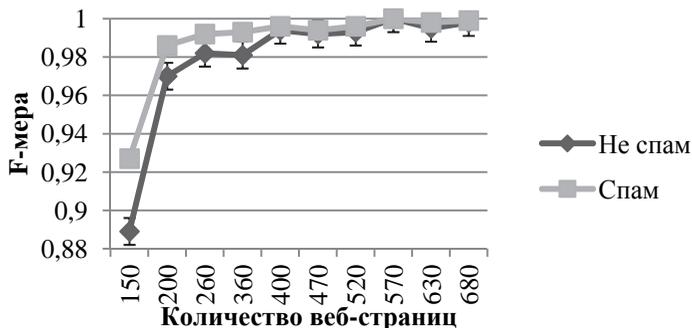


Рис. 3. График зависимости f-мера = f(n выборки)

Были проведены эксперименты по подбору типа ядра классификатора, дающего наиболее высокие показатели. Как видно из гистограммы (рисунок 4), лучше всего с задачей классификации справляется классификатор с заданной функцией сигмоида (SIGMOID), а также радиальной базисной функцией ядра (RBF). Менее значительными являются результаты работы классификатора с полиномиальным (POLY) и линейным (LINEAR) ядром.

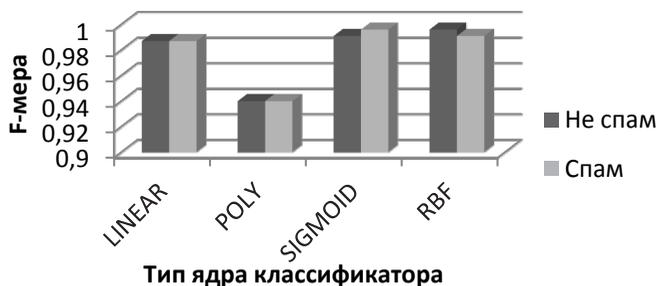


Рис. 4. Гистограмма значений f-меры в зависимости от типа ядра классификатора

В таблице 2 представлен перечень всех рассчитанных метрик для функций ядра: сигмоид и RBF.

Таблица 2. Метрики классификации для двух типов ядер классификатора

Ядро	Аккуратность		Точность		Полнота		F-мера	
	Не спам	Спам	Не спам	Спам	Не спам	Спам	Не спам	Спам
SIGMOID	0,383	0,612	1	0,991	0,986	1	0,993	0,996
RBF	0,357	0,598	1	0,983	0,973	1	0,986	0,991

В ходе экспериментов были подсчитаны метрики классификации, для двух разных групп наборов характеристик веб-страницы (таблица 3). Как видно из гистограммы (рисунок 5) наименее показательными являются характеристики анкоров веб-страницы.

Таблица 3. Группы характеристик веб-страницы

Main	Anchor	Text
Доля текста на странице	Анкор: число ссылок, без анкоров, количество внешних, внешние без анкоров.	Среднее количество знаков пунктуации на предложение
Мета - тег «keywords» (50-80 символов): количество слов, плотность слов.	Доля анкорного текста	Число слов
Тег «title» (50-80 символов) количество слов, плотность слов.	-	Средняя длина слова
Мета-тег «description»: количество слов, плотность слов;	-	Количество длинных слов
Доля видимого текста	-	Количество знаков экспрессивной пунктуации («!», «?», «...»)
-	-	Количество слов, начинающихся с заглавной буквы

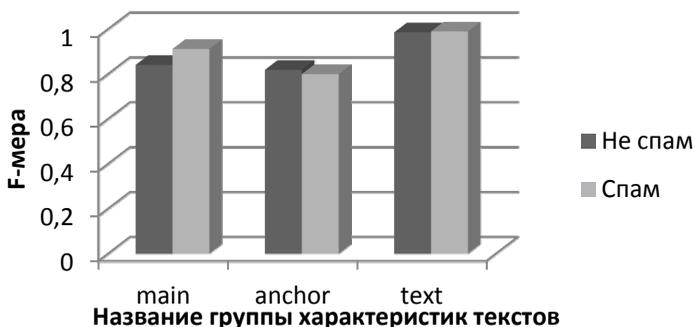


Рис. 5. Гистограмма значений величины f-меры в зависимости от группы исследуемых характеристик

Сравнение результативности методов, основанных на схожем подходе к определению поискового спама, приведено в таблице 4.

Таблица 4. Результаты классификации спама разными методами

	Метрика	Метод обнаружения поискового спама, порожденного с помощью цепей Маркова	Метод контентного анализа характеристик веб-страницы как HTML-объекта	Метод данной работы
Спам	Точность	92,8%	97,8%	98,3%
	Полнота	59,4%	98,7%	100%
	F-мера	72,5%	98,3%	99,1%

Величины метрик классификации в данной работе были взяты для выборки, состоящей из 244 страниц не спама и 120 страниц спама. Полученные результаты в большинстве носят закономерный характер и позволяют сделать вывод о работоспособности предложенного подхода для оценки принадлежности веб-страницы к поисковому спаму. Более того, данная система позволила добиться лучших результатов при использовании аналогичных методов по сравнению с предыдущими исследованиями.

**5. Заключение.** В ходе данной работы были решены следующие задачи:

- анализ и систематизация знаний о предметной области;
- исследование в рамках методов выявления поискового спама, связанных с анализом текстовой составляющей;
- разработка методики выявления поискового спама. В ходе экспериментов были отобраны 18 характеристик, отличающих

поисковый спам от не спама. Была сформирована обучающая база, состоящая из 244 примеров веб-страниц, не относящихся к спаму, и 120 примеров поискового спама;

– анализ результатов автоматизированной классификации поискового спама.

Показатели точности и полноты работы метода для отобранных методом экспертной оценки данных являются довольно высокими. При объеме выборки, составляющей 244 отобранных веб-страниц в качестве представителей класса «Не спам» и 120 страниц-представителей «Поискового спама», F-мера классификации «Не спам» составила 98,6%, «Спама» - 99,1%.

Для классификации веб-страниц из других наборов данных требуются дополнительные эксперименты с увеличением доли поискового спама в выборках для лучшего обучения классификатора.

В дальнейшем планируется провести модификацию разработанной методики и программной системы с целью улучшения показателей классификации. В частности будут проанализированы дополнительные характеристики веб-страницы, обработка которых пока не реализована в существующей версии программы, а также морфологические и синтаксические характеристики самого текста [17].

### Литература

1. *Ronzhin A.L., Karpov A.A.* Russian voice interface // Pattern Recognition and Image Analysis. 2007. vol. 17. no. 2. pp. 321-336.
2. Лицензия на использование поисковой системы Яндексa. URL: <http://legal.yandex.ru/termsfuse/> (дата обращения: 13.01.14).
3. *Gyongyi Z., Garcia-Molina H.* Web Spam Taxonomy // Chiba: First International Workshop on Adversarial Information Retrieval on the Web. 2005. URL: <http://infolab.stanford.edu/> (дата обращения: 26.01.2014).
4. *Зеленков Ю.Г., Сегалович И.В.* Сравнительный анализ методов определения нечетких дубликатов для Web-документов // Тр. IX Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» (RCDL'2007). Переславль. 2007. Т. 1. С. 166–174.
5. *Золтан Д., Гарсиа-Молина Г.* Таксономия веб-спама // Стенфорд: Кафедра информационных технологий Стенфордского университета. URL: <http://wseob.ru/seo/web-spam-taxonomy> (дата обращения: 20.02.2014).
6. Дорвей // свободная статья из Википедиа. URL: <http://www.webeffector.ru/wiki/Дорвей> (дата обращения: 20.11.2013).
7. *Abernethy J., Chapelle O., Castillo C.* WITCH: A new approach to Web spam detection // Proc. Of the 4th Int. Workshop on Adversarial Information Retrieval on the web. Beijing: ACM. 2008. pp. 61–62.
8. *Nioulas A., Najork M., Manasse M., Fetterly D.* Detecting spam Web pages through content analysis // Proc. Of the 15th Int. Conference on World Wide Web. Edinburgh: ACM. 2006. pp. 83–92.

9. *Biro I., Siklosi D., Szabo J., Benczur A.A.* Linked latent Dirichlet allocation in Web spam filtering // Proc. of the 5th Int. Workshop on Adversarial Information Retrieval on the Web. Madrid: ACM. 2009. pp. 37–40.
10. *Гречников Е.А., Гусев Г., Кустарев А.А., Райгородский А.М.* Поиск неестественных текстов // Тр. XI Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции»: Петрозаводск. 2009. С. 306–308.
11. *Павлов А.С., Добров Б.В.* Методы обнаружения поискового спама, порожденного с помощью цепей Маркова // Тр. XI Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции»: Петрозаводск. 2009. С. 311–317.
12. *Павлов А.С., Добров Б.В.* Метод обнаружения массово порожденных неестественных текстов на основе анализа тематической структуры // Тр. XII Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции». Петрозаводск. 2010. С. 210–218.
13. *Романов А.С., Мещеряков Р.В.* Идентификация автора текста с помощью аппарата опорных векторов // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог 2009» (Бекасово, 27-31 мая 2009 г.). М.: РГГУ 2009. Вып. 8 (15). С. 432–437.
14. *Романов А.С., Мещеряков Р.В.* Идентификация авторства коротких текстов методами машинного обучения // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Бекасово, 26-30 мая 2010 г.). М.: Изд-во РГГУ 2010. Вып. 9 (16). С. 407–413.
15. *Романов А.С., Мещеряков Р.В.* Определение пола автора короткого электронного сообщения // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Бекасово, 25 - 29 мая 2011 г.). М.: Изд-во РГГУ. 2011. Вып. 10 (17). С. 620–626.
16. *Романов А.С., Резанова З.И., Мещеряков Р.В.* Методика проверки однородности текста и выявления плагиата на основе метода опорных векторов и фильтра быстрой корреляции // Доклады томского государственного университета систем управления и радиоэлектроники. Томск: Издательство Томского государственного университета систем. 2014. № 2(32). С. 264-269.
17. *Karpov A., Kipyatkova I., Ronzhin A.* Very Large Vocabulary ASR for Spoken Russian with Syntactic and Morphemic Analysis // In Proc. 12th International Conference INTERSPEECH-2011. ISCA Association. Florence. Italy. 2011. pp. 3161-3164.

## References

1. Ronzhin A.L., Karpov A.A. Russian voice interface. Pattern Recognition and Image Analysis. 2007. vol. 17. no. 2. pp. 321-336.
2. Licenzija na ispol'zovanie poiskovoj sistemy Yandexa [License to use the search engine Yandex]. Available at: <http://legal.yandex.ru/termsfuse/> (accessed 26.10.2014) (In Russ.).
3. Gyongyi Z., Garcia-Molina H. Web Spam Taxonomy Chiba: First International Workshop on Adversarial Information Retrieval on the Web. 2005. Available at: <http://infolab.stanford.edu/> (accessed 26.10.2014).
4. Zelenkov Ju.G., Segalovich I.V. [Comparative analysis of methods of near-duplicate detection for Web-documents] *Tr. IX conference "Elektronnye biblioteki:*

- perspektivnye metody i tehnologii, jelektronnye kollekcii*" [Digital Libraries: Advanced Methods and Technologies] (RCDL'2007). Pereslavl. 2007. vol. 1. pp. 166–174. (In Russ.).
5. Zoltan D., Garsia-Molina G. [Taxonomy of web-SPAM]. Stanford: Dept. information technology Stanford university. Available at: <http://wseob.ru/seo/web-spam-taxonomy> (accessed 26.10.2014).
  6. Dorvej [Doorway Page]. Available at: <http://www.webeffector.ru/wiki/Dorvej>. (accessed 26.10.2014) (In Rus).
  7. Abernethy J., Chappelle O., Castillo C. WITCH: A new approach to Web spam detection. Proc. Of the 4th Int. Workshop on Adversarial Information Retrieval on the web. Beijing: ACM. 2008. P. 61–62.
  8. Ntoulas A. M., Najork M., Manasse D., Fetterly D. Detecting spam Web pages through content analysis. Proc. Of the 15th Int. Conference on World Wide Web. Edinburgh: ACM. 2006. pp. 83–92.
  9. Biro I., Siklosi D., Szabo J., Benczur A.A. Linked latent Dirichlet allocation in Web spam filtering. Proc. of the 5th Int. Workshop on Adversarial Information Retrieval on the Web. Madrid: ACM. 2009. pp. 37–40.
  10. Grechnikov E.A., Gusev G.G., Kustarev A.A., Rajgorodskij A.M. [Search unnatural texts]. *Tr. XI conference "Jelektronnye biblioteki: perspektivnye metody i tehnologii, jelektronnye kollekcii"* [Digital Libraries: Advanced Methods and Technologies]. Petrozavodsk. 2009. pp. 306–308. (In Russ.).
  11. Pavlov A.S., Dobrov B.V. [Detecting web spam created with Markov Chains]. *Tr. XI conference "Jelektronnye biblioteki: perspektivnye metody i tehnologii, jelektronnye kollekcii"* [Digital Libraries: Advanced Methods and Technologies]. Petrozavodsk. 2009. pp. 311–317. (In Russ.).
  12. Pavlov A.S., Dobrov B.V. [Detection method massively generated unnatural texts based on an analysis of the thematic structure] *Tr. XII conference "Jelektronnye biblioteki: perspektivnye metody i tehnologii, jelektronnye kollekcii"* [Digital Libraries: Advanced Methods and Technologies]. Petrozavodsk. 2010. pp. 210–218. (In Russ.).
  13. Romanov A.S., Meshcheryakov R.V. [Identification of the author of text by using support vector machine]. *Komp'juternaja lingvistika i intellektual'nye tehnologii: Po materialam ezhegodnoj Mezhdunarodnoj konferencii «Dialog»* [Computational Linguistics and Intellectual Technologies]. M.: RGGU. 2009. vol. 8(15). pp. 432–437. (In Russ.).
  14. Romanov A.S., Meshcheryakov R.V. [The identification of authorship short texts machine learning methods]. *Komp'juternaja lingvistika i intellektual'nye tehnologii: Po materialam ezhegodnoj Mezhdunarodnoj konferencii «Dialog»* [Computational Linguistics and Intellectual Technologies]. M.: Izd-vo RGGU. 2010. vol. 9 (16). pp. 407–413. (In Russ.).
  15. Romanov A.S., Meshcheryakov R.V. *Opredelenie pola avtora korotkogo jelektronnoogo soobshhenija* [Sexing the author of a short e-mail message]. *Komp'juternaja lingvistika i intellektual'nye tehnologii: Po materialam ezhegodnoj Mezhdunarodnoj konferencii «Dialog»* [Computational Linguistics and Intellectual Technologies]. M.: Izd-vo RGGU. 2011. vol. 10 (17). pp. 620–626. (In Russ.).
  16. Romanov A.S., Rezanova Z.I., Meshcheryakov R.V. [Methodology for testing homogeneity of the text and plagiarism detection method based on support vector machines and fast correlation filter]. *Doklady TUSUR – Reports of TUSUR*. Tomsk: Izdatel'stvo Tomskogo gosudarstvennogo universiteta sistem. 2014. vol. 2(32). pp. 264–269. (In Russ.).

17. Karpov A., Kipyatkova I., Ronzhin A. Very Large Vocabulary ASR for Spoken Russian with Syntactic and Morphemic Analysis. In Proc. 12th International Conference INTERSPEECH-2011. ISCA Association. Florence. Italy. 2011. pp. 3161-3164.

**Созинова Ирина Сергеевна** — инженер кафедры Комплексной информационной безопасности электронно-вычислительных систем ТУСУР. Область научных интересов: информационная безопасность. Число научных публикаций — 5. irishechka7371@gmail.com; 634045, г. Томск, ул. Красноармейская 146, ауд. 509; р.т. +7 (3822) 900-111, факс +7 (3822) 900-111.

**Sozinova Irina Sergeevna** — engineer, Dept. of Complex Security of Electronic-computing Systems of Tomsk State University of Control Systems and Radioelectronics (TUSUR). Research interests: information security. The number of publications — 5. alexx.romanov@gmail.com; KIBEVS Dept. TUSUR, 146, Red Army street Tomsk, 634045, Russia; office phone +7(3822)900-111, fax +7(3822)900-111.

**Романов Александр Сергеевич** — к-т техн. наук, доцент кафедры Комплексной информационной безопасности электронно-вычислительных систем ТУСУР. Область научных интересов: информационная безопасность, интеллектуальный анализ данных, искусственный интеллект, обработка текста. Число научных публикаций — 41. alexx.romanov@gmail.com; 634045, г. Томск, ул. Красноармейская 146, ауд. 509; р.т. +7(3822) 900-111, факс +7 (3822) 900-111.

**Romanov Aleksandr Sergeevich** — Ph.D., associate professor, Dept. of Complex Security of Electronic-computing Systems of Tomsk State University of Control Systems and Radioelectronics (TUSUR). Research interests: information security, data mining, AI, word processing. The number of publications — 41. alexx.romanov@gmail.com; KIBEVS Dept. TUSUR, 146, Red Army street Tomsk, 634045, Russia; office phone +7(3822)900-111, fax +7(3822)900-111.

**Мещеряков Роман Валерьевич** — д-р техн. наук, профессор кафедры Комплексной информационной безопасности электронно-вычислительных систем ТУСУР. Область научных интересов: системный анализ, информационная безопасность, вопросы обработки информации в интеллектуальных системах, особое внимание уделяется вопросам создания информационно-безопасных систем. Число научных публикаций — 247. mrv@security.tomsk.ru; 634050, г. Томск, пр. Ленина, 40, ауд. 210; р.т. +7(3822)900111, факс +7 (3822) 900-111.

**Meshcheriakov Roman Valerievich** — Ph.D., Dr. Sci., professor, Dept. of Complex Security of Electronic-computing Systems of Tomsk State University of Control Systems and Radioelectronics (TUSUR). Research interests: speech analysis, speech recognition, medical technology, information security. The number of publications — 247. mrv@security.tomsk.ru; KIBEVS Dept. TUSUR, 40, Lenin-avenue Tomsk, 634050, Russia; office phone +7(3822)413426, fax +7(3822)900-111.

## РЕФЕРАТ

### *Созинова И.С., Романов А.С., Мещеряков Р.В.* **Определение поискового спама с использованием метода опорных векторов.**

Поисковый спам – это попытки обмана поисковой системы и манипулирования результатами поиска в целях изменения позиции того или иного веб-сайта. Решение данной проблемы лежит на стыке нескольких областей знаний и носит междисциплинарный характер, при этом имеет большое значение в сфере информационной безопасности, защищая интересы как конечных пользователей поисковой системы, так и владельцев веб-сайтов, не относящихся к поисковому спаму.

Общепринятая классификация подразумевает деление спама на два вида: контентный и ссылочный. Важными направлениями в борьбе с поисковым спамом являются методы обнаружения дубликатов текстов, автоматически сгенерированных и неестественных текстов.

В данной работе используется подход, связанный с углублением в характеристики спам-контента, отличающих его «легальных» веб-страниц. В ходе экспериментов были отобраны 18 характеристик, отличающих поисковый спам от не спама. На основе данных характеристик производится классификация с использованием метода опорных векторов (SVM). Показатели точности и полноты работы метода для отобранных методом экспертной оценки данных являются довольно высокими. При объеме выборки, составляющей 244 отобранных веб-страниц в качестве представителей класса «Не спам» и 120 страниц-представителей «Поискового спама», F-мера классификации «Не спама» составила 98,6%, «Спама» – 99,1%.

## SUMMARY

### *Sozinova I.S., Romanov A.S., Meshcheryakov R.V.* **Search engine spam detection using support vector machine.**

Search engine spam is attempts to deceive search engine and manipulate search results in order to change the position of a website. The solution of this problem lies at the intersection of several disciplines and interdisciplinary character, and has a great importance in the field of information security, protecting the interests of the search engine end-user and website owners, not related to search engine spam.

Common classification implies the division of spam into two types: a content and reference. Important areas in the fight against search spam are detection methods of duplicate texts, automatically generated and unnatural texts.

In this paper we use the approach of a detailed examination of the characteristics of spam content, distinguishing it from "legal" web pages. During the experiments were selected 18 characteristics that distinguish search engine spam from not spam. Based on these characteristics, we classify web pages using support vector machine (SVM). Accuracy and precision of the method are quite high. When the sample size is 244 selected web pages as representatives of a class "Not Spam" and 120 pages, representatives of the "search spam", F-measure of classification "Not Spam" amounted to 98.6%, "spam" – 99.1%.

Д.А. ВАЖЕНИНА, И.С. КИПЯТКОВА, К. МАРКОВ, А.А. КАРПОВ  
**МЕТОДИКА ВЫБОРА ФОНЕМНОГО НАБОРА ДЛЯ  
АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ РУССКОЙ РЕЧИ**

---

*Важенина Д.А., Кипяткова И.С., Марков К., Карпов А.А. Методика выбора фонемного набора для автоматического распознавания русской речи.*

**Аннотация.** В статье описывается выбор оптимального фонемного набора для системы автоматического распознавания русской речи. При создании акустических моделей был предложен комбинированный метод для выбора наилучшего фонемного набора, объединяющий статистическую информацию и фонетические знания. В результате применения данного метода к русскому фонетическому набору алфавита IPA (International Phonetic Alphabet) был получен набор из 47 фонологических единиц, который был преобразован в несколько фонемных наборов с разным размером от 27 до 47 единиц. Эксперименты по распознаванию речи показали, что использование сокращенных фонемных наборов позволяет увеличить точность распознавания фонем. В ходе экспериментов с применением расширенной языковой модели и сверхбольшим словарем точность распознавания слов составила 73,1%. Полученные результаты соответствуют качеству распознавания слитной русской речи, полученному на настоящий момент другими организациями.

**Ключевые слова:** автоматическое распознавание русской речи, акустическое моделирование, выбор фонемного набора.

*Vazhenina D.A., Kipyatkova I.S., Markov K., Karpov A.A. Technique for Phoneme Set Selection for Automatic Russian Speech Recognition.*

**Abstract.** In the paper, selection of best phoneme set for Russian automatic speech recognition is described. For the acoustic modeling, we describe a method based on combination of knowledge-based and statistical approaches to create several different phoneme sets. Applying this method to the Russian phonetic set of the IPA (International Phonetic Alphabet) alphabet, we first reduced it to 47 phonological units and derived several other phoneme sets with different number of phonological units from 27 till 47. Speech recognition experiments using these sets showed that reduced phoneme sets are better for phoneme recognition task and as good for word level speech recognition. For experiment with extra-large vocabulary, we used syntactico-statistical language model, which allowed us to achieve the word recognition accuracy of 73.1%. The results correspond to continuous Russian speech recognition quality obtained by other organizations up to date.

**Keywords:** automatic Russian speech recognition, acoustic modeling, phoneme set selection.

---

**1. Введение.** Автоматическое распознавание слитной русской речи представляет собой очень сложную задачу из-за ряда особенностей языка. Русский язык является флективным языком со сложной системой словообразования, что приводит к существенному увеличению размера словаря системы распознавания, а также к увеличению коэффициента неопределенности (replexity) статистических  $n$ -граммных моделей языка [1]. Для того чтобы создать словарь большого объема, обычно используются автоматические транскрипторы, которые преобразуют слова в фонематическую транскрипцию, используя правила транскрибирования. Для русского языка правила транскриби-

рования не являются сложными, основная проблема состоит в том, чтобы определить положение ударения в слове, кроме того, в сложных словах может быть несколько ударных гласных [2].

При создании фонематических транскрипций обычно используется Международный фонетический алфавит (International Phonetic Alphabet) IPA/МФА, русский вариант которого состоит из 55 фонологических единиц: 38 согласных и 17 гласных. Большое количество согласных, которые более сложны для распознавания за счет меньшей длительности звучания, обусловлено специфичной для русского языка палатализацией большинства твердых согласных. Такие пары имеют небольшое различие в дополнительной артикуляции: подъем средней части языка к небу во время основной артикуляции согласного. Наличие таких пар согласных усложняет распознавание, поскольку увеличивает схожесть согласных между собой. Для сравнения: в американском варианте английского алфавита IPA содержится 49 фонологических единиц: 24 согласных и 25 гласных и дифтонгов. Таким образом, при меньшем общем числе фонем количество согласных значительно меньше, чем в IPA для русского языка, и их соотношение к гласным и дифтонгам, которые менее сложны для распознавания, примерно равное.

Набор гласных фонем в алфавите IPA включает также их редуцированные варианты. Все безударные гласные в русском языке подвержены редукции в длительности и все, кроме [u], подвержены артикуляторной редукции. Так безударный гласный звук [e] становится близким по произношению к [i], а безударный гласный [o] в большинстве случаев произносится как [a], исключая случаи с иностранными словами, такими как “радио”.

Выбор фонемного набора является одним из первых шагов при разработке системы распознавания речи. Под фонемным набором подразумевается совокупность фонологических единиц, которые используются как базовые акустические единицы при создании акустических моделей в системе распознавания речи. От выбора этого набора может зависеть точность и быстродействие всей системы.

В данной статье представлено описание методики выбора наилучшего набора фонем для систем распознавания русской речи. Предложенная методика основана на использовании как фонетических знаний, так и статистической информации. Для оценки эффективности методики были проведены эксперименты с использованием контекстно-независимых моделей для распознавания фонем и контекстно-зависимых моделей для распознавания слов.

**2. Обзор систем распознавания русской речи.** Обзор систем автоматического распознавания русской речи приведен в [3-5]. Ниже рассмотрены некоторые новые работы в этой области.

В [6] описывается система распознавания русской речи, разработанная в ходе проекта Quaero (<http://www.quaero.com/>). Система использует два различных подхода для обучения акустических моделей. 4-граммная модель языка со словарем 500 тыс. слов была обучена на материалах радионовостей, интернет-данных, книгах, транскрипциях аудиоданных. Точность распознавания слов, полученная на корпусе Quaero 2010, составила 80 %.

В работе [7] представлена модель русского языка на основе максимума энтропии, использующая признаки, которые позволяют учесть такие особенности русского языка как флективность и свободный порядок слов. Эта модель, объединенная с моделью языка, основанной на частях слов, использовалась для переоценки списка лучших  $N$  гипотез, что позволило увеличить точность распознавания слов на 1,2 %.

Система распознавания слитной русской речи с большим словарем, использующая слоговую модель языка, представлена в [8]. Словарь системы содержит 12 тыс. слогов. В работе предложен метод объединения слов и коррекции ошибок. Гипотеза произнесенной фразы создается из распознанных символов путем применения коэволюционного асимптотического вероятностного генетического алгоритма CAPGA (co-evolutionary asymptotic probabilistic genetic algorithm).

В [9] описывается метод учета синтаксических связей в модели языка, при котором используются следующие стадии обработки обучающего текстового корпуса: обозначение части речи, анализ синтаксических зависимостей и создание факторной модели языка для переоценки гипотез распознавания. Наилучшая точность распознавания, полученная при проведении экспериментов на части Национального корпуса русского языка, составила 91,77 %, что на 1,26 % лучше, чем точность, полученная при применении базовой модели.

Распознавание слитной русской речи с применением глубоких нейронных сетей (deep neural networks) совместно со скрытыми марковскими моделями представлено в [10]. На первой стадии распознавания для вычисления векторов признаков использовались глубокие нейронные сети. На второй стадии декодер Витерби осуществлял генерацию распознанной последовательности слов, используя вероятности, полученные на первой стадии. Эксперименты были выполнены на корпусе телефонной речи, содержащем 25 часов обучающих данных, 1 час данных для настройки системы и 1 час тестовых данных. Дополнительно 17 часов неразмеченных речевых данных использовались для

предварительного обучения глубоких нейронных сетей. Наивысшая точность распознавания составила 45 %, при этом использовалась следующая конфигурация нейронных сетей: 5 слоев с 1000 элементами на один слой.

Для голосового поиска в сети Интернет компания Google разработала функцию Voice Search [11], которая позволяет пользователям находить нужную им информацию, произнося слова или фразы вслух. Для создания модели языка использовались запросы пользователей к поисковой системе Google. Также эта технология применяется для других сервисов Google, например, она позволяет использовать речевые запросы для поиска места на карте.

Недавно компания Яндекс представила свою систему распознавания речи SpeechKit [12]. В настоящее время данное приложение позволяет искать общую информацию (то, что люди обычно ищут в Интернете) и геоинформацию (адреса, названия организаций). Точность распознавания общей информации составляет 84%, геоинформации – 94%.

В работе [13] представлена система автоматического создания в режиме реального времени субтитров для телевизионных новостей. Система работает следующим образом: специально обученный диктор (респикер) пересказывает речь диктора новостей, преобразуя ее в форму, более соответствующую письменной речи. Система распознавания речи обучена на голос респикера, что позволяет увеличить точность распознавания (точность распознавания слов составила 94,5 %). Затем редактор субтитров исправляет ошибки, допущенные системой распознавания. Таким образом, количество точность распознавания слов в субтитрах после их проверки редактором составляет 99,98%. Данная система использовалась в ходе паралимпийских игр в Сочи.

**3. Акустическое моделирование и выбор фонемного алфавита.** Создание моделей акустических единиц речи является первым этапом обучения системы автоматического распознавания речи. В зависимости от задачи и объема обучающих данных необходимо определить список моделируемых фонетических единиц, в качестве которых могут использоваться части слова, слоги, контекстно-независимые фонемы или контекстно-зависимые фонемные реализации. Преимуществом использования контекстно-зависимых единиц является их способность моделировать эффекты коартикуляции между соседними звуками, поэтому в современных системах распознавания речи контекстно-независимые модели (монофоны), которые соответствуют фонологическим единицам фонемного набора, часто заменяются контекстно-зависимыми моделями (трифонами). Как видно из примера, приведен-

ного в таблице 1, монофон /a/ встречается в модели дважды, но, с учетом контекста левой и правой фонем, они образуют разные трифоны.

Таблица 1. Фонетическое моделирование слова "пара" с использованием контекстно-независимых и контекстно-зависимых моделей

	Фонетические единицы			
Монофоны	/p/	/a/	/r/	/a/
Трифоны	/sil-p+a/	/p-a+r/	/a-r+a/	/r-a+sil/

**3.1. Разработка дерева решений для русского языка.** При использовании трифонов может возникнуть проблема дефицита обучающих данных. Для обучения каждой модели трифона необходимо наличие большого числа наблюдений соответствующего трифона, и зачастую не все возможные трифоны присутствуют в обучающих данных. Наиболее распространенным решением данной проблемы является кластеризация и связывание состояний моделей трифонов, чьи контексты относятся к одному кластеру. Обычно кластеризация контекстов осуществляется с помощью деревьев решений [14,15], при этом используются вопросы о том, имеет ли левая или правая фонема определенные фонетические признаки (например, является ли левая/правая фонема звонкой).

В данном исследовании в качестве основы для разработки фонетического дерева для русского языка было взято фонетическое дерево для английского языка [16], которое было дополнено вопросами, относящимися к специфике фонологии русского языка, например, является ли левая/правая контекстная фонема мягкой. Некоторые вопросы, которые относятся к особенностям английского языка, были удалены. В таблице 2 приведены основные различия между наборами вопросов дерева решений для английского и русского языков. Таким образом, дерево решений для русского языка состоит из 38 общих вопросов плюс по одному вопросу для каждой единицы фонемного набора отдельно для левого и правого контекста.

**3.2. Методика выбора оптимального фонемного набора.** Размер используемого фонемного набора определяет количество контекстно-независимых моделей и также влияет на число контекстно-зависимых моделей. Если их число излишне велико, возрастает вычислительная сложность системы распознавания. Напротив, если их количество слишком мало, может снизиться точность системы, так как акустически схожие модели будут чаще распознаваться неправильно (спутываться). На данный момент существует два основных подхода к выбору такого набора для акустического моделирования: основанный на знаниях и статистический.

Таблица 2. Различия между наборами вопросов дерева решений для английского и русского языков

Удаленные вопросы	Добавленные вопросы для русского
Является ли левый/правый гласный звук долгим? Является ли левый/правый гласный звук коротким? Является ли левый/правый гласный звук дифтонгом? Является ли левый/правый гласный звук редуцированным?	Является ли левый/правый гласный звук ударным? Является ли левый/правый гласный звук безударным?
Является ли левый/правый согласный звук синлабическим? Является ли левый/правый согласный звук непрерывным?	Является ли левый/правый согласный звук мягким? Является ли левый/правый согласный звук твердым? Является ли левый/правый согласный звук дрожащим?

Для русского языка часто используются фонемные наборы, разработанные экспертами на основе лингвистических и фонологических правил [17], поскольку правила преобразования орфографического текста в фонемное представление для русского являются относительно несложными. Прямое преобразование графем в акустические единицы дает набор из 49 единиц, такой набор фонем использовался в [18] для сравнения с графемной системой распознавания речи. В работе [19] использовался набор из 43 фонологических единиц, который являлся стандартным алфавитом SAMPA для русского языка с добавлением фонемы /ʏ/. Для распознавания слитной русской речи с большим словарем в работе [20] был предложен фонемный набор из 59 единиц. В большинстве случаев исследователи используют расширенные наборы для представления гласных, включающие их ударные и безударные варианты [20-22].

Однако для других языков существуют исследования, в которых для получения фонемного набора используется статистическая информация. Для китайского языка было предложено использовать взаимную информацию между частями слов и их фонетическими транскрипциями в обучающем тексте [23]. Фонемные наборы создавались путем объединения тонально-зависимых фонем, которое приводит к минимальному уменьшению значения взаимной информации. Это позволило значительно сократить количество контекстно-зависимых моделей.

Для английского языка было предложено автоматически генерировать набор акустических единиц, основываясь на речевых данных [24].

Фонематические транскрипции в этом исследовании создавались с применением полученных акустических единиц, что позволило увеличить точность системы распознавания. Однако это является и главным недостатком системы из-за сложности добавления в нее новых слов.

В данной работе использовался комбинированный метод, в котором объединяется информация, полученная из фонологических знаний, и статистические данные от фонемной матрицы спутывания. Фонологические знания включают в себя правила произношения и информацию о фонологических чередованиях. Данный подход позволяет определить акустически близкие фонологические единицы. С помощью матрицы спутывания (см. рисунок 1) определяются наиболее часто несовпадающие монофоны. Для выбора наилучшего фонемного набора вначале используется наибольший набор, затем его размер постепенно уменьшается путем удаления или объединения некоторых фонологических единиц.

	a	a'	g	g'	e	e'	s	s'
a	6969	626	14	0	82	77	20	7
a'	685	3076	0	0	25	102	3	0
g	40	2	562	0	4	1	0	0
g'	2	0	1	38	1	0	1	0
e	121	11	5	0	1016	102	7	8
e'	100	40	1	0	125	2168	5	0
s	19	5	0	0	3	2	2248	59
s'	6	2	0	0	5	7	123	1293

Рис. 1. Пример матрицы спутывания монофонов (количество распознанных фонем внутри пары выделено квадратом)

Методика выбора фонем (см. рисунок 2) включает в себя следующие шаги:

1. Из алфавита IPA выбирается набор P0 путем применения фонологических правил произношения, чтобы найти наиболее акустически близкие единицы, например, гласные, различающиеся уровнем редукации.

2. В соответствии с фонологической спецификой языка определяются пары фонологических единиц, являющиеся кандидатами на объединение. Эти пары включают в себя мягкие/твердые согласные и ударные/безударные гласные.

3. С помощью набора P0 выполняется фонемное распознавание, и создается матрица спутывания монофонов. Для выбранных пар

вычисляется коэффициент спутывания (CR), который определяется следующим образом:

$$CR = \frac{M_1 + M_2}{H_1 + M_1 + H_2 + M_2} \cdot 100\%$$

где  $H_1$  – количество правильно распознанных появлений первого монофона в паре (например, /а/ распознана как /а/),  $H_2$  – количество правильно распознанных появлений второго монофона в паре (например, /а!/ распознана как /а!/),  $M_1$  – количество неправильно распознанных появлений первого монофона в паре (например, /а/ распознана как /а!/),  $M_2$  – количество неправильно распознанных появлений второго монофона в паре (например, /а!/ распознана как /а/). Чем больше коэффициент спутывания, тем больше монофоны в паре не совпадают, что делает их вероятными кандидатами для объединения.

4. Пары монофонов сортируются по уменьшению коэффициента спутывания.

5. Выбираются первые  $N$  пар монофонов, и после объединения соответствующих им фонологических единиц получается новый фонемный набор.

Выбор различных значений  $N$  позволяет получить несколько фонемных наборов. Наиболее подходящий набор можно определить, оценивая точность и скорость распознавания речи, полученную при его использовании.

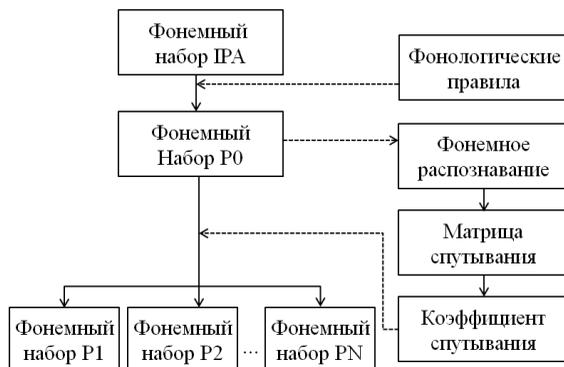


Рис. 2. Методика выбора фонем для системы автоматического распознавания русской речи

**3.3. Моделирование произношения.** Основными проблемами автоматического создания фонематических транскрипций для слов русского языка являются: плавающее ударение, замена буквы "ё" на

"е", большое количество омографов, а также явления редукции и ассимиляции звуков, возникающие в разговорной речи.

В данной работе фонематические транскрипции слов создавались автоматически путем применения к списку слов правил транскрибирования [25]. Для автоматической генерации транскрипций использовалась база данных словоформ русского языка с отмеченным ударением, которая была создана путем объединения двух баз данных, свободно доступных в Интернете: морфологической базы данных проекта STARLING ([starling.rinet.ru](http://starling.rinet.ru)) и морфологической базы данных проекта AOT ([www.aot.ru](http://www.aot.ru)). Объем полученной базы данных составляет 2,3 млн словоформ. Подробнее процесс создания фонематических транскрипций слов описан в [26].

**4. Базы данных слитной русской речи.** Для обучения акустических моделей и проведения экспериментов по распознаванию русской речи использовались два речевых корпуса. Оба корпуса используют частоту дискретизации 16 кГц. Первый корпус – это корпус русской речи GlobalPhone [18], который содержит записи чтения русских газетных статей. Корпус состоит из 12321 фразы, в записи корпуса приняли участие 115 человек (61 мужчина и 54 женщины). Общий объем записей составил 26 часов. Из корпуса были удалены фразы, содержащие речевые сбои и шум, оставшаяся часть корпуса разделена на обучающий корпус (15 часов 25 минут) и тестовый корпус (1 час 40 минут), который содержит записи 5 мужчин и 5 женщин.

Второй корпус был записан в СПИИРАН [27]. В записи корпуса приняли участие 50 дикторов - носителей русского языка (25 женщин и 25 мужчин). Корпус содержит 327 фонетически сбалансированных фраз на русском языке (321 фраза состоит из одного предложения, остальные фразы являются короткими текстами, содержащими по несколько предложений). Общий объем корпуса -13,5 Гб, длительность записей – более 21 часа. Запись проводилась в специальной звукоизолированной комнате, отношение сигнал/шум составляло больше 35 дБ. Дополнительно для экспериментов по распознаванию русской речи были записаны 100 фраз, произнесенных двумя дикторами (мужчиной и женщиной). Длительность записи составила 30 минут. Фразы взяты из материалов интернет-газеты «Фонтанка.ru» ([www.fontanka.ru](http://www.fontanka.ru)). Каждая фраза в этом тестовом корпусе состоит из одного предложения длиной от 6 до 20 слов.

## **5. Описание экспериментов и их результаты.**

**5.1. Условия проведения эксперимента.** Для каждого набора фонем была создана одна система распознавания речи, использующая монофоны, и три различные системы, использующие трифоны, в которых применялась одинаковая акустическая модель, но разные по раз-

меру словаря и модели языка. В первой системе использовалась закрытая биграммная модель языка, обученная на транскрипциях речевой базы данных СПИИРАН, размер словаря составлял 1146 слов, значение коэффициента неопределенности было равно 241. Оценка системы производилась на тестовой части базы данных СПИИРАН. Вторая система построена аналогичным образом, но для обучения биграммной модели языка использовались транскрипции базы данных GlobalPhone, размер словаря составил 20 тыс. слов, значение коэффициента неопределенности – 137. Тестирование системы производилось на тестовой части базы данных GlobalPhone. Третья система использовала биграммную модель языка со словарем 204 тыс. слов, эта модель применялась для задачи распознавания речи со сверхбольшим словарем. Для оценки третьей системы также использовалась тестовая часть базы данных GlobalPhone, значение коэффициента неопределенности составило 844, количество внесловарных слов – 3,41%.

Акустические модели создавались с помощью инструментария НТК (Hidden Markov Model Toolkit) [15]. Для извлечения информативных признаков речевого сигнала использовались энергия сигнала и 12 мел-частотных кепстральных коэффициентов с их 1-й и 2-й производными. Данные признаки были вычислены путем анализа с помощью 26-канальной гребенки фильтров сегментов речи длительностью 20 мс с 10 мс перекрытием между сегментами. В качестве акустических моделей использовались лево-правые скрытые марковские модели с тремя состояниями.

**5.2. Выбор фонемных наборов.** На основе фонологических особенностей алфавита IPA был получен фонемный набор P0 путем выполнения следующих шагов [28]:

1. Ударные гласные звуки [a], [æ] и [ɑ] были объединены, поскольку они являются акустически близкими и различаются только от их местоположения в слове.

2. Безударные гласные звуки [e] и [ə] были объединены, поскольку их произношения отличаются лишь незначительно в зависимости от расстояния от ударного слога. Аналогично были объединены [u] и [o].

3. Звук [ø] является похожим на комбинацию звуков [j] и [o], и соответственно он был представлен как 2 фонемы.

4. Согласные звуки [z] и [ʒ] были исключены, поскольку они используются только в некоторых диалектах русского языка.

Получившийся фонемный набор P0 представлен в таблице 3. Символ «!» обозначает ударные гласные, символ «'» — мягкие (палатализованные) согласные.

Таблица 3. Фонемный набор P0

Классы фонологических единиц		Список фонологических единиц
Согласные	Твердые	/b/, /v/, /g/, /d/, /zh/, /z/, /k/, /l/, /m/, /n/, /p/, /r/, /s/, /t/, /f/, /h/, /c/, /sh/
	Мягкие	/b'/, /v'/, /g'/, /d'/, /z'/, /j/, /k'/, /l'/, /m'/, /n'/, /p'/, /r'/, /s'/, /t'/, /f'/, /h'/, /ch/, /sch/
Гласные	Ударные	/a!/, /e!/, /i!/, /o!/, /u!/, /y!/
	Безударные	/a/, /e/, /i/, /u/, /y/

Набор P0 содержит 47 фонологических единиц: 6 ударных и 5 безударных гласных и 36 согласных. В качестве кандидатов на объединение были выбраны пары твердых и мягких согласных, потому что они акустически схожи между собой, а также пары ударных и безударных гласных, поскольку они различаются только в длительности звука.

С использованием набора P0 была создана система распознавания фонем и выполнены эксперименты по их распознаванию. Из полученной матрицы спутывания был сформирован список фонемных пар, отсортированный по величине коэффициента спутывания (см. рисунок 3). Из рисунка видно, что парами с наибольшим процентом спутывания являются пары гласных, поэтому они были объединены, в результате чего был получен фонемный набор P1.

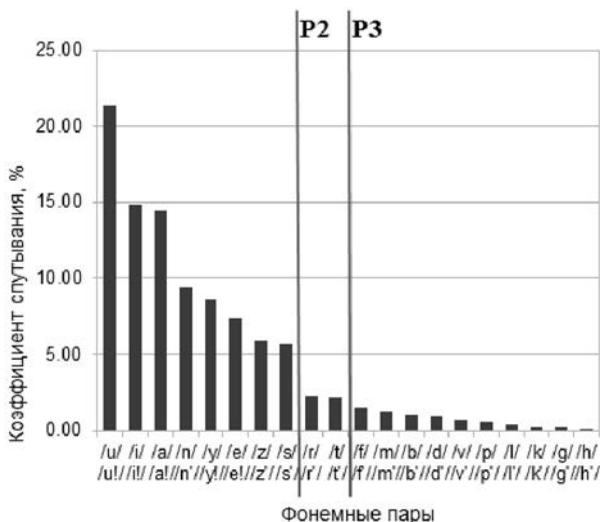


Рис. 3. Коэффициент спутывания для всех пар-кандидатов

Также из рисунка 3 видно, что пары согласных /n/-n'/, /z/-z'/, /s/-s'/, /r/-r'/ и /t/-t'/ имеют более высокое значение процента спутывания, чем другие. Кроме того, различие в значении процента спутывания между парами /s/-s'/ и /r/-r'/ достаточно большое. Поэтому были сформированы три дополнительные фонемные наборы представлены в таблице 4.

Таблица 4. Описание созданных наборов фонем

Фонемный набор	Число фонологических единиц	Описание
P0	47	См. таблицу 3
P1	42	P0 без /a'/, /e'/, /i'/, /u'/, /y'/
P2	39	P1 без /n'/, /s'/, /z'/
P3	37	P2 без /r'/, /t'/
P4	27	P3 без /b'/, /g'/, /d'/, /k'/, /l'/, /m'/, /p'/, /f'/, /h'/

Набор P2 сформирован путем объединения пар /n/-n'/, /z/-z'/, /r/-r'/ и набор P3, дополнительно объединяющий пары /s/-s'/ и /t/-t'/.

Также был создан набор P4, в котором были объединены все пары-кандидаты.

**5.3. Оценивание эффективности применения фонемных наборов.** Кроме системы распознавания, использующей фонемный набор P0, было создано еще 4 системы с фонемными наборами P1-P4. Точность распознавания фонем, полученная с использованием моделей монофонов, представлена в верхней части таблицы 5. В качестве модели языка использовалась фонемная биграммная модель, обученная на фонемных транскрипциях обеих баз данных. Тестовые части аудио-корпусов были также объединены в этом эксперименте. Наилучшая точность распознавания была достигнута при использовании набора P2. Большое различие в точности распознавания, полученной при использовании набора P0 и остальных наборов, объясняется отсутствием в наборах P1-P4 наиболее путающихся пар гласных. Низкая точность, полученная с набором P4, вызвана тем, что слишком много согласных было объединено, что привело к уменьшению разрешения пространства фонологических единиц. Результаты экспериментов показали, что пары-кандидаты были выбраны правильно, и их объединение дает наилучшие результаты.

В нижней части таблицы 5 приведены результаты экспериментов, полученные с применением трифонных моделей. Биграммная модель языка была обучена на транскрипциях как обучающих, так и тестовых частей речевых корпусов, таким образом, в данном случае используется модель языка с закрытым словарем. Однако обучающие и тестовые части корпуса акустически различны. Результаты распозна-

вания, полученные с применением различных фонемных наборов, приблизительно одинаковые, но при использовании набора P0 точность оказалась чуть выше. Высокая точность распознавания, полученная для базы данных СПИИРАН, связана с тем, что все дикторы произносили одни и те же фразы, то есть лексическое содержание фраз было одинаковым, а также размер словаря значительно меньше, чем в речевом корпусе GlobalPhone.

Таблица 5. Точность распознавания речи, полученная при использовании различных фонемных наборов

Тестовые корпуса/ Фонемные наборы	P0	P1	P2	P3	P4
Точность распознавания фонем, %					
СПИИРАН + GlobalPhone	48,92	52,43	53,20	53,06	52,93
Точность распознавания слов, %					
СПИИРАН	96,64	96,55	96,41	96,62	96,40
GlobalPhone	80,92	80,35	80,67	79,82	78,52

Количество контекстно-зависимых трифонов для каждого фонемного набора представлено в таблице 6, где также указано количество связанных состояний соответствующей акустической модели.

Таблица 6. Количество трифонов и состояний для различных фонемных наборов

Набор фонем	Количество трифонов	Число состояний
P0	112849	5342
P1	81314	5356
P2	65562	5359
P3	56279	5335
P4	22709	5342

Поскольку различие в точности распознавания слов, полученной при применении различных фонемных наборов, для речевой базы данных СПИИРАН оказалось достаточно небольшим (менее 1% в абсолютных значениях), были проведены дополнительные эксперименты по определению скорости распознавания речи. В качестве показателя скорости распознавания использовался показатель реального времени (real time factor - RTF). На рисунке 4 представлен график зависимости точности распознавания слов от значения RTF. Для значения RTF=1,2 и ниже фонемный набор P3 показал лучшие результаты распознавания, в то время как для больших значений RTF (более медленное распознавание) лучшие результаты показал набор P0.

Высокая точность распознавания слов объясняется тем, что применялась закрытая модель языка с малым размером словаря. Для создания открытой модели языка со сверхбольшим размером словаря использовалась статистическая биграммная модель, описанная в работе [29]. Результаты, полученные на тестовой части корпуса GlobalPhone, представлены в таблице 7. Хотя точность распознавания существенно ниже, полученные результаты показывают такую же тенденцию, как и предыдущие эксперименты.

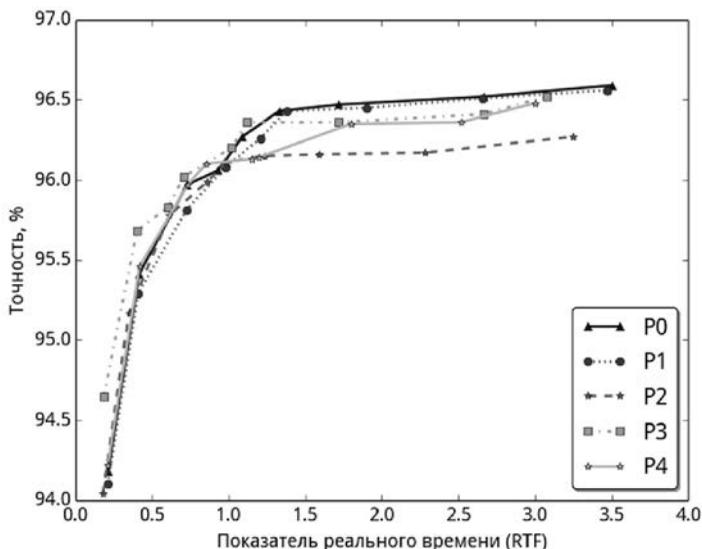


Рис. 4. Точность распознавания слов в зависимости от значения RTF, полученная с использованием речевого корпуса СПИИРАН

Таблица 7. Точность распознавания речи, полученная при использовании модели языка с открытым словарем

Фонемный набор	P0	P1	P2	P3	P4
	67,07	66,78	66,53	65,93	65,28

Таким образом для дальнейших экспериментов по распознаванию речи со сверхбольшим словарем с трифонными акустическими моделями был выбран набор P0, состоящий из 47 фонологических единиц.

**5.4. Результаты экспериментов по распознаванию слитной русской речи со сверхбольшим словарем.** Во всех предыдущих экспериментах использовались стандартные биграммные языковые модели.

Нами была также создана синтактико-статистическая модель языка на основе базовой биграммной модели, полученной в результате автоматического статистического анализа обучающего текстового корпуса, собранного с интернет-сайтов электронных газет. Базовая биграммная модель была расширена за счет выполнения синтаксического анализа обучающего текстового корпуса, в ходе которого выявляются грамматически связанные пары слов, разделенные в тексте другими словами. Таким образом, синтаксический анализ позволяет учесть дальнедействующие грамматические связи между словами. Затем была проведена линейная интерполяция статистической триграммной модели с синтаксическо-статистической моделью с коэффициентом интерполяции 0,27. Размер словаря составил 204 тыс. слов, относительное количество внесловарных слов для тестовых данных составило 2,5 %. При создании модели языка использовалось сглаживание Несер-Нея (Kneser-Neu). Подробно процесс создания синтаксическо-статистической модели описан в [29, 30].

Для экспериментов с модифицированной моделью языка использовался декодер речи Julius 4.2 [31]. Для настроек параметров системы были использованы записи 4 дикторов из тестовой части GlobalPhone общей длительностью 40 минут. Записи остальных дикторов были объединены с дополнительной частью базы данных СПИИРАН в тестовый набор, включающий в себя записи 8 дикторов (4 мужчин и 4 женщин) общей длительностью 1,5 часа.

Результаты распознавания слов и графем (под графемами понимаются буквы и знак пробела) представлены в таблице 8, в которой также даны значения коэффициента неопределенности, относительно количества совпадений  $n$ -грамм (т.е. количества  $n$ -грамм в тестовых данных, которые присутствуют в модели языка) и относительного количества внесловарных слов для тестовых данных.

Таблица 8. Результаты распознавания слитной русской речи со сверхбольшим словарем

Коэффициент неопределенности модели языка	Относительное кол-во совпадений $n$ -грамм, %	Точность распознавания слов, %	Точность распознавания графем, %
549	39,2 (+44,0 для 2-грамм)	73,1	91,5

Скорость распознавания речи на компьютере с многоядерным процессором с тактовой частотой 1,8 ГГц составила около 2,0 RTF.

**6. Заключение.** В статье представлен выбор оптимального фонемного набора для системы распознавания слитной русской речи. Для процесса акустического моделирования было разработано новое фоне-

тическое дерево решений для создания трифонных моделей, учитывающее специфику фонологии русского языка. Была предложена комбинированная методика выбора оптимального фонемного набора, в которой используются как фонетические знания, так и статистическая информация. Эксперименты показали, что разница в точности распознавания слов системами с разными фонемными наборами незначительна, кроме системы с набором P4, в котором все пары-кандидаты были объединены. В оффлайн системах распознавания речи, где точность является важнейшим фактором, может быть рекомендовано использование фонемного набора P0, поскольку он позволяет достичь наибольшей точности. В системах распознавания реального времени, в которых необходим компромисс между скоростью и точностью распознавания, может быть рекомендовано применение набора P3, поскольку точность распознавания соответствующей системы при  $RTF \approx 1$  была выше.

В ходе экспериментов по распознаванию речи с применением синтактико-статистической языковой модели со сверхбольшим словарем, где использовался фонемный набор P0, состоящий из 47 фонем, наибольшая точность распознавания слов русской речи составила 73,1%. Полученные результаты соответствуют качеству распознавания слитной русской речи, полученному на настоящий момент в мире другими организациями.

### Литература

1. *Whittaker E.W.D.* Statistical language modelling for automatic speech recognition of Russian and English // Ph.D. thesis. Cambridge Univ. 2000. 140 p.
2. *Karpov A., Kipyatkova I., Ronzhin A.* Speech Recognition for East Slavic Languages: The Case of Russian // Proceedings of 3rd International Workshop on Spoken Languages Technologies for Under-resourced Languages SLTU'2012. Cape Town. RSA. 2012. pp. 84–89.
3. *Кипяткова И.С., Карпов А.А.* Аналитический обзор систем распознавания русской речи с большим словарем // Труды СПИИРАН. 2010. №1(12). С.7–20.
4. *Kipyatkova I., Karpov A.* Study of Morphological Factors of Factored Language Models for Russian ASR // Edited by A. Ronzhin et al. // SPECOM 2014. Springer International Publishing Switzerland. LNAI 8773. 2014. pp. 451–458.
5. *Vazhenina D., Kipyatkova I., Markov K., Karpov A.* State-of-the-art Speech Recognition Technologies for Russian Language // Proceedings of the Joint International Conference on Human-Centered Computer Environments HCCE-2012. Aizu-Wakamatsu. Japan. 2012. pp. 59–63.
6. *Titov Y., Kilgour K., Stüker S., Waibel A.* The 2011 kit quaero speech-to-text system for the Russian language // Proceedings of the 14th International Conference “Speech and Computer” (SPECOM'2011). 2011. pp. 136–143.
7. *Shin E., Stüker S., Kilgour K., Fügen C., Waibel A.* Maximum Entropy Language Modeling for Russian ASR // Proceedings of the International Workshop for Spoken Language Translation (IWSLT 2013). 2013.

8. *Zablotskiy S., Shvets A., Sidorov M., Semenkin E., Minker W.* Speech and Language Resources for LVCSR of Russia // Proceedings of LREC'2012. Istanbul. Turkey. 2012. pp. 3374–3377.
9. *Zulkarneev M., Satunovsky P., Shamraev N.* The use of d-gram language model for speech recognition in Russian // SPECOM 2013. Springer LNAI 8113. 2013. pp. 362–366.
10. *Zulkarneev M., Grigoryan R., Shamraev N.* Acoustic modeling with deep belief networks for Russian Speech // SPECOM 2013. Springer LNAI 8113. 2013. pp. 17–23.
11. *Schalkwyk J., Beeferman D., Beaufays F., Byrne B., Chelba C., Cohen M., Kamvar M., Strope B.* Google Search by Voice: A Case Study // Advances in Speech Recognition: Mobile Environments, Call Centers and Clinics. 2010. pp. 61–90.
12. SpeechKit API. URL: <http://api.yandex.ru/speechkit/> (дата обращения: 25.04.2014).
13. *Levin K., Ponomareva I., Bulusheva A., Chernykh G., Medennikov I., Merkin N., Prudnikov A., Tomashenko N.* Automated closed captioning for Russian live broadcasting // Proceedings of Interspeech'2014. 2014. pp. 1438–1442.
14. *Young S., Odell J., Woodland P.* Tree-based state tying for high accuracy acoustic modeling // Proceedings of Int. Workshop on Human Language Technology HLT'1994. Stroudsburg. PA. USA. 1994. pp. 307–312.
15. *Young S. et al.* The HTK book // Cambridge Univ. Press. 2009. 384 p.
16. *Odell J.* The use of context in large vocabulary speech recognition // Ph.D. thesis. Cambridge Univ., 1995.
17. *Cubberley P.* Russian: a linguistic introduction // Cambridge University Press. 2002.
18. *Schultz T., Waibel A.* Development of Multilingual acoustic models in the GlobalPhone project // Proceedings of TSD'1998. Brno. Czech Republic. 1998. pp. 311–316.
19. *Psutka J., Ircing P., Psutka J.V., Hajic J., Byrne W.J., Mirovsky J.* Automatic transcription of Czech, Russian, and Slovak spontaneous speech in the MALACH project // Proceedings of Interspeech'2005. Lisbon. Portugal. 2005. pp. 1349–1352.
20. *Tatarnikova M., Tampel I., Oparin I., Khokhlov Y.* Building acoustic models for a large vocabulary continuous speech recognizer for Russian // Proceedings of SPECOM'2006. St. Petersburg, Russia. 2006. pp. 83–87.
21. *Stüker S., Schultz T.* A grapheme based speech recognition system for Russian // Proceedings of International Conference SPECOM'2004. St. Petersburg, Russia. 2004. pp. 297–303.
22. *Ronzhin A., Karpov A.* Russian Voice Interface // Pattern Recognition and Image Analysis. 2007. vol. 17. no. 2. pp. 321–336.
23. *Zhang J.S., Hu X.H., Nakamura S.* Using mutual information criterion to design an efficient phoneme set for Chinese speech recognition // IEICE Transactions on Information and Systems. 2008. vol. E91-D. no. 3. 2008. pp. 508–513.
24. *Singh R., Raj B., Stern R.* Automatic generation of subword units for speech recognition systems // IEEE Transactions on Acoustics, Speech and Signal Processing. 2002. vol. 10(2). pp. 89–99.
25. *Шведова Н.Ю. (гл. ред.) и др.* Русская грамматика // М.: Наука. 1980. 783 с.
26. *Кияткова И.С., Карпов А.А.* Модуль фонематического транскрибирования для системы распознавания разговорной русской речи // Искусственный интеллект, Донецк. Украина. 2008. № 4. 2008. С. 747–757.
27. *Jokisch O., Wagner A., Sabo R., Jaeckel R., Cylwik N., Rusko M., Ronzhin A., Hoffmann R.* Multilingual speech data collection for the assessment of pronunciation and prosody in a language learning system // Proceedings of SPECOM'2009. St. Petersburg, Russia. 2009. pp. 515–520.
28. *Vazhenina D., Markov K.* Phoneme set selection for Russian speech recognition // Proceedings of Int. Conf. on Natural Language Processing and Knowledge Engineering NLP-KE. Tokushima. Japan. 2011. pp. 475–478.

29. *Karpov A., Markov K., Kipyatkova I., Vazhenina D., Ronzhin A.* Large vocabulary Russian speech recognition using syntactico-statistical language modeling // *Speech Communication*. 2014. vol. 56. pp. 213–228.
30. *Кипяткова И.С.* Программно-алгоритмическое обеспечение создания синтаксическо-статистической модели русского языка по текстовому корпусу // *Труды СПИИРАН*. 2013. Вып. 24. С. 332–348.
31. *Lee A., Kawahara T.* Recent Development of Open-Source Speech Recognition Engine Julius // *Proceedings of Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC 2009)*. Sapporo. Japan. 2009. pp.131–137.

## References

1. Whittaker E.W.D. Statistical language modelling for automatic speech recognition of Russian and English. Ph.D. thesis. Cambridge Univ. 2000. 140 p.
2. Karpov A., Kipyatkova I., Ronzhin A. Speech Recognition for East Slavic Languages: The Case of Russian. *Proceedings of 3rd International Workshop on Spoken Languages Technologies for Under-resourced Languages SLTU'2012*. Cape Town. RSA. 2012. pp. 84–89.
3. Kipyatkova I.S., Karpov A.A. [An Analytical Survey of Large Vocabulary Russian Speech Recognition Systems]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2010. vol. 12. pp.7–20. (In Russ.).
4. Kipyatkova I., Karpov A. Study of Morphological Factors of Factored Language Models for Russian ASR. Edited by A. Ronzhin et al. *SPECOM 2014*. Springer International Publishing Switzerland. LNAI 8773. 2014. pp. 451–458.
5. Vazhenina D., Kipyatkova I., Markov K., Karpov A. State-of-the-art Speech Recognition Technologies for Russian Language. *Proceedings of the Joint International Conference on Human-Centered Computer Environments (HCCE-2012)*. Aizu-Wakamatsu. Japan. 2012. pp. 59–63.
6. Titov Y., Kilgour K., Stüker S., Waibel A. The 2011 kit quero speech-to-text system for the Russian language. *Proceedings of the 14th International Conference “Speech and Computer” (SPECOM'2011)*. 2011. pp. 136–143.
7. Shin E., Stüker S., Kilgour K., Fügen C., Waibel A. Maximum Entropy Language Modeling for Russian ASR. *Proceedings of the International Workshop for Spoken Language Translation (IWSLT 2013)*. 2013.
8. Zablotskiy S., Shvets A., Sidorov M., Semenko E., Minker W. Speech and Language Resources for LVCSR of Russia. *Proceedings of LREC'2012*. Istanbul. Turkey. 2012. pp. 3374–3377.
9. Zulkarneev M., Satunovsky P., Shamraev N. The use of d-gram language model for speech recognition in Russian. *SPECOM 2013*. Springer LNAI 8113. 2013. pp. 362–366.
10. Zulkarneev M., Grigoryan R., Shamraev N. Acoustic modeling with deep belief networks for Russian Speech. *SPECOM 2013*. Springer LNAI 8113. 2013. pp. 17–23.
11. Schalkwyk J., Beeferman D., Beaufays F., Byrne B., Chelba C., Cohen M., Kamvar M., Strope B. Google Search by Voice: A Case Study. *Advances in Speech Recognition: Mobile Environments, Call Centers and Clinics*. 2010. pp. 61–90.
12. SpeechKit API. Available at: <http://api.yandex.ru/speechkit/> (accessed 25.04.2014).
13. Levin K., Ponomareva I., Bulusheva A., Chernykh G., Medennikov I., Merkin N., Prudnikov A., Tomashenko N. Automated closed captioning for Russian live broadcasting. *Proceedings of Interspeech'2014*. 2014. pp. 1438–1442.
14. Young S., Odell J., Woodland P. Tree-based state tying for high accuracy acoustic modelling. *Proceedings of Int. Workshop on Human Language Technology (HLT'1994)*. Stroudsburg. PA. USA. 1994. pp. 307–312.
15. Young S. et al. *The HTK book*. Cambridge Univ. Press. 2009. 384 p.

16. Odell J. The use of context in large vocabulary speech recognition. Ph.D. thesis. Cambridge Univ. 1995.
17. Cubberley P. Russian: a linguistic introduction. Cambridge University Press. 2002.
18. Schultz T., Waibel A. Development of Multilingual acoustic models in the GlobalPhone project. Proceedings of TSD'1998. Brno. Czech Republic. 1998. pp. 311–316.
19. Psutka J., Ircing P., Psutka J.V., Hajic J., Byrne W.J., Mirovsky J. Automatic transcription of Czech, Russian, and Slovak spontaneous speech in the MALACH project. Proceedings of Interspeech'2005. Lisbon. Portugal. 2005. pp. 1349–1352.
20. Tatarnikova M., Tampel I., Oparin I., Khokhlov Y. Building acoustic models for a large vocabulary continuous speech recognizer for Russian. Proceedings of SPECOM'2006. St. Petersburg, Russia. 2006. pp. 83–87.
21. Stüker S., Schultz T. A grapheme based speech recognition system for Russian. Proceedings of International Conference SPECOM'2004. St. Petersburg, Russia. 2004. pp. 297–303.
22. Ronzhin A., Karpov A. Russian Voice Interface. Pattern Recognition and Image Analysis. 2007. vol. 17. no. 2. pp. 321–336.
23. Zhang J.S., Hu X.H., Nakamura S. Using mutual information criterion to design an efficient phoneme set for Chinese speech recognition. IEICE Transactions on Information and Systems. 2008. vol. E91-D. no. 3. pp. 508–513.
24. Singh R., Raj B., Stern R. Automatic generation of subword units for speech recognition systems. IEEE Transactions on Acoustics, Speech and Signal Processing. 2002. vol. 10(2). pp. 89–99.
25. Shvedova N.Yu. et. al. Russkaya grammatika [Russian Grammar]. M.: Nauka, 1980. 783 p. (In Russ.)
26. Kipyatkova I.S., Karpov A.A. [The module of phonemic transcription for conversational Russian speech recognition system]. *Iskusstvenny intellekt – Artificial intelligence*. Donetsk. Ukraine. 2008. vol. 4., pp. 747–757. (In Russ.)
27. Jokisch O., Wagner A., Sabo R., Jaeckel R., Cylwik N., Rusko M., Ronzhin A., Hoffmann R. Multilingual speech data collection for the assessment of pronunciation and prosody in a language learning system. Proceedings of SPECOM'2009. St. Petersburg, Russia. 2009. pp. 515–520.
28. Vazhenina D., Markov K. Phoneme set selection for Russian speech recognition. Proceedings of Int. Conf. on Natural Language Processing and Knowledge Engineering NLP-KE. Tokushima. Japan. 2011. pp. 475–478.
29. Karpov A., Markov K., Kipyatkova I., Vazhenina D., Ronzhin A. Large vocabulary Russian speech recognition using syntactico-statistical language modeling. Speech Communication. 2014. vol. 56. pp. 213–228.
30. Kipyatkova I.S. [Software for Creation of Sintactico-Statistical Russian Language Model Based on the Text Corpus]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2013. vol. 24. pp.332–348. (In Russ.)
31. Lee A., Kawahara T. Recent Development of Open-Source Speech Recognition Engine Julius. Proceedings of Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC 2009). Sapporo. Japan. 2009. pp.131–137.

**Важенина Дарья Анатольевна** — магистр техн. наук, аспирант лаборатории пользовательских интерфейсов отделения информационных систем Университета Аизу, Япония. Область научных интересов: автоматическое распознавание речи, статистические языковые модели, машинное обучение. Число научных публикаций — 8. d8132102@u-aizu.ac.jp. Университет Аизу, Тсуруга, Икки-мачи, г.Аизу-Вакаматсу, преф.Фукушима, 965-8580, Япония; р.т.+81(242)372762, факс: +81(242)372549.

**Vazhenina Daria Anatolievna** — M.Sci., Ph.D. student, Human Interface Laboratory of the Information Systems Division, University of Aizu, Japan. Research interests: automatic speech recognition, statistical language modeling, machine learning. The number of publications — 8. d8132102@u-aizu.ac.jp; The University of Aizu, Tsuruga, Ikki-machi, Aizu-Wakamatsu-shi, Fukushima-ken, 965-8580, Japan; fax: +81(242)372549.

**Кипяткова Ирина Сергеевна** — к-т техн. наук, старший научный сотрудник лаборатории речевых и многомодальных интерфейсов Федерального государственного бюджетного учреждения науки Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: автоматическое распознавание речи, статистические модели языка. Число научных публикаций — 50. kipyatkova@iias.spb.su; СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-7081, факс +7(812)328-7081.

**Kipyatkova Irina Sergeevna** — Ph.D., senior researcher, Laboratory of Speech and Multimodal Interfaces St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS). Research interests: automatic speech recognition, statistical language models. The number of publications — 50. kipyatkova@iias.spb.su; SPIIRAS, 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)328-7081, fax +7(812)328-7081.

**Марков Константин** — к-т техн. наук, доцент лаборатории пользовательских интерфейсов отделения информационных систем Университета Аизу, Япония. Область научных интересов: обработка аудио сигналов, Байесовские статистические модели, машинное обучение, распознавание образов. Число научных публикаций — 100. markov@u-aizu.ac.jp; Университет Аизу, Тсуруга, Икки-мачи, г.Аизу-Вакаматсу, преф.Фукушима, 965-8580, Япония; р.т.+81(242)372762, факс: +81(242)372549.

**Markov Konstantin** — Ph.D., associate professor, Human Interface Laboratory of the Information Systems Division, University of Aizu, Japan. Research interests: audio signal processing, Bayesian statistical modeling, machine learning and pattern recognition. The number of publications — 100. markov@u-aizu.ac.jp; The University of Aizu, Tsuruga, Ikki-machi, Aizu-Wakamatsu-shi, Fukushima-ken, 965-8580, Japan; office phone +81(242)372762, fax: +81(242)372549.

**Карпов Алексей Анатольевич** — д-р техн. наук, доцент, ведущий научный сотрудник лаборатории речевых и многомодальных интерфейсов Федерального государственного бюджетного учреждения науки Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: автоматическое распознавание речи, многомодальные интерфейсы, аудиовизуальное распознавание речи. Число научных публикаций — 190. karpov@iias.spb.su; СПИИРАН, 14-я линия В.О., д. 39, Санкт-Петербург, 199178, РФ; р.т. +7(812)328-7081, факс +7(812)328-7081.

**Karpov Alexey Anatolyevich** — Ph.D., Dr. Sci., associate professor, leading researcher, Laboratory of Speech and Multimodal Interfaces St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS). Research interests: automatic speech recognition, multimodal interfaces, audio-visual speech recognition. The number of publications — 190. karpov@iias.spb.su; SPIIRAS, 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)328-7081, fax +7(812)328-7081.

**Поддержка исследований.** Работа выполнена при финансовой поддержке фонда РФФИ (проект № 12-08-01265-а).

**Acknowledgements.** This research is supported by RFBR (project № 12-08-01265-а).

## РЕФЕРАТ

### *Важенина Д.А., Кипяткова И.С., Марков К., Карпов А.А.* **Методика выбора фонемного набора для автоматического распознавания русской речи.**

Статья посвящена обучению акустических моделей и выбору оптимального фонемного алфавита. Для создания моделей акустических единиц было разработано фонетическое дерево решений, учитывающее специфику фонологии русского языка. Для выбора оптимального фонемного набора предложен комбинированный метод, который объединяет информацию, полученную из фонологических знаний, и статистические данные из матрицы спутывания монофонов. Матрица спутывания позволяет определить наиболее часто несовпадающие монофоны. Для выбора наилучшего фонемного набора вначале использовался наибольший фонемный набор, затем его размер постепенно уменьшался путем удаления или объединения некоторых фонологических единиц. Применяя данный метод к русскому варианту ПРА/МФА, фонемный набор был уменьшен до 47 фонологических единиц и был использован как исходный. За счет применения матрицы спутывания были созданы 4 набора с разным количеством фонологических единиц: 42, 39, 37, 27. Система распознавания речи была обучена с использованием каждого фонемного набора. Эксперименты по распознаванию русской речи, которые проводились на речевых корпусах СПИИРАН и GlobalPhone, показали, что использование сокращенных фонемных наборов позволяет увеличить точность распознавания фонем и практически не влияет на точность распознавания слов. При этом использование сокращенного фонемного набора позволяет повысить скорость распознавания речи, что важно для систем, работающих в реальном времени.

Для проведения экспериментов по распознаванию русской речи со сверхбольшим словарем был выбран набор, состоящий из 47 фонологических единиц. Распознавание речи проводилось с использованием синтаксическо-статистической модели языка, которая была создана путем добавления к базовой биграммной модели синтаксически связанных пар слов, которые были разделены в обучающем тексте другими словами. Размер словаря составил 204 тыс. слов. В результате экспериментов точность распознавания слов составила 73,1%.

## SUMMARY

### *Vazhenina D.A., Kipyatkova I.S., Markov K., Karpov A.A.* **Technique for phoneme set selection for automatic Russian speech recognition.**

The paper is devoted to training the acoustic models and selection of the optimal phoneme alphabet. The phonetic tree that takes into account specifics of the Russian phonology was developed to create context-dependent acoustic models. A combined method, which utilizes information obtained from phonological knowledge and statistical data from the confusion matrix of monophones, was proposed for selection of the optimal phoneme set. Confusion matrix allows to determine the most frequently mismatched monophones. For the best phoneme set selection, the largest phoneme set was used and then its size was gradually decreases by deleting or merging some phonemes. Applying this approach to the IPA Russian phonetic set, we first reduced it to 47 phonological units, which were used as initial set. Based on the phoneme confusion results 4 sets with the number of phonological units of 42, 39, 27, 27 were created. The speech recognition systems were trained separately using each phoneme set. Experiments on Russian speech recognition conducted on SPIIRAS and GlobalPhone speech corpora showed that usage of reduced phoneme sets allows to increase phoneme recognition accuracy and almost has no influence on word recognition accuracy. At the same time using the reduced phoneme set allows to increase decoding speed what is important for real-time systems.

For experiments concerning very large vocabulary Russian speech recognition the set consisting of 47 phonological units was selected. Speech recognition was conducted using the syntactico-statistical language model. This model was created by adding grammatically-connected word pairs, which were separated by other words in the training corpus, to the baseline bigram model. Vocabulary size was 204 K words. Word recognition accuracy of 73.1 % was achieved.

К.В. ЕВГРАФОВА, Т.В. КАЧКОВСКАЯ  
**ДИФФЕРЕНЦИАЛЬНЫЕ ПОРОГИ ВОСПРИЯТИЯ  
ДЛИТЕЛЬНОСТИ ЯЗЫКОВЫХ ЕДИНИЦ В ЗАВИСИМОСТИ  
ОТ ФОНЕТИЧЕСКОГО КОНТЕКСТА**

---

*Евграфова К.В., Качковская Т.В. Дифференциальные пороги восприятия длительности языковых единиц в зависимости от фонетического контекста.*

**Аннотация.** В данной статье описаны результаты исследования «Перцептивная шкала длительностей фонетических единиц языка». В ходе исследования были проведены два психоакустических эксперимента, в результате которых был установлен порог восприятия длительности ударных гласных. В среднем он составляет 2–4 периода основного тона, что приблизительно соответствует 20–40 мс. Показано, что величина порога не зависит от количества слогов в слове и места ударения. В разделах статьи представлены: краткий обзор литературных данных по проблеме, а также описание психоакустических экспериментов, целью которых было установить слуховые дифференциальные пороги длительности для русских гласных с учетом фонетического контекста, статистический анализ результатов эксперимента, а также выводы по полученным данным.

**Ключевые слова:** восприятие речи, акустика, дифференциальные пороги, длительность фонетических звуков.

*Evgrafova K.V., Kachkovskaya T.V. Differential Thresholds of Perception of Sound Duration in Different Phonetic Contexts.*

**Abstract.** The paper presents the results of the study which was aimed at obtaining the values of differential thresholds for Russian stressed vowels. Two types of psychoacoustic experiments were conducted. The experimental material was formed taking into account the segmental and prosodic context. The segmental context of the vowel stimuli was relatively similar while prosodic context varied. Each vowel was presented in several word types which involved one-syllable words and three-syllable words with different stress patterns. Each word was embedded into an affirmative utterance (initial/medial/final position), an interrogative utterance (final position), and an utterance with a contrastive stress on the word. The subject listened to a sample word containing a target vowel and its modifications with increased or decreased vowel duration. The gradual increment or decrement was 1 F0 period or 10 ms on average. The task was to click the modification when the duration difference was perceived. Then the differential thresholds are calculated by means of statistical analysis. The average differential threshold was found to be 2–4 F0 periods which is approximately about 20–40 ms. The threshold value does not seem to be influenced by word duration (in syllables) or stress pattern.

**Keywords:** speech perception, acoustics, differential thresholds, phonetic segment duration.

---

**1. Введение.** В данной статье представлены результаты исследования «Перцептивная шкала длительностей фонетических единиц языка». Целью данного исследования являлось изучение параметра длительности звуковых сегментов высказывания с учетом артикуляционно-акустических и акустико-слуховых отношений. Длительность фонетических единиц (в особенности гласных) играет важную роль в формировании временной организации высказывания и при реализации словесного ударения в русском языке. Изучение длительности с

точки зрения восприятия необходимо для установления границ зон акустического варьирования этого параметра. Изменение величины этого параметра (уменьшение или увеличение) в границах такой зоны должно быть неощутимо на слух, а выход за пределы зоны должен приводить к скачкообразным изменениям. Установление этих границ основывается на данных о возможностях слуховой системы, а также данных, как по абсолютным, так и по дифференциальным слуховым порогам. Установление слуховых порогов изменения длительности звуков необходимо для точного моделирования временной организации высказывания при разработке новых систем синтеза и распознавания речи, а также улучшения качества уже существующих.

В следующих разделах статьи представлены: краткий обзор литературных данных по проблеме, описание психоакустических экспериментов, целью которых было установить слуховые дифференциальные пороги длительности для русских гласных с учетом фонетического контекста, статистический анализ результатов эксперимента, а также выводы по полученным данным.

**2. Краткий обзор литературных данных по проблемам восприятия параметра длительности и методов анализа.** На данный момент существует многочисленный ряд исследований механизмов восприятия речи в целом, а также временной структуры отдельного высказывания и его элементов в частности [2–4, 6, 7, 11–14]. Тем не менее, многие положения относительно восприятия человеком единиц речевой последовательности имеют дискуссионный характер. Отмечается недостаточность теоретической и экспериментальной разработки проблемы, а также то, что «используемые экспериментальные методики не позволяют переносить результаты лабораторных исследований в модели механизмов естественного речевого общения» [3, с. 201].

Способность слуховой системы различать малые изменения во временной структуре сигнала является в настоящее время предметом многочисленных исследований. Однако, данные, о чувствительности слуха к различиям в длительности разных звуковых объектов, которые приводят авторы различных психоакустических исследований, не совпадают. По мнению Клатта, только изменения порядка 20 % и более могут быть значимы для восприятия речи [11]. И. Лехисте в обзоре, посвященном дифференциальной чувствительности слуха по длительности [13], отмечает, что в диапазоне длительностей, типичном для ударных гласных в связной речи (60–200 мс), пороговое значение, равное 20 %, соответствует признанию того, что минимальные различия, обнаруживаемые на слух, составляют 12–40 мс.

Различия в результатах, по-видимому, связано с тем, что слуховая чувствительность к различиям звуковых единиц по длительности определяется условиями и методикой эксперимента. Кроме того, важными факторами также являются: физическая природа предъявляемых сигналов (естественные или синтезированные), набор исследуемых длительностей, частота повторения стимулов в наборе и т. д. Как отмечено в [4, с. 166], «получаемые в опыте величины порогов отражают одновременно как свойства сенсорной системы, так и особенности воспринимаемого набора сигналов».

Как правило, распространенные методики построения психоакустических экспериментов основываются на так называемых классических методах психофизики, используемым для определения абсолютных и разностных порогов.

Порог определяется как барьер, разделяющий ощущаемые раздражители от неощущаемых. Под абсолютным порогом понимается величина раздражителя, дающего 50 % случаев обнаружения. Аналогично этому дифференциальный порог — это минимальное изменение раздражения, обнаруживаемое в 50 % случаев. Ниже будет приведено описание только одного из них — метода границ или минимальных изменений, так как именно он был использован как основа для построения методики экспериментов в данном исследовании.

При использовании этого метода границ производится изменение интенсивности раздражителя в какую-либо сторону (увеличение — восходящая стимуляция или уменьшение — нисходящая стимуляция). Таким образом, устанавливается барьер, или границу, где начинает ощущаться различие (для дифференциальных порогов). В случае определения порогов путем увеличения переменного стимула они носят название инкрементных порогов, в случае уменьшения переменного стимула — декрементных. Правильность установления пороговых точек гарантируется их многократной регистрацией. Это помогает избежать влияния таких факторов, как приобретение опыта в данной сенсорной деятельности и развитие утомления. Окончательно место пороговой точки определяется как среднее всех произведенных замеров. Разброс получаемых показателей характеризует меру точности в оценке величины порога и носят наименование переменной ошибки. Чем меньше «шаг» изменения стимула, тем выше точность измерения [10].

**3. Материал исследования.** Как было указано выше, помимо методики на характер полученных результатов оказывает большое влияние также и материал исследования.

Для данного исследования было принято решение использовать речевой материал, а не синтезированные стимулы. Набор фонетических контекстов, влияющих на реализацию акустической длительности гласных звуков, был составлен на основе теоретических описаний. Затем в ходе акустического анализа гласных из корпуса CORPRES оценивалось влияние выделенных факторов на изменение объективной длительности гласного. Таким образом, был уточнен набор контекстов, на основе которого был сформирован экспериментальный звуковой материал [5].

Таким образом, в экспериментальном материале каждый из шести русских гласных был представлен под ударением в 4 типах слов: односложное, трехсложное с ударением на первом слоге, трехсложное с ударением на втором слоге, трехсложное с ударением на третьем слоге. Консонантный контекст мог включать в себя любые твердые шумные согласные кроме /v/. Список использованных слов представлен в таблице 1.

Таблица 1. Список слов, использующихся в эксперименте

Односложное слово	Трехсложное слово		
	с ударением на первом слоге	с ударением на втором слоге	с ударением на третьем слоге
<i>сад</i>	<i>заговор</i>	<i>закатов</i>	<i>палисад</i>
<i>ТЭЦ</i>	<i>цезарем</i>	<i>катетер</i>	<i>рикошет</i>
<i>сыч</i>	<i>пыжиков</i>	<i>копытом</i>	<i>коротыш</i>
<i>кот</i>	<i>фокусов</i>	<i>походов</i>	<i>переход</i>
<i>суп</i>	<i>кузовом</i>	<i>сосудов</i>	<i>лилипут</i>
<i>кит</i>	<i>титулов</i>	<i>напиток</i>	<i>гепатит</i>

Эти слова были прочитаны диктором-профессионалом как изолированно, так и в составе утвердительного высказывания, где слово находилось в начальной, конечной позиции и в середине, вопросительного (конечная позиция и одновременно интонационный центр), а также в высказывании, где слово находилось под логическим ударением. Диктор при записи получил задание читать все предложения в одинаковом темпе, максимально единообразно. При этом утвердительные высказывания читались с нейтральным просодическим оформлением, а вопросительные высказывания и высказывания с логическим ударением — с ярким просодическим оформлением. Сформированный материал был использован в двух видах экспериментов.

**4. Метод.** Для проведения эксперимента по определению порогов восприятия длительности гласных рассматриваемые слова были

вырезаны из контекста. Далее выполнялась модификация длительности ударных гласных. Модификации представляли собой естественно звучащие варианты слова-эталона, в которых от одной модификации к другой автоматически была изменена первоначальная длительность гласного (увеличена или сокращена) на один период частоты основного тона, что в среднем составляло 10 мс. Одна часть была направлена на установление инкрементного, а другая — декрементного порога. Таким образом, в ходе эксперимента испытуемый в случайном порядке анализировал стимулы двух видов. Чередование нисходящих и восходящих рядов стимуляции обычно применяется для предотвращения ошибок привыкания и ошибок ожидания (предвосхищения) или, по крайней мере, для их взаимного уравнивания. Ошибки привыкания вызываются психологической инерцией суждений. Продолжить ряд одних и тех же оценок испытуемому обычно легче, чем прервать его, высказав отличное от предыдущих суждений. Ошибки ожидания объясняются тем, что испытуемый на основе полученного в эксперименте опыта понимает, что на каком-либо шаге изменения стимула ощущение различия должно появиться. Ожидание появления этого психологического эффекта с каждым шагом все напряженнее. В результате появляется возможность того, что он сообщит о появлении у него этого ощущения раньше, чем это происходит в действительности.

Длина рядов (число шагов изменения стимула) в эксперименте также варьировалась. Это было обусловлено собственной длительностью гласных. В целом количество модификаций (шагов) зависело от длины каждого конкретного гласного. Модификаций каждого слова-эталона было ровно столько, чтобы последовательно увеличить или уменьшить первоначальную длительность гласного в два раза. При сокращении длительности количество модификаций ограничивалось необходимостью сохранить не менее двух-трех периодов ЧОТ стационарного участка гласного. В связи с этим, в инкрементном наборе стимулов количество модификаций могло достигать 17, а в декрементном наборе — 3.

Для проведения эксперимента была составлена презентация в программе PowerPoint, позволяющая прослушивать эталон и модификации в том порядке, в котором это было удобно испытуемому. Пример слайда представлен на рисунке 1.

В задачи испытуемых входило послушать слово-эталон, обращая внимание на длину ударного гласного в нем. Далее необходимо было прослушать модификацию 1 и сравнить длительность гласного в эталоне и в модификации. Если испытуемый не слышал разницу в длительности, он переходил к сравнению эталона и модификации 2.

Таким образом, он продолжал сравнивать эталон с модификациями по порядку до тех пор, пока длина ударного гласного в эталоне и модификации для него не отличалась на слух. В этом случае он отмечал данную модификацию. Если ни одна из модификаций для него не отличалась от эталона по длительности ударного гласного, он выбирал соответствующую отметку.

Эталон	Модификации
<input type="checkbox"/> кот	<input type="radio"/> Модификация 1
	<input type="radio"/> Модификация 2
	<input checked="" type="radio"/> Модификация 3
	<input type="radio"/> Модификация 4
	<input type="radio"/> Модификация 5
	<input type="radio"/> Модификация 6
	<input checked="" type="checkbox"/> Ни одна из модификаций не отличается от эталона
	<input type="button" value="Submit"/>

Рис. 1. Пример слайда из презентации, использующейся для проведения перцептивных экспериментов

Поскольку такой выбор требует от испытуемого большой концентрации внимания, при построении эксперимента необходимо ограничить объем материала таким образом, чтобы испытуемый не ощущал дискомфорта или утомления. В данном исследовании сессия состояла из 48 стимулов; в середине сессии испытуемому предлагалось сделать перерыв и посмотреть короткий мультипликационный фильм.

**5. Эксперимент 1.** Целью первого эксперимента было установление дифференциального порога восприятия длительности ударных гласных в изолированно произнесенных словах с разным количеством слогов и с разными моделями словесного ударения, но при этом имеющие нейтральное периодическое оформление.

В эксперименте приняло участие 24 испытуемых.

В таблице 2 представлены данные о доле испытуемых, отметивших, что ни одна из модификаций не отличается от эталона, для всех контекстов. Эти данные показывают, что большая часть таких

случаев приходится на те стимулы, в которых длительность ударного гласного последовательно *уменьшалась*.

Таблица 2. Доля испытуемых, ответивших, что ни одна из модификаций не отличается от эталона, для слов, произнесенных изолированно, при увеличении длительности ударного гласного («вверх») и при уменьшении длительности ударного гласного («вниз»)

	вниз	вверх		вниз	Вверх
<i>Сад</i>	4%	4%	<i>кот</i>	0%	4%
<i>заговор</i>	0%	0%	<i>фокусов</i>	17%	4%
<i>закатов</i>	0%	0%	<i>походов</i>	13%	0%
<i>палисад</i>	4%	0%	<i>переход</i>	4%	0%
<i>ТЭЦ</i>	0%	0%	<i>суп</i>	25%	0%
<i>цезарем</i>	8%	0%	<i>кузовом</i>	13%	0%
<i>катетер</i>	13%	4%	<i>сосудов</i>	13%	4%
<i>рикошет</i>	21%	0%	<i>лилпут</i>	21%	4%
<i>Сыч</i>	8%	0%	<i>кит</i>	8%	0%
<i>пыжиков</i>	0%	0%	<i>титулов</i>	8%	13%
<i>копытом</i>	8%	4%	<i>напиток</i>	29%	0%
<i>коротыши</i>	25%	0%	<i>гепатит</i>	13%	4%

В ходе эксперимента 17 из 24 испытуемых выбрали вариант ответа «Ни одна из модификаций не отличается от эталона» хотя бы для одного стимула. Таким образом, дисперсионный анализ с повторными измерениями был возможен только для оставшихся 7 испытуемых. Результаты этого анализ показали, что длина слова и место ударения не влияют на значение порога восприятия длительности. Влияние типа гласного подтвердилось только для односложных слов при последовательном *увеличении* длительности ( $p=0,019$ , однако попарные сравнения не позволили выявить значимых отличий); в остальных случаях тип гласного не влиял на значение порога восприятия.

Следует отметить, что большое число испытуемых (71 %) хотя бы для одного стимула не смогли отличить эталон от модификаций даже при условии, что длительность гласного изменялась *в два раза*. Это говорит о том, что способность к различению длительностей носит индивидуальный характер и, возможно, не является основополагающим фактором при восприятии речи.

Средние значения порогов для всех шести гласных и всех рассмотренных типов слов представлены на рисунке 2. В большинстве случаев значения порога, вычисленные при последовательном *умень-*

шении длительности гласного, оказались ниже, чем при *увеличении*. Статистически это различие было подтверждено только для ударных гласных в словах *суп*, *ТЭЦ* и *коротыш*.

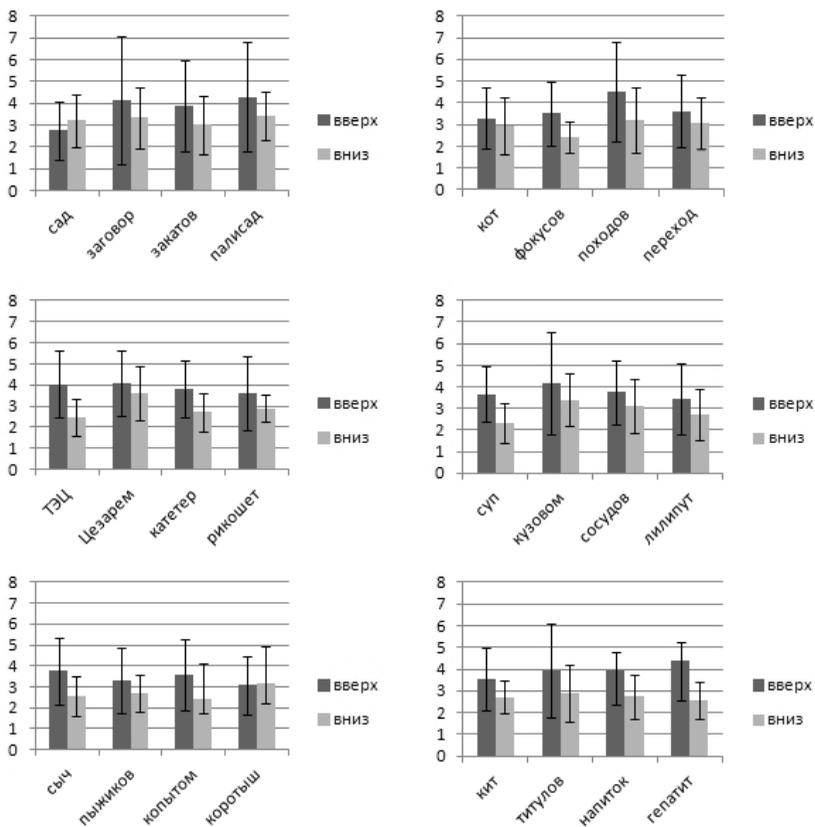


Рис. 2. Средние значения порога восприятия длительности ударного гласного в периодах основного тона при увеличении длительности ударного гласного («вверх») и при уменьшении длительности ударного гласного («вниз») (вертикальные линии соответствуют значениям стандартного отклонения)

В среднем порог восприятия длительности ударных гласных составляет от 2 до 4 периодов основного тона, что приблизительно соответствует 20–40 мс. Такой результат не противоречит имеющимся данным о дифференциальных порогах восприятия длительности.

**6. Эксперимент 2.** Целью второго эксперимента было установление дифференциального порога восприятия длительности ударных

гласных, встретившихся в различных просодических контекстах. В качестве материала для второго эксперимента были выбраны не изолированные слова, а слова, вырезанные из предложений. Поскольку характер эксперимента накладывает ограничения на объем материала, предъявляемого испытуемому, было принято решение использовать стимулы только для одного гласного; в качестве этого гласного был выбран /а/ как наиболее частотный.

В эксперименте приняло участие 20 испытуемых. В таблице 3 представлены данные о доле испытуемых, отметивших, что ни одна из модификаций не отличается от эталона, для всех контекстов. Как и в предыдущем эксперименте, эти данные показывают, что большая часть таких случаев приходится на те стимулы, в которых длительность ударного гласного последовательно *уменьшалась*.

Таблица 3. Доля испытуемых, ответивших, что ни одна из модификаций не отличается от эталона, для слов с ударным гласным /а/, произнесенных в различных контекстах, при увеличении длительности ударного гласного («вверх») и при уменьшении длительности ударного гласного («вниз»)

Изолированное произнесение			На конце утвердительного предложения		
	вниз	вверх		Вниз	вверх
<i>сад</i>	4%	4%	<i>сад</i>	8%	0%
<i>заговор</i>	8%	0%	<i>заговор</i>	0%	0%
<i>закатов</i>	0%	0%	<i>закатов</i>	4%	0%
<i>палисад</i>	0%	0%	<i>палисад</i>	0%	0%
На конце вопросительного предложения			В середине утвердительного предложения		
	вниз	вверх		Вниз	вверх
<i>сад</i>	13%	4%	<i>сад</i>	4%	4%
<i>заговор</i>	17%	0%	<i>заговор</i>	17%	0%
<i>закатов</i>	4%	0%	<i>закатов</i>	25%	4%
<i>палисад</i>	4%	4%	<i>палисад</i>	4%	4%
Под логическим ударением			В начале утвердительного предложения		
	вниз	вверх		Вниз	вверх
<i>сад</i>	4%	0%	<i>сад</i>	8%	0%
<i>заговор</i>	0%	8%	<i>заговор</i>	13%	0%
<i>закатов</i>	0%	0%	<i>закатов</i>	13%	4%
<i>палисад</i>	4%	4%	<i>палисад</i>	0%	0%

Как и в предыдущем эксперименте, большое число испытуемых (16 из 20) выбрали вариант ответа «Ни одна из модификаций не отличается от эталона» хотя бы для одного стимула. Данных для ос-

тавшихся 4 испытуемых было недостаточно для проведения дисперсионного анализа с повторными измерениями.

Средние значения порогов для гласного /а/ в 4 типах слов и шести рассмотренных контекстов представлены на рисунке 3. Как и в предыдущем эксперименте, здесь наблюдается тенденция к более низким значениям порога, вычисленным при последовательном *уменьшении* длительности гласного, чем при *увеличении*.

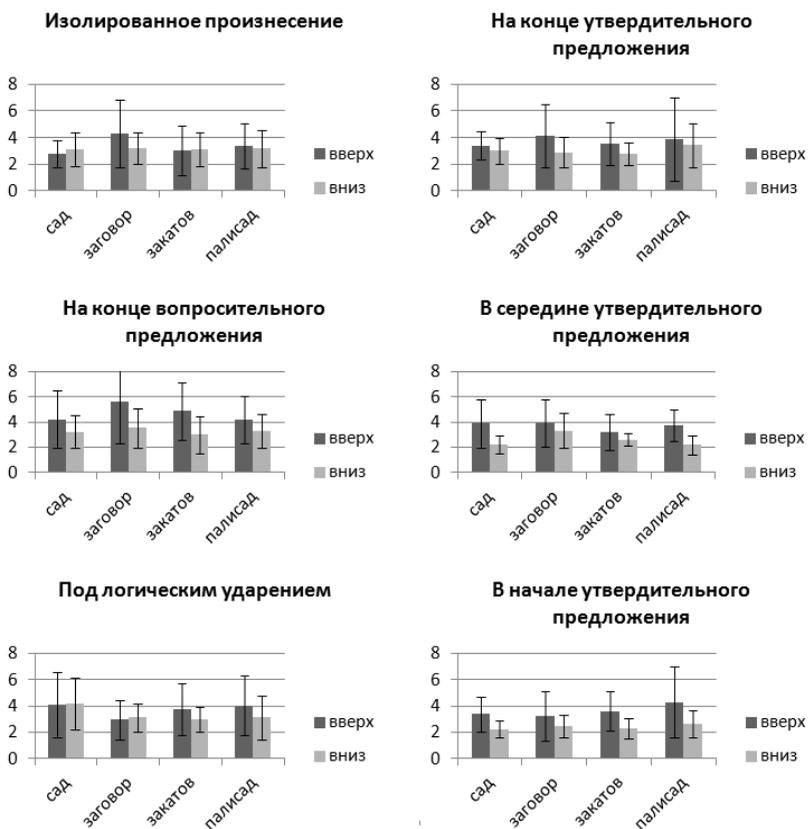


Рис. 3. Средние значения порога восприятия длительности ударного гласного в периодах основного тона при увеличении длительности ударного гласного («вверх») и при уменьшении длительности ударного гласного («вниз») (вертикальные линии соответствуют значениям стандартного отклонения)

Данные, представленные на рисунке 3, позволяют предположить наличие тенденции к более высоким значениям порога для гласного /а/ в словах, находящихся на конце *вопросительного* предложе-

ния под синтагматическим ударением. В таких предложениях была реализована ИК-3 в системе Е. А. Брызгуновой, для которой характерен резкий высокий подъем тона на ударном слоге. Такой результат, по-видимому, связан с тем, что значение длины периода основного тона в этом контексте оказывается меньше, чем в остальных контекстах, за счет резкого изменения частоты основного тона.

В остальных контекстах значения порога восприятия длительности ударных гласных были в среднем такими же, как и в предыдущем эксперименте: от 2 до 4 периодов основного тона, что приблизительно соответствует 20–40 мс.

**7. Заключение.** По результатам двух перцептивных экспериментов был установлен порог восприятия длительности ударных гласных. В среднем он составляет 2–4 периода основного тона, что приблизительно соответствует 20–40 мс. Величина порога не зависит от количества слогов в слове и места ударения. Влияние *типа гласного* подтверждается результатами статистического анализа. Установление более точных соотношений между значениями порога для разных гласных требует дополнительных данных, полученных от значительно большего числа испытуемых.

Наблюдается тенденция к более высоким значениям инкрементного порога по сравнению с декрементным. Это говорит о том, что слушающий более чувствителен к уменьшению длительности звука.

Приблизительно 70 % испытуемых не всегда ощущают разницу по длительности, даже если она меняется в два раза; чаще всего это происходит в декрементном ряду стимулов (при уменьшении длительности). Это может говорить о том, что при естественном процессе восприятия речи слушающий не оценивает длительность отдельных звуковых единиц. Такой результат согласуется с положениями, изложенными в ряде лингвистических и психологических работ (см. работы А. А. Леонтьева, С. И. Бернштейна, Р. О. Якобсона и др.) и показывающими, что слушающий не воспринимает в речи фонемы одну за другой, а идентифицирует и детерминирует значащие звуковые единства, т. е. слова [1, 8, 9].

### Литература

1. *Бернштейн С.И.* Вопросы обучения произношению применительно к преподаванию русского языка иностранцам // Вопросы фонетики и обучение произношению. М. 1937. С. 5–61
2. *Бондарко Л.В., Загоруйко Н.Г., Кожевников В.А., Молчанов А.П., Чистович Л.А.* Модель восприятия речи человеком // Новосибирск. 1968. 59 с.
3. *Венцов А.В.* Исследование восприятия речи. XXI век // Идеи и методы экспериментального изучения речи. СПб. 2008. С. 193–203.
4. *Венцов А.В., Касевич В.Б.* Проблемы восприятия речи // М. 2003. 240 с.

5. *Евграфова К.В., Качковская Т.В.* Построение набора фонетических контекстов для исследования объективной и субъективной длительности звуковых единиц // Научное мнение. СПб. 2013. №12. С. 46–52.
6. *Кривнова О.Ф.* Временная структура синтагмы в связанном тексте // Экспериментальный фонетический анализ: проблемы и методы. Л. 1989. С. 128–139.
7. *Кривнова О.Ф.* Препозуальное продление гласных в связанном тексте // Тезисы докладов Всес. школы-семинара «АРСО- 16». М. 1991. С. 153–154.
8. *Леонтьев А.А.* Язык, речь, речевая деятельность // М.: Просвещение. 1969. 214 с.
9. *Якобсон Р.О.* Избранные работы. М. 1985. 460 с.
10. *Gescheider G. A.* Psychophysics: The Fundamentals // 1997. 448 p.
11. *Klatt D.* Interaction between two factors that influence vowel duration // JASA. 1973. vol. 54. pp. 1102–1104.
12. *Klatt D.* Linguistic uses of segmental duration in English: acoustic and perceptual evidence // JASA. 1976. vol. 59. pp. 1208–1221.
13. *Lehiste I.* Suprasegmentals // Cambridge. Mass. London. 1970. 194 p.
14. *Lindblom B., Lyberg B., Holmgren K.* Durational patterns of Swedish Phonology: do they reflect short-term motor memory process // Bloomington. 1981. 85 p.

## References

1. Bernstein S.I. [Problems of teaching pronunciation applied to teaching Russian to foreigners]. *Voprosy fonetiki i obuchenie proiznosheniju. Moscow – The Problems of Phonetics and Pronunciation Teaching.* 1937. pp. 5–61. (In Russ.).
2. Bondarko L.V., Zagorujko N.G., Kozhevnikov V.A. et al. *Model' vospriyatija rechi chelovekom* [The Model of Speech Perception]. Novosibirsk. 1968. 59 p. (In Russ.).
3. Ventsov A. V. [The study of speech perception. XXI Century]. *Idei i metody jeksperimental'nogo izuchenija rechi – The Ideas and Methods of Experimental Study of Speech.* Saint-Petersburg. 2008. pp. 192–203. (In Russ.).
4. Ventsov A.V., Kasevich V.B. *Problemy vospriyatija rechi* [The Problems of Speech Perception]. Moscow. 2003. 240 p. (In Russ.).
5. Evgrafova K.V. Kachkovskaya T.V. [The construction of a set of phonetic contexts for the study of objective and subjective duration of the sound units]. *Nauchnoe mnenie – The Scientific Opinion.* Saint-Petersburg. 2013. vol. 12. pp. 46–52. (In Russ.).
6. Krivnova O.F. [Temporal structure of syntagma in a coherent text]. *Jeksperimental'nyj foneticheskij analiz: problemy i metody – The Experimental Phonetic Analysis: Problems and Methods.* Leningrad. 1989. pp. 128–139. (In Russ.).
7. Krivnova O.F. [The pre-pausal extension of vowels in a coherent text]. *Tezisy dokladov Vses. shkoly-seminara «ARSO- 16» – The Proceedings of All-Soviet Union School-Seminar «ARSO - 16».* Moscow. 1991. pp. 153–154. (In Russ.).
8. Leontev A.A. *Jazyk, rech', rechevaja dejatel'nost'* [Language, Speech, the Activity of Speaking]. Moscow. 1969. 214 p. (In Russ.).
9. Jakobson R. O. *Izbrannye raboty* [Selected Writings]. Moscow. 1985. 460 p. (In Russ.).
10. Gescheider G. A. Psychophysics: The Fundamentals. 1997. 448 p.
11. Klatt D. Interaction between two factors that influence vowel duration. JASA. 1973. vol. 54. pp. 1102–1104.
12. Klatt D. Linguistic uses of segmental duration in English: acoustic and perceptual evidence. JASA. 1976. vol. 59. pp. 1208–1221.
13. Lehiste I. Suprasegmentals. Cambridge. Mass. London. 1970. 194 p.
14. Lindblom B., Lyberg B., Holmgren K. Durational patterns of Swedish Phonology: do they reflect short-term motor memory process. Bloomington. 1981. 85 p.

**Евграфова Карина Владимировна** — к-т филол. наук, доцент кафедры фонетики и методики преподавания иностранных языков, заведующая лабораторией экспериментальной фонетики СПбГУ. Область научных интересов: фонетика, акустика, речевые технологии. Число научных публикаций — 30. [karinaevgr@mail.ru](mailto:karinaevgr@mail.ru); СПбГУ, Университетская наб., д. 7-11, г. Санкт-Петербург, 199034, РФ; р.т. +7(812)328-9565.

**Evgrafova Karina Vladimirovna** — Ph.D., associate professor, Department of Phonetics SPbSU, head of the Laboratory of Experimental Phonetics SPbSU. Scientific interests: phonetics, acoustics, speech perception, speech technologies. The number of publications — 30. [karinaevgr@mail.ru](mailto:karinaevgr@mail.ru); 195426, 67, 11-2, Industrialny prospect; office phone +7(931)3437976.

**Качковская Татьяна Васильевна** — младший научный сотрудник, Санкт-Петербургский государственный Университет. Область научных интересов: фонетика, акустика, просодика. Число научных публикаций — 7. [tania200288@rambler.ru](mailto:tania200288@rambler.ru); СПбГУ, Университетская наб., д. 11, г. Санкт-Петербург, 199034, РФ; р.т. +7(812)328-9565.

**Kachkovskaya Tatiana Vasiljevna** — junior researcher, Department of Phonetics SPbSU researcher at the Laboratory of Experimental Phonetics SPbSU. Scientific interests: phonetics, acoustics, prosody. The number of publications — 7. [tania200288@rambler.ru](mailto:tania200288@rambler.ru); Saint-Petersburg, University embankment, 7-11; office phone +7(812)328-9565.

**Поддержка исследований.** Работа выполнена при финансовой поддержке Гранта Президента РФ для поддержки молодых российских ученых кандидатов наук (проект МК-1398.2013.6).

**Acknowledgements.** This research is supported by the President Grant for Government Support of Young Russian Scientists (МК-1398.2013.6).

## РЕФЕРАТ

### *Евграфова К.В., Качковская Т.В.* Дифференциальные пороги восприятия длительности языковых единиц в зависимости от фонетического контекста.

В данной статье представлены результаты исследования «Перцептивная шкала длительностей фонетических единиц языка». Целью данного исследования являлось изучение параметра длительности звуковых сегментов высказывания с учетом артикуляционно-акустических и акустико-слуховых отношений.

В разделах статьи представлены: краткий обзор литературных данных по проблеме, описание психоакустических экспериментов, целью которых было установить слуховые дифференциальные пороги длительности для русских гласных с учетом фонетического контекста, статистический анализ результатов эксперимента, а также выводы по полученным данным.

По результатам двух перцептивных экспериментов был установлен порог восприятия длительности ударных гласных. В среднем он составляет 2–4 периода основного тона, что приблизительно соответствует 20–40 мс. Величина порога не зависит от количества слогов в слове и места удара. Влияние типа гласного подтверждается результатами статистического анализа. Установление более точных соотношений между значениями порога для разных гласных требует дополнительных данных, полученных от значительно большего числа испытуемых.

Наблюдается тенденция к более высоким значениям инкрементного порога по сравнению с декрементным. Это говорит о том, что слушающий более чувствителен к уменьшению длительности звука.

Приблизительно 70 % испытуемых не всегда ощущают разницу по длительности, даже если она меняется в два раза; чаще всего это происходит в декрементном ряду стимулов (при уменьшении длительности). Это может говорить о том, что при естественном процессе восприятия речи слушающий не оценивает длительность отдельных звуковых единиц.

## SUMMARY

### *Evgrafova K.V. Kachkovskaya T.V.* Differential Thresholds of Perception of Sound Duration in Different Phonetic Contexts.

The paper presents the results of the research «Perceptual Scale of Sound Duration». The goal of the research was to study acoustic and perceptive variability of sound duration. The divisions of the paper present a short review of relevant studies, the description of two psychoacoustic experiments, the statistical analysis of their results and the conclusions based on the analysis. The experiments were aimed at obtaining differential thresholds for Russian vowel duration with the account of phonetic context. The study showed the following results.

On average, the differential threshold value is 2–4 F0 periods which is about 20–40 ms. The value does not depend on word duration (in syllables) and stress pattern and seems to depend on vowel type. The increment threshold values tend to be higher than the decrement ones. About 70 % of the subjects were not able to perceive the difference in duration even it is increased or decreased by two times.

И.С. АЗАРОВ, А.А. ПЕТРОВСКИЙ  
**ФОРМИРОВАНИЕ ПЕРСОНАЛЬНОЙ МОДЕЛИ ГОЛОСА  
ДИКТОРА С УНИВЕРСАЛЬНЫМ ФОНЕТИЧЕСКИМ  
ПРОСТРАНСТВОМ ПРИЗНАКОВ НА ОСНОВЕ  
ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ**

---

*Азаров И.С., Петровский А.А. Формирование персональной модели голоса диктора с универсальным фонетическим пространством признаков на основе искусственной нейронной сети.*

**Аннотация.** В работе исследуется возможность формирования модели голоса заданного диктора на основе записей образцов его голоса с транскрипцией. В работе предлагается практический способ построения голосовой модели и результаты экспериментов ее применения к задаче конверсии голоса. Модель использует искусственную нейронную сеть, устроенную по принципу автоматического кодера, устанавливающую соответствие между пространством речевых параметров и пространством возможных фонетических состояний, унифицированным для произвольного голоса.

**Ключевые слова:** конверсия голоса, синтез речевого сигнала, искусственная нейронная сеть.

*Azarov E., Petrovsky A. Training Personal Voice Model of a Speaker with Unified Phonetic Space of Features Using Artificial Neural Network.*

**Abstract.** The paper investigates possibility of creating a personal voice model using transcribed speech samples of a specified speaker. The paper presents a practical way of building such speech model and some experimental results of applying the model to voice conversion. The model uses an artificial neural network organized as autoencoder that establishes correspondence between space of speech parameters and space of possible phonetic states, unified for any voice.

**Keywords:** voice conversion, speech synthesis, artificial neural network.

---

**1. Введение.** В настоящее время большую часть существующих задач обработки речевых сигналов можно условно разделить на следующие основные направления: синтез речи по тексту, распознавание речи, создание различных звуковых эффектов (например, конверсия голоса), кодирование речи и улучшение речевых характеристик (например, повышение разборчивости речи и шумоподавление).

Синтез речи по тексту и распознавание речи являются, вероятно, наиболее важными из всех, поскольку в перспективе могут привести к организации полноценного голосового интерфейса между человеком и компьютером. Последние коммерческие решения в данных областях являются многообещающими: представлены синтезаторы речи по тексту (<http://www.speechpro.ru/product/recognition/tts>, <https://play.google.com/store/apps/details?id=com.google.android.tts&hl=ru>), позволяющие генерировать практически натуральную речь с низким уровнем слышимых артефактов, и современные распознаватели, по-

зволяющие распознавать фразы слитной речи с недостижимой ранее высокой вероятностью. Основной причиной такого заметного прогресса является появление возможности использования огромных речевых корпусов, обусловленное многократным удешевлением и ускорением вычислительного процесса. Так, для качественного синтеза речи по тексту теперь вместо аллофонного синтеза, использующего речевую базу аллофонов общей продолжительностью в несколько минут, применяется корпусный синтез, использующий десятки часов речевого материала состоящий из всевозможных комбинации аллофонов и отдельных слов в разных контекстах. Создание такой базы данных для каждого диктора является очень сложным процессом, требующим привлечения большого числа специалистов. Недавно предложен способ автоматического создания синтезатора речи по тексту, позволяющая формировать голосовую и языковую модели на основе речевых записей и их транскрипций [1]. Что касается распознавателей слитной речи, то для них при обучении системы распознавания теперь используются намного большие речевые выборки, содержащие различные голоса.

Задача изменения голоса (конверсии голоса) появилась сравнительно недавно, тем не менее развитие данного направления происходит очень активно. Целью конверсии голоса является замена личности говорящего при сохранении содержания исходного речевого сообщения. Решение данной задачи подразумевает установление соответствия между голосом исходного диктора и целевого на основе некоторого обучающего речевого материала. Большинство из всего многообразия предложенных способов решения можно разделить на несколько основных групп: использование модели Гауссовых смесей [2], масштабирование частотной шкалы [3] и использование нейронных сетей [4,5]. Последние опубликованные результаты показывают что данные подходы позволяют достигать средней разборчивости и узнаваемости целевого диктора более 75%, что, учитывая сложность задачи, является достаточно высоким показателем. Основной проблемой в конверсии голоса является несоответствие обучающих данных. Исходный и целевой дикторы даже чисто теоретически не могут говорить одинаково, вследствие этого функция конверсии страдает чрезмерным усреднением данных, что неизбежно сказывается на качестве выходного сигнала. Наиболее качественный результат конверсии достигается при использовании параллельных обучающих фраз, т.е. когда исходный и целевой дикторы произносят одни и те же обучающие фразы. Однако, наиболее интересными, исходя из практических соображений, являются решения, позволяющие выполнять обучение при по-

мощи произвольных текстовых корпусов не соответствующих друг другу [6]. Далее эти два подхода будем называть «текстозависимым» и «текстонезависимым» соответственно.

Кодирование речи является одним из самых первых направлений, сформировавшихся в цифровой обработке речевых сигналов. Наиболее эффективными здесь оказались решения, основанные на параметрическом моделировании речевого сигнала. Предложено большое количество различных моделей: линейное предсказание с различными способами возбуждения [7], кепстральный анализ [8], синусоидальное представление [9], гармоника+шум [10–12] и т.д. Современные вокодерные системы обеспечивают удовлетворительное кодирование узкополосного речевого сигнала со скоростью потока 2.4–12кбит/с и широкополосного со скоростью потока выше 8кбит/с. Не решенной в полной мере остается задача кодирования речи на сверхнизких скоростях потока менее 1.2кбит/с.

Автоматическое повышение разборчивости речи и шумоподавление представляют собой область, в которой нашли применение различные подходы: 1) универсальные методы такие как спектральное взвешивание [13–15] и обработка в подпространствах [16,17]; 2) методы основанные на особенностях слухового восприятия, такие как фильтрация в модуляционной области [18–20]; 3) методы основанные на особенностях речеобразования [21].

Все вышеперечисленные задачи относятся к одному и тому же объекту исследования (речевому сигналу), и, несмотря на все имеющиеся различия, между ними существует внутренняя взаимосвязь. Дальнейший успех в решении каждой из этих задач зависит от того, насколько удачно моделируется речь как феномен в различных его аспектах: представление процесса речеобразования, интерпретации содержания речевого сообщения (в том числе фонетического, смыслового, эмоционального) и процесса восприятия. Потому разработка способов для наиболее адекватного и универсального моделирования речевого сигнала представляется очень перспективным научным направлением.

В данной работе исследуется возможность создания параметрического описания речевого сигнала, основанного на раздельном моделировании голоса диктора и содержания речевого сообщения. Речевой сигнал представляется в виде некоторой последовательности данных, характеризующей содержание сообщения, и модели голоса диктора, при помощи которого данная последовательность преобразовывается в речь. Процесс речеобразования рассматривается как некоторая систе-

ма, позволяющая озвучивать речевые сообщения заданным голосом как показано на рисунке 1.

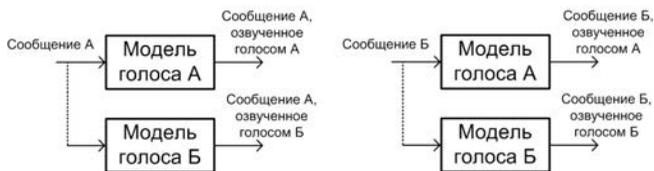


Рис. 1. Моделирование процесса речеобразования

Предполагается возможность применения модели голоса и в обратном направлении, т.е. не только для озвучивания сообщений, но и для декодирования их из речевого сигнала.

Фундаментальной проблемой является разделение параметров речевого сигнала на индивидуальные характеристики диктора и фонетическое содержимое. В чистом виде такого рода разделение естественным образом возникает в задаче синтеза речи по тексту, поскольку отдельно существует текст (содержание сообщения) и модель голоса (речевая база данных, соответствующая заданному диктору), которая используется при озвучивании этого текста. В распознавании речи решается похожая задача разделения, однако в обратном направлении: на вход поступает речевой сигнал, на выходе нужно сформировать текст сообщения.

*Постановка задачи.* В отличие от классических задач преобразования речь→текст и текст→речь в настоящей работе предлагается использовать несколько другую логику разделения параметров речи. Предполагается, что озвучиваемое сообщение содержит не только текстовую информацию и состоит не из последовательности фонем, а скорее из последовательности состояний речевого тракта, выполняющих универсальную фонетическую и просодическую функции для любого голоса. Таким образом, ставится задача реализации пары преобразований речь→фонемы+просодика и фонемы+просодика→речь с использованием параметрического описания голоса исходного и целевого дикторов. Постановка задачи в таком виде делает возможную область применения модели голоса очень широкой, поскольку во-первых, модель является фонетически мотивированной, однако не подразумевает явного преобразования речи в текст; во-вторых, позволяет сохранять и использовать просодику исходного речевого сообщения; в-третьих, теоретически после преобразования сохраняется возможность восстановления исходного сигнала с субъективно незначительными потерями, что невозможно достичь в результате последовательного выполнения преобразований речь→текст и текст→речь. Для

решения поставленной задачи необходимо создать параметрическую модель голоса диктора, а также соответствующие методы анализа (извлечения параметров модели из речевого сигнала) и синтеза (генерирование речевого сигнала из параметров модели). Модель голоса, предлагаемая в данной работе использует нейронную сеть, построенную по принципу автоматического кодера, реализованного в виде искусственной нейронной сети (autoencoder) [22]. В отличие от Гауссовых смесей и масштабирования частотной шкалы, нейронная сеть обладает большими возможностями, поскольку позволяет формировать более сложную функцию отображения входных и выходных характеристических векторов. Основной идеей является использование свойства кодера автоматически находить и упорядочивать похожие данные. В работе также приводятся практические результаты применения предлагаемой модели для конверсии голоса.

**2. Использование модели голоса в задачах обработки речевых сигналов. Возможные области применения.** В данном разделе приводятся некоторые возможные области применения параметрической модели голоса, и дается краткая характеристика вероятных преимуществ.

*Конверсия голоса.* При использовании параметрических моделей голоса задача конверсии голоса сводится к последовательному использованию 2-х моделей: исходного и целевого дикторов, как показано на рисунке 2. Причем, по типу обучения такая система конверсии голоса является текстонезависимой, поскольку модели голосов формируются независимо друг от друга. Такой способ обучения системы является предпочтительным, так как позволяет использовать различные речевые сообщения исходного и целевого дикторов.



Рис. 2. Использование персональной модели голоса для конверсии голоса

*Кодирование речи.* Разделение голоса диктора и сообщения может оказаться полезным при кодировании речи для сверхнизких скоростей передачи. Основопологающей идеей здесь служит то, что диктор не меняется (либо меняется очень редко) во время сеанса связи. Следовательно, голос диктора может быть долговременным параметром, обновляющимся значительно реже, чем параметры сообщения. При обучении системы на стороне кодера и декодера формируются универсальные базы возможных голосов. При кодировании сообщения выполняется выбор голоса из базы данных, наиболее близкого к голосу говорящего. Далее декодеру передается выделенное из речи сооб-

щение и индекс найденного голоса, как показано на рисунке 3. Учитывая, что состояния модели, используемые для описания сообщения, могут описываться набором параметров малой размерности, а так же и то, что эти параметры могут быть грубо проквантованы, данная схема кодирования может оказаться эффективной.



Рис. 3. Использование персональной модели голоса для кодирования речевого сигнала

*Синтез речи по тексту.* Использование модели голоса представляет интерес и для задачи синтеза речи по тексту. Процесс синтеза может быть организован таким образом, чтобы формирование речевого сообщения выполнялось отдельно от озвучивания его заданным голосом, как показано на рисунке 4.



Рис. 4. Использование персональной модели голоса для синтеза речи по тексту

В результате процесс добавления в базу данных синтезатора речи по тексту новых дикторов упрощается и унифицируется. При качественной автоматизации формирования модели голоса на основе речевых записей возможно создание синтезаторов речи по тексту с функцией добавления пользовательских голосов.

*Шумоподавление.* Модель голоса также может применяться и для очистки речи от шума – рисунок 5. Основная идея заключается в том, что речевое сообщение представляет собой последовательность состояний, которые имеют значительную продолжительность (от 10 до 50 мс), а продолжительность переходного процесса соответствующего смене состояний можно считать короткой (менее 1 мс). Таким образом, чистое сообщение должно иметь некоторую сегментную структуру. Учитывая это, процедуру фильтрации можно организовать в виде двух следующих шагов: 1) аппроксимация зашумленного сообщения наиболее вероятной последовательностью, состоящей из сегментов с ограничением минимальной длины путем решения задачи оптимизации; 2) сглаживание каждого выделенного сегмента фильтром низких частот.

Критическим фактором в данном походе является точность декодирования сообщения из входного речевого сигнала. Поскольку точность будет сильно зависеть от интенсивности и характера шума, вероятно, что данный подход может применяться только для шумов определенного типа. Интересным вопросом для исследования является также возможность подавления акустического эха и реверберации при помощи предложенной схемы.

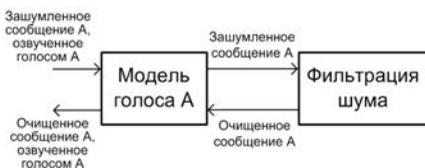


Рис. 5. Использование персональной модели голоса для шумоподавления

*Повышение разборчивости.* Рассмотрим возможность применения голосовой модели для повышения разборчивости речевого сигнала. Соответствующая схема обработки речи, представленная на рисунке 6, выполняет сложное преобразование просодики входящего речевого сообщения, моделирующие неторопливую и внятную речь. Используя подобную схему можно выполнять как изменение общих характеристик голоса, так и корректировку звучания отдельных фонем. Отдельной задачей является повышение разборчивости речи для людей с патологиями слуха [23]. Потеря слуха чаще всего характеризуется утратой чувствительности в высокочастотной области спектра. В результате этого в первую очередь ухудшается способность воспринимать фонемы, для которых характерны высокочастотные звуки – например [х]/[ф] и [з]/[ж]. Изменяя звучание этих отдельных фонем можно добиться повышения разборчивости. Другими возможными применениями этого подхода могут быть автоматическая коррекция акцента и искусственное расширение частотного диапазона в узкополосной телефонии.



Рис. 6. Схема использования персональной модели голоса для повышения разборчивости

Во всех перечисленных выше приложениях для повышения точности декодирования сообщения дополнительно с входным рече-

вым сигналом может использоваться и его транскрипция, в случаях, когда это допустимо.

Целью данной работы является исследование возможности автоматического создания персональной модели голоса диктора по записанным образцам его голоса с транскрипцией, а также возможности практического использования такой модели для решения задачи конверсии голоса с текстонезависимым обучением.

**3. Параметрическое представление речевого сигнала.** Для выполнения обработки речевого сигнала необходимо выполнить его параметрическое описание, т.е. представить в виде последовательности характеристических векторов одинаковой размерности. В контексте решаемой задачи речевой сигнал удобно рассматривать как процесс имеющий квазипериодические (детерминированные) и шумовые (стохастические) составляющие [24,25]. Можно считать, что квазипериодические составляющие порождаются периодическими колебательными движениями голосовых связок и характерны для гласных (вокализованных звуков), в то время как шумовые возникают вследствие непериодических колебаний и характерны для шипящих согласных (невокализованных звуков). Моделирование процесса обеспечивается путем раздельного представления каждой из этих составляющих при помощи разных средств описания. Этот подход широко применяется в современных системах обработки речи [26,27]. В данной работе используется аналогичный способ параметрического описания, реализованный при помощи алгоритма оценки параметров периодической модели с многокомпонентным гармоническим возбуждением [28].

Речевой сигнал разбивается на перекрывающиеся фрагменты каждый из которых описывается набором параметров: спектральной огибающей, мгновенной частотой основного тона (если фрагмент вокализованный) и типом возбуждения, который может быть вокализованным, невокализованным либо смешанным.

Квазипериодическая составляющая речевого сигнала  $s(n)$  представляется в виде суммы синусоид или действительной части комплексных экспонент с непрерывной амплитудой, частотой и фазой, а шумовая как случайный процесс с заданной спектральной плотностью мощности (СПМ):

$$s(n) = \sum_p^P A_p(n) \cos \varphi_k(n) + r(n) = \operatorname{Re} \left[ \sum_p^P A_p(n) e^{j\varphi_p(n)} \right] + r(n),$$

где  $P$  – число синусоид (комплексных экспонент),  $A_p(n)$  – мгновенная амплитуда  $p$ -ой синусоиды,  $\varphi_p(n)$  – мгновенная фаза  $p$ -ой синусоиды,  $r(n)$  – аperiodическая составляющая. Мгновенная частота  $F_p(n)$ , на-

ходящаяся в интервале  $[0, \pi]$  ( $\pi$  соответствует частоте Найквиста), является производной от мгновенной фазы. Предполагается, что амплитуда изменяется медленно, что означает ограничение частотной полосы каждой из составляющих. Используя полученные гармонические амплитуды вокализованной и СПМ невокализованной составляющих, формируется общая спектральная огибающая.

Характеристический вектор состоит из значения частоты основного тона, спектральной огибающей и признака вокализованности текущего речевого фрагмента.

Данный набор параметров, выделяется из речевого сигнала при помощи алгоритма, состоящего из следующих шагов:

- оценка мгновенной частоты основного тона при помощи устойчивого к ошибкам алгоритма слежения за мгновенной частотой основного тона IRAPT (Instantaneous Robust Algorithm for Pitch Tracking) [29];

- деформация временной оси сигнала для обеспечения стационарности частоты основного тона;

- оценка мгновенных гармонических параметров речевого сигнала с использованием ДПФ-модулированного банка фильтров – каждая гармоника основного тона вокализованной речи попадает в отдельный канал банка фильтров, где преобразуется в аналитический комплексный сигнал, из которого выделяется мгновенная амплитуда, фаза и частота;

- на основе анализа полученных значений мгновенной частоты различные области спектра классифицируются как периодические и аperiodические;

- гармоники, принадлежащие периодическим областям спектра синтезируются и вычитаются из исходного сигнала;

- остаток переводится в частотную область при помощи кратковременного преобразования Фурье;

- параметры синтезированных гармоник и СПМ остатка объединяются в одну общую спектральную огибающую и переводятся в логарифмическую шкалу;

- смежные спектральные огибающие анализируются для определения способа возбуждения всего анализируемого фрагмента сигнала.

Каждая спектральная огибающая представляется в виде вектора логарифмических значений энергии равнорасположенных в шкале мелов. Для речевого сигнала с частотой дискретизации 44.1 кГц используется 100-мерный вектор. Размерность вектора определяет компромисс между качеством реконструкции сигнала и вычислительной сложностью. На основе практических экспериментов установлено, что

выбранная размерность является достаточной для реконструкции натуральной речи.

**4. Модель голоса с фонетической привязкой на основе нейронной сети.** Голосовая модель использует нейронную сеть, построенную по принципу автоматического кодера. Автоматический кодер представляет собой многослойную нейронную сеть, которая преобразовывает многомерные данные в коды меньшей размерности и затем восстанавливает их в первоначальном виде [22] – рисунок 7.

Понижение размерности данных широко используется в задачах классификации, связи, распознавания и др. Достаточно простым и широко применяемым способом является анализ главных компонент, выделяющий в обучающей выборке направления с максимальной дисперсией и описывающий каждый элемент выборки через координаты по этим направлениям. В работе [22] показано, что системы понижения размерности данных на основе нейронных сетей обладают гораздо более широкими возможностями, поскольку, в отличие от метода анализа главных компонент позволяют выполнять нелинейные преобразования.

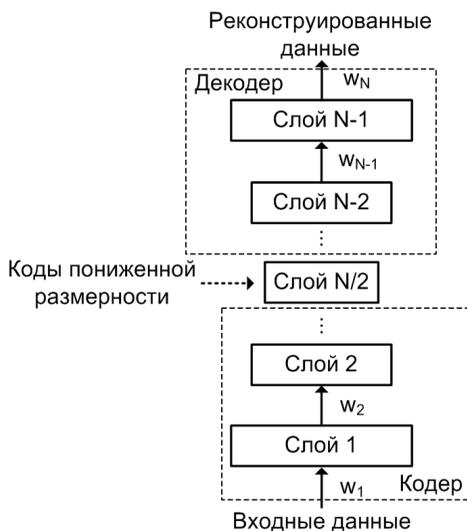


Рис. 7. Автоматический кодер на основе нейронной сети

Обе сети, кодер и декодер можно обучить вместе, уменьшая разницу между исходными данными и их реконструкцией. Частные производные всех параметров легко вычисляются, используя правило дифференцирования сложной функции, для распространения произ-

водной ошибки сперва через сеть декодирования, а затем через сеть кодирования.

В контексте конверсии голоса полезным свойством автоматического кодера является способность самостоятельно классифицировать, упорядочивать и находить похожие данные без "учителя", т.е. каких-либо заранее заданных целевых меток в обучающей выборке. Показано, что при достижении хорошей реконструкции данных, близкие коды пониженной размерности соответствуют схожим входным данным. Это может быть использовано для построения производительных ассоциативных поисковых систем [30].

В задаче конверсии голоса основной проблемой является поиск соответствия между параметрами исходного голоса и целевого. Проблема возникает из-за того, что параметры голоса выделяются на основе речевых фрагментов, соответствующим разным состояниям исходного и целевого дикторов. Каждый диктор имеет некоторое пространство возможных вариаций при произношении одних и тех же фонетических единиц, обусловленных многообразием речевых оттенков, выражением различных эмоций и интонаций. Таким образом одна и та же фонема, находясь в одном и том же фонетическом контексте может звучать по-разному. Учитывая это, очень сложно найти соответствие между состояниями разных дикторов.

Основная идея применения автоматического кодера в данной работе заключается в использовании фонетически мотивированных кодов пониженной размерности, (имеющих фонетическую интерпретацию). Предлагается организовать процесс обучения нейронной сети таким образом, чтобы коды, соответствующие одной фонеме компактно располагались в одной определенной области пространства, причем границы, примыкающих к нему областей, обеспечивали плавный переход к другим фонемам – рисунок 8.

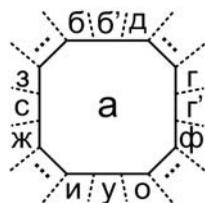


Рис. 8. Пространство кодов пониженной размерности с фонетической привязкой

Для того, чтобы обеспечить возможность непрерывного перехода из любой фонемы в любую размерность пространства равна числу используемых фонем. Расположение каждой фонемы в про-

странстве кодов фиксировано вдоль координат пространства в интервале от 0 до 1.

Использованная конфигурация искусственной нейронной сети показана на рисунке 9. Кодер выполняет функцию отображения:

$$H = (w_4 RL(w_3 RL(w_2 RL(w_1 X + b_1) + b_2) + b_3) + b_4) \otimes M$$

где  $X$  – характеристический вектор речевого сигнала,  $H$  – вектор пониженной размерности,  $M$  – вектор фонетической маски,  $w_{1-4}$  и  $b_{1-4}$  – весовые коэффициенты и смещения соответствующих сигналов сети,  $\otimes$  обозначает поэлементное умножение. В сети используется кусочно-линейная функция активации  $RL(x) = \max(0, x)$ , поскольку показано, что она обеспечивает более эффективное внутреннее представление речевых данных по сравнению с логистической и позволяет ускорить процесс обучения [31]. На выходе кодера формируются коды пониженной размерности, на которые накладываются ограничения для того, чтобы выполнить фонетическую привязку (см. рисунок 9). Наложение ограничения выполняется путем перемножения сигнала  $H$  на фонемную маску, представляющую собой разреженную матрицу, и формируемую на основе фонетической разметки речевого корпуса.

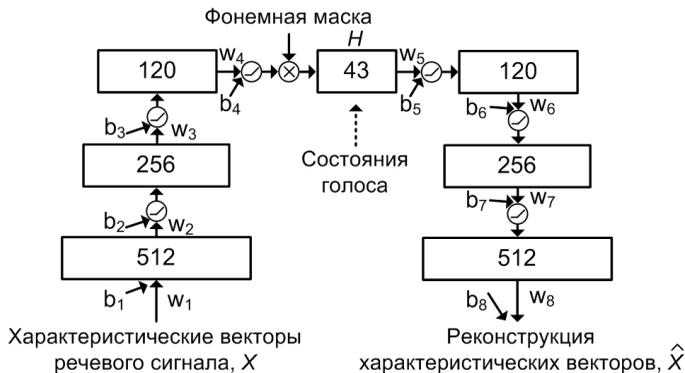


Рис. 9. Формирование состояний голоса на основе нейронной сети

Декодер выполняет реконструкцию кодов пониженной размерности в характеристические векторы  $\hat{X}$ . Соответствующая функция отображения имеет следующий вид:

$$\hat{X} = (w_8 RL(w_7 RL(w_6 RL(w_5 H + b_5) + b_6) + b_7) + b_8).$$

Использовалось следующее число нейронов в каждом скрытом слое нейронной сети: 512-256-120-43-120-256-512. Обучение сети включает несколько этапов, которые кратко описаны ниже.

*Предварительная сегментация обучающего речевого корпуса на фонемы.* Для того чтобы при обучении нейронной сети было возможно осуществить фонетическое разграничение состояний голоса, необходимо установить соответствие каждого обучающего характеристического вектора определенной фонеме. С этой целью выполняется сегментация обучающего речевого корпуса на фонемы, используя фонемную транскрипцию. Задача определения границ фонем в речевом сигнале имеет классическое решение, основанное на использовании скрытой марковской модели [32]. Для повышения точности анализа может применяться предварительная ручная (частичная либо полная) разметка.

*Инициализация параметров сети и предварительное обучение.* Поиск оптимальных коэффициентов многослойного кодера является сложной задачей, поскольку для того, чтобы алгоритм обратного распространения ошибки был эффективным, требуется хорошее начальное приближение. При использовании больших начальных коэффициентов процесс обучения обычно сходится к плохому локальному минимуму, использование малых коэффициентов приближает градиент в начальных слоях к нулю, что делает невозможным обучение сети с большим количеством слоев. Известен метод раздельного предварительного обучения слоев при помощи ограниченной машины Больцмана, успешно использованный в различных задачах машинного обучения [22]. В настоящей работе применяется схема предварительного обучения, основанная на частичном обнулении данных. Каждая пара матриц коэффициентов инициализируется случайными числами, и тренируется отдельно в виде нейронной сети с одним скрытым слоем. На вход сети подаются данные, часть из которых случайным образом обнуляется, на выходе сети восстанавливаются полные исходные данные. Причем обеспечивается равенство соответствующих матриц кодера и декодера  $w_1 = w_8^T$ ,  $w_2 = w_7^T$ ,  $w_3 = w_6^T$ ,  $w_4 = w_5^T$ . Обучение выполняется при помощи алгоритма обратного распространения ошибки с накоплением градиента (momentum). При обучении матриц коэффициентов  $w_4 = w_5^T$  частные производные по внутренним сигналам вычисляются с дополнитель-

ным слагаемым, обеспечивающим повышение активности в точках, обозначенных единицами в фонетической маске и понижение в точках, обозначенных нулями. Фонетическая маска формируется на основании предварительной сегментации речевого корпуса. Пространство кодов состояний голоса и маска имеют размерность 43 (равную числу фонем русского языка). Каждая фонема соответствует отдельной координате. Маска содержит единицы в координатах, соответствующих фонеме каждого вектора состояния, как показано на рисунке 10. Промежуточный переход между двумя соседними фонемами задается двумя единицами.

Номер вектора состояния

		1	2	3	4	5	6	7	8	9	10	11	12	13		
Фонемы	а	0	0	1	1	1	1	1	0	0	0	1	1	1	...	1
	э	0	0	0	0	0	0	0	0	0	0	0	0	0	...	2
	и	0	0	0	0	0	0	0	0	0	0	0	0	0	...	3
	о	0	0	0	0	0	0	0	0	0	0	0	0	0	...	4
	⋮															
	⋮															
	⋮															
	⋮															
	⋮															
	⋮															
	⋮															
	⋮															
	⋮															
б'	0	0	0	0	0	0	0	0	0	0	0	0	0	...	41	
м	1	1	1	1	0	0	1	1	1	1	1	0	0	...	42	
м'	0	0	0	0	0	0	0	0	0	0	0	0	0	...	43	
		/ м		а		м		а/								

Рис. 10. Маска кодов пониженной размерности для осуществления фонетической привязки

*Обучение системы кодера/декодера.* После предварительного определения коэффициентов каждого слоя выполняется подгонка параметров всей модели при помощи модификации алгоритма обратного распространения ошибки RPROP (Resilient back PROPagation) [33]. При перемножении внутреннего сигнала на маску каждый отдельный вектор попадает либо в определенную плоскость пространства, соответствующую двум смежным фонемам, либо на координатную ось, соответствующую одной фонеме. В процессе подгонки параметров происходит уменьшение ошибки реконструкции и упорядочение характеристических векторов каждой фонемы вдоль осей пространства – рисунок 11.

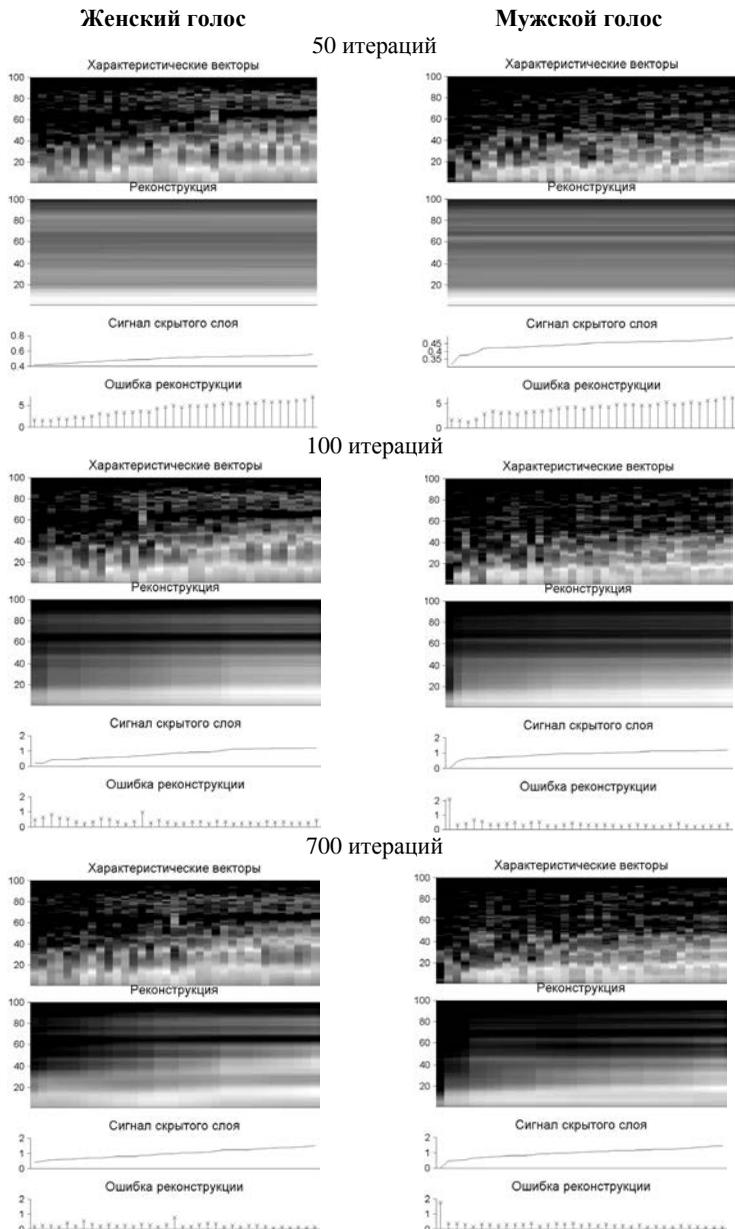


Рис. 11. Расположение характеристических векторов фонемы [a] вдоль соответствующей оси пространства кодов пониженной размерности

Процедура предварительного обучения создает достаточно хорошее начальное приближение и характеристические векторы в начале процедуры уже определенным образом упорядочены. Однако, как показано на рисунке 11, после выполнения некоторого числа итераций происходит перестановка векторов, уменьшающая ошибку реконструкции и разницу между соседними векторами. В результате обучения формируется модель голоса, которая включает модель каждой отдельной фонемы и переходов между ними, содержащихся в обучающей выборке.

**5. Результаты экспериментов.** Целью выполненных экспериментов является оценка практической применимости предложенной модели для решения задачи конверсии голоса с текстонезависимым обучением. Обученные модели голосов использовались для поиска соответствия между характеристическими векторами исходного и целевого дикторов. Модели обучались независимо друг от друга. Конверсия речевого сигнала выполнялась с использованием ручной фонетической разметки, в которой выделялся характерный «центральный» характеристический вектор каждой фонемы. На основании разметки автоматически определялись «переходные» характеристические векторы, относящиеся к границам между фонемами. Соответствие между центральными и переходными характеристическими векторами исходного и целевого диктора устанавливалось следующим образом. На вход кодера исходного диктора подавался входной характеристический вектор и вычислялся вектор пониженной размерности, умноженный на фонемную маску. Из обучающей выборки целевого диктора извлекался характеристический вектор, наиболее близкий к полученному в пространстве кодов пониженной размерности. Все необходимые для синтеза характеристические векторы вычислялись путем интерполяции между конвертированными центральными и переходными векторами.

*Речевая база и оценка качества конверсии.* Для экспериментальной оценки качества конверсии использовалась речевая база на русском языке. База содержит широкополосную речь 6-и дикторов (3 мужчин и 3 женщины), записанную с частотой дискретизации 44,1 кГц. Для каждого из дикторов в базе содержится по 26 фраз для обучения и по 4 фразы для конверсии.

В ходе экспериментов выполнено сравнение двух методов: 1) изложенный выше метод конверсии с текстонезависимым обучением на основе автоматического кодера (далее обозначается как ‘ТН’) и метод с текстозависимым обучением на основе нейронной сети с кусочно-линейной функцией активации (далее обозначается как ‘ТЗ’) [5].

Для обоих методов использовались одинаковые алгоритмы параметризации и синтеза речевого сигнала, одинаковые обучающие

последовательности исходного и целевого дикторов, одинаковые сигналы возбуждения и одинаковые целевые контуры основного тона. Результаты конверсии оценивались субъективно в терминах узнаваемости целевого диктора и натуральности звучания реконструированной речи с использованием средних значений оценок экспертов MOS (mean opinion score). В ходе эксперимента задействовано четверо слушателей, которые ставили оценки (по шкале от 1 до 5). Усреднение оценок выполнено отдельно по четырем группам в зависимости от направления конверсии мужчина-мужчина, мужчина-женщина, женщина-мужчина, женщина-женщина (обозначенных как "мм", "мж", "жм" и "жж" соответственно) для каждого из методов конверсии. Средние оценки показаны на рисунке 12.

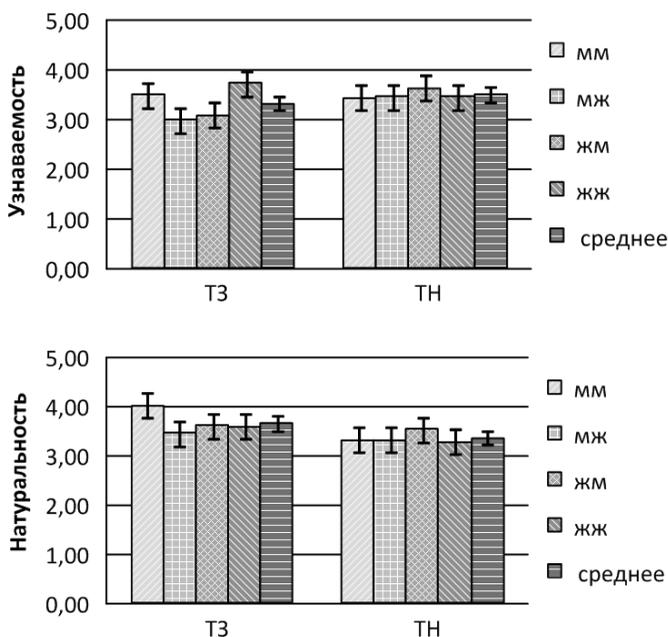


Рис. 12. Субъективная оценка узнаваемости и натуральности конвертированной речи. Средние значения оценок экспертов (доверительный интервал 95%).

На основании прослушивания и анализа оценок можно сделать вывод, что метод на основе автоматического кодера обеспечивает немного более высокую узнаваемость целевого диктора, однако несколько проигрывает по натуральности звучания. Повышение средней узнаваемости скорее всего обусловлено тем, что описанный способ позво-

ляет ослабить эффект усреднения спектральной огибающей, характерный для систем с текстозависимым обучением. Понижение натуральности обусловлено в первую очередь ошибками полуавтоматической сегментации речевого корпуса и тем, что использовалась простая модель сегментации, выделяющая только границы и центр каждой из фонем. Следует также отметить, что в методе ТЗ используется разделение параметров огибающей на высокочастотные и низкочастотные, а так же последовательность состояний диктора, генерируемых автоматически на основе текущих значений основного тона и признаков вокализованности. Таким образом, на вход системы конверсии ТЗ поступает намного больше характеристических признаков. В тоже время, необходимо учитывать, что в методе ТН использована ручная фонемная разметка, которая значительно упрощает поиск соответствия между исходными и целевыми данными.

**6. Заключение.** В работе исследуется возможность создания персональной модели голоса с фонетической привязкой на основе искусственной нейронной сети, построенной по принципу автоматического кодера. Приводятся результаты экспериментального применения данного подхода к решению задачи конверсии голоса с текстонезависимым обучением. Показано, что формирование унифицированных состояний в виде кодов пониженной размерности позволяет установить соответствие между различными голосами и может использоваться в системах синтеза речи по тексту и конверсии голоса. Особенностью полученной модели является относительная инвариантность к характеру произношения, что достигается за счет привязки внутренних состояний к фонетическому содержанию, что может использоваться в различных системах обработки речи, таких как системы автоматического распознавания и кодирования.

## Литература

1. *Watts O., Stan A., Clark R., Mamiya Y., Giurgiu M., Yamagishi J., King S.* Unsupervised and lightly supervised learning for rapid construction of TTS systems in multiple languages from 'found' data: evaluation and analysis // In: Proc. 8th ISCA Speech Synthesis Workshop. 2013. pp. 101–106.
2. *Toda T., Black A.W., Tokuda K.* Voice conversion based on maximum likelihood estimation of spectral parameter trajectory // IEEE Trans. Audio, Speech and Language Processing. 2007. vol. 15. no. 8. pp. 2222–2235.
3. *Godoy E., Rossec O., Chonavel T.* Spectral envelope transformation using DFW and amplitude scaling for voice conversion with parallel or nonparallel corpora // Proc. INTERSPEECH. Florence. Italy. 2011. pp. 673–676.
4. *Desai S., Black A.W., Yegnanarayana B., Prahallad B.* Spectral mapping using artificial neural networks for voice conversion // IEEE Trans. Audio, Speech and Language Processing. 2010. vol. 18. no. 5. pp. 954–964.
5. *Azarov E., Vashkevich M., Likhachov D., Petrovsky A.* Real-time Voice Conversion Using Artificial Neural Networks with Rectified Linear Units // Proc. INTERSPEECH Lyon. France. 2013. pp. 1032–1036.

6. *Erro D., Moreno A., Bonafonte A.* INCA Algorithm for Training Voice Conversion Systems From Nonparallel Corpora // IEEE Transactions on Audio, Speech, and Language Processing. 2010. vol. 18. no .5. pp. 944–953.
7. *Yeldener S., De Martin J.C., Viswanathan V.* A mixed sinusoidally excited linear prediction coder at 4 kb/s and below // Proc. ICASSP'98. 1998. vol. 2. pp. 589–592.
8. *Boucheron L.E., De Leon P.L., Sandoval S.* Low Bit-Rate Speech Coding Through Quantization of Mel-Frequency Cepstral Coefficients // IEEE Transactions on Audio, Speech, and Language Processing. 2012. vol. 20. no. 2. pp. 610–619.
9. *Etmoglu C.O., Cuperman V.* Matching pursuits sinusoidal speech coding // IEEE Transactions on Speech and Audio Processing. 2003. vol. 11, no. 5. pp. 413–424.
10. *Shlomot E., Cuperman V., Gersho A.* Hybrid coding: combined harmonic and wave-form coding of speech at 4 kb/s // IEEE Transactions on Speech and Audio Processing. 2001. vol. 9. no. 6. pp. 632–646.
11. *Sercov V.V., Petrovsky A.A.* An improved speech model with allowance for time-varying pitch harmonic amplitudes and frequencies in low bit-rate MBE coders // Proc. of the 6th European conference on “Speech communication and technology” (Eurospeech'99). Budapest. Hungary. 1999. pp. 1479–1482.
12. *Петровский А.А., Серков В.В.* Низкоскоростной вокодер с моделью речеобразования «гармоники+шум» // Цифровая обработка сигналов. Москва. 2002. №2. С. 61-74.
13. *Udrea R.M., Ciochina S.* Speech enhancement using spectral over-subtraction and residual noise reduction // International Symposium on Signals, Circuits and Systems. 2003. vol. 1. pp. 165–168.
14. *Петровский А.А., Борович А., Парфенюк М.* Дискретное преобразование Фурье с неравномерным частотным разрешением в перцептуальных системах редактирования шума в речи // Речевые технологии. Москва. 2008. №3. С. 16–26.
15. *Borowicz A., Parfieniuk M., Petrovsky A.A.* An application of the warped discrete Fourier transform in the perceptual speech enhancement // Speech Communication. ELSEVIER. 2006. vol. 48. pp. 1024–1036.
16. *Hansen P.S.K., Hansen P.C., Hansen S.D., Sorensen J.A.* Experimental comparison of signal subspace based noise reduction methods // IEEE International Conference on Acoustics, Speech, and Signal Processing. 1999. vol. 1, pp. 101–104.
17. *Borowicz A., Petrovsky A.* Signal subspace approach for psychoacoustically motivated speech enhancement // Speech Communication. Elsevier. 2011. vol. 53. pp. 210–219.
18. *Yu W., Brookes M.* Speech enhancement using a robust Kalman filter post-processor in the modulation domain // ICASSP–2013. 2013. pp.7457–7461.
19. *Bielawski K., Petrovsky A.A.* Speech enhancement system for hands-free telephone based on the psychoacoustically motivated filter bank with allpass frequency transformation // Proc. of the 6th European conference on “Speech communication and technology” (Eurospeech'99). Budapest. Hungary. 1999. pp.2555–2558.
20. *Петровский А.А., Бауун Я.М.* Пре-процессор повышения качества зашумленной и реверберирующей речи для систем улитковой имплантации // Цифровая обработка сигналов. 2002. №2, Москва. С.48-61.
21. *Zorila T.-C., Kandida V., Stylianou Y.* Speech-in-noise intelligibility improvement based on spectral shaping and dynamic range compression // In Proc. Interspeech. Portland. Oregon. 2012. pp. 635–638.
22. *Hinton G.E., Salakhutdinov R.R.* Reducing the Dimensionality of Data with Neural Networks // Science. 2006. vol. 313 no. 786. pp. 504–507.
23. *Arifianto D.* Speech intelligibility improvement of cochlear implant using release of masking // ICACIS–2013. 2013. pp.207–211.
24. *D'Alessandro C., Yegnanarayana B., Darsinos V.* Decomposition of speech signals into deterministic and stochastic components // ICASSP-95. 1995 vol.1. pp. 760–763.

25. *Petrovsky Al., Azarov E., Petrovsky A.* Hybrid signal decomposition based on instantaneous harmonic parameters and perceptually motivated wavelet packets for scalable audio coding // *Signal Processing. Fourier Related Transforms for Non-Stationary Signals.* Elsevier. 2011. vol. 91. Issue 6. pp. 1489–1504.
26. *Kawaahra H., Nisimura R., Irino T., Morise M., Takahashi T., Banno B.* Temporally variable multi-aspect auditory morphing enabling extrapolation without objective and perceptual breakdown // *Proc. ICASSP.* Taipei. Taiwan. 2009. pp: 3905–3908.
27. *Pantazis Y., Stylianou Y.* Improving the modeling of the noise part in the harmonic plus noise model of speech // *Proc. ICASSP–2008.* 2008. pp. 4609–4612.
28. *Azarov E., Vashkevich M., Petrovsky A.* Guslar: a framework for automated singing voice correction // *The 39th International Conference on Acoustics, Speech and Signal Processing (ICASSP 2014).* Florence. Italy. 2014. pp. 7969–7973.
29. *Azarov E., Vashkevich M., Petrovsky A.* Instantaneous pitch estimation based on RAPT framework // *Proc. EUSIPCO'12.* Bucharest. Romania. 2012. pp. 2787–2791.
30. *Nair V., Hinton G.E.* Rectified linear units improve restricted Boltzmann machines // *Proc. ICML.* Haifa. Israel. 2010.
31. *Zeiler M.D., Ranzato M., Monga R., Mao M., Yang K., Le Q.V., Nguyen P., Senior A., Vanhoucke V., Dean J., Hinton G.* On Rectified Linear Units for Speech Processing // *Proc. ICASSP.* Vancouver. Canada. 2013.
32. *Rabiner L.R., Juang B-H.* Fundamentals of speech recognition // Pearson Education. 1993. 507 p.
33. *Осовский С.* Нейронные сети для обработки информации // Москва: "Финансы и статистика". 2002. 344 с.

## References

1. *Watts O., Stan A., Clark R., Mamiya Y., Giurgiu M., Yamagishi J., King S.* Unsupervised and lightly supervised learning for rapid construction of TTS systems in multiple languages from 'found' data: evaluation and analysis. In: *Proc. 8th ISCA Speech Synthesis Workshop.* 2013. pp. 101–106.
2. *Toda T., Black A.W., and Tokuda K.* Voice conversion based on maximum likelihood estimation of spectral parameter trajectory. *IEEE Trans. Audio, Speech and Language Processing.* 2007. vol. 15. no. 8. pp. 2222–2235.
3. *Godoy E., Rosac O., Chonavel T.* Spectral envelope transformation using DFW and amplitude scaling for voice conversion with parallel or nonparallel corpora. *Proc. INTERSPEECH.* Florence. Italy. 2011. pp. 673–676.
4. *Desai S., Black A.W., Yegnanarayana B., Prahallad B.* Spectral mapping using artificial neural networks for voice conversion. *IEEE Trans. Audio, Speech and Language Processing.* 2010. vol. 18. no. 5. pp. 954–964.
5. *Azarov E., Vashkevich M., Likhachov D., Petrovsky A.* Real-time Voice Conversion Using Artificial Neural Networks with Rectified Linear Units. *Proc. INTERSPEECH Lyon.* France. 2013. pp. 1032–1036.
6. *Erro D., Moreno A., Bonafonte A.* INCA Algorithm for Training Voice Conversion Systems From Nonparallel Corpora. *IEEE Transactions on Audio, Speech, and Language Processing.* 2010. vol. 18. no. 5. pp. 944–953.
7. *Yeldener S., De Martin J.C., Viswanathan V.* A mixed sinusoidally excited linear prediction coder at 4 kb/s and below. *Proc. ICASSP'98.* 1998. vol. 2. pp. 589–592.
8. *Boucheron L.E., De Leon P.L., Sandoval S.* Low Bit-Rate Speech Coding Through Quantization of Mel-Frequency Cepstral Coefficients. *IEEE Transactions on Audio, Speech, and Language Processing.* 2012. vol. 20. no. 2. pp. 610–619.
9. *Etmoglu C.O., Cuperman V.* Matching pursuits sinusoidal speech coding. *IEEE Transactions on Speech and Audio Processing.* 2003. vol. 11, no. 5. pp. 413–424.

10. Shlomot E., Cuperman V., Gersho A. Hybrid coding: combined harmonic and wave-form coding of speech at 4 kb/s. *IEEE Transactions on Speech and Audio Processing*. 2001. vol. 9. no. 6. pp. 632–646.
11. Sercov V.V., Petrovsky A.A. An improved speech model with allowance for time-varying pitch harmonic amplitudes and frequencies in low bit-rate MBE coders. Proc. of the 6th European conference on “Speech communication and technology” (Eurospeech’99). Budapest. Hungary. 1999. pp. 1479–1482.
12. Petrovsky A.A., Sercov V.V. [Low – bit rate vocoder with phonation model "harmonics+noise"]. *Cifrovaya obrabotka signalov – Digital signal processing*. 2002. vol. 2. Moscow. pp. 61-74. (In Russ).
13. Udrea R.M., Ciocina S. Speech enhancement using spectral over-subtraction and residual noise reduction. *International Symposium on Signals, Circuits and Systems*. 2003. vol. 1. pp. 165–168.
14. Petrovsky A.A., Borowicz A., Parfieniuk M. [Discrete Fourier transform with non-uniform frequency resolution in perceptual noise reduction systems]. *Recheviye technologii – Speech technologies*. Moscow. 2008. vol. 3. pp. 16–26. (In Russ).
15. Borowicz A., Parfieniuk M., Petrovsky A.A. An application of the warped discrete Fourier transform in the perceptual speech enhancement. *Speech Communication*. ELSEVIER. 2006. vol. 48. pp. 1024–1036.
16. Hansen P.S.K., Hansen P.C., Hansen S.D., Sorensen J.A. Experimental comparison of signal subspace based noise reduction methods. *IEEE International Conference on Acoustics, Speech, and Signal Processing*. 1999. vol. 1, pp. 101–104.
17. Borowich A., Petrovsky. A. Signal subspace approach for psychoacoustically motivated speech enhancement. *Speech Communication*. Elsevier. 2011. vol. 53. pp. 210–219.
18. Yu W., Brookes M. Speech enhancement using a robust Kalman filter post-processor in the modulation domain. *ICASSP–2013*. 2013. pp.7457–7461.
19. Bielawski K., Petrovsky A.A. Speech enhancement system for hands-free telephone based on the psychoacoustically motivated filter bank with allpass frequency transformation. Proc. of the 6th European conference on “Speech communication and technology” (Eurospeech’99). Budapest. Hungary. 1999. pp.2555–2558.
20. Petrovsky A.A., Bashun Y.M. [Pre-processor for quality improvement of noisy and reverberant speech for hearing implants]. *Cifrovaya obrabotka signalov – Digital signal processing*. Moscow. 2002. vol. 2. pp. 48-61. (In Russ).
21. Zorila T.-C., Kandida V., Stylianou Y. Speech-in-noise intelligibility improvement based on spectral shaping and dynamic range compression. In Proc. *Interspeech*. Portland. Oregon. 2012. pp. 635–638.
22. Hinton G.E., Salakhutdinov R.R. Reducing the Dimensionality of Data with Neural Networks. *Science*. 2006. vol. 313 no. 786. pp. 504–507.
23. Arifianto D. Speech intelligibility improvement of cochlear implant using release of masking. *ICACIS–2013*. 2013. pp.207–211.
24. D’Alessandro C., Yegnanarayana B., Darsinos V. Decomposition of speech signals into deterministic and stochastic components. *ICASSP-95*. 1995 vol.1. pp. 760–763.
25. Petrovsky Al., Azarov E., Petrovsky A. Hybrid signal decomposition based on instantaneous harmonic parameters and perceptually motivated wavelet packets for scalable audio coding. *Signal Processing. Fourier Related Transforms for Non-Stationary Signals*. Elsevier. 2011. vol. 91. Issue 6. pp. 1489–1504.
26. Kawaahra H., Nisimura R., Irino T., Morise M., Takahashi T., Banno B. Temporally variable multi-aspect auditory morphing enabling extrapolation without objective and perceptual breakdown. Proc. *ICASSP*. Taipei. Taiwan. 2009. pp: 3905–3908.
27. Pantazis Y., Stylianou Y. Improving the modeling of the noise part in the harmonic plus noise model of speech. Proc. *ICASSP–2008*. 2008. pp. 4609–4612.

28. Azarov E., Vashkevich M., Petrovsky A. Guslar: a framework for automated singing voice correction. The 39th International Conference on Acoustics, Speech and Signal Processing (ICASSP 2014). Florence. Italy. 2014. pp. 7969–7973.
29. Azarov E., Vashkevich M., Petrovsky A. Instantaneous pitch estimation based on RAPT framework. Proc. EUSIPCO'12. Bucharest. Romania. 2012. pp. 2787–2791.
30. Nair V., Hinton G.E. Rectified linear units improve restricted Boltzmann machines. Proc. ICML. Haifa. Israel. 2010.
31. Zeiler M.D., Ranzato M., Monga R., Mao M., Yang K., Le Q.V., Nguyen P., Senior A., Vanhoucke V., Dean J., Hinton G. On Rectified Linear Units for Speech Processing. Proc. ICASSP. Vancouver. Canada. 2013.
32. Rabiner L.R., Juang B-H. Fundamentals of speech recognition. Pearson Education. 1993. 507 p.
33. Osovsky S. Nejronnye seti dlja obrabotki informacii [Neural networks for information processing]. Moskva: "Finansyi i statistika". 2002. 344 p. (In Russ).

**Азаров Илья Сергеевич** — к-т техн. наук, доцент кафедры электронных вычислительных средств БГУИР. Область научных интересов: цифровая обработка речевых сигналов. Число научных публикаций — 42. azarov@bsuir.by, www.bsuir.by; БГУИР, ул. П. Бровки 6, г. Минск, 220013, РБ; р.т. +375 (17) 293-8805.

**Azarov Elias** — Ph.D., associate professor of computer engineering department, BSUIR. Research interests: digital speech processing. The number of publications — 40. azarov@bsuir.by, www.bsuir.by; BSUIR, 6, P.Brovky str., 220013, Minsk, RB; office phone +375 (17) 293-8805.

**Петровский Александр Александрович** — д-р техн. наук, профессор, заведующий кафедрой электронных вычислительных средств БГУИР. Область научных интересов: цифровая обработка сигналов. Число научных публикаций — более 600. palex@bsuir.by, www.bsuir.by; БГУИР, ул. П. Бровки 6, г. Минск, 220013, РБ; р.т. +375 (17) 293-2340.

**Petrovsky Alexander** — Ph.D., Dr. Sci., professor, head of computer engineering department, BSUIR. Research interests: digital speech processing. The number of publications — more than 600. palex@bsuir.by, www.bsuir.by; BSUIR, 6, P.Brovky str., 220013, Minsk, RB; office phone +375 (17) 293-2340.

## РЕФЕРАТ

*Азаров И.С., Петровский А.А.* **Формирование персональной модели голоса диктора с универсальным фонетическим пространством признаков на основе искусственной нейронной сети.**

В работе исследуется возможность формирования модели голоса заданного диктора на основе записей образцов его голоса с транскрипцией. Необходимость решения данной задачи возникает во многих речевых приложениях таких как конверсия голоса, коррекция акцента, синтез речи по тексту, кодирование и др.

Разработанная схема моделирования голоса основывается на нейронной сети с кусочно-линейной функцией активации, построенной по принципу автоматического кодера. В результате преобразования характеристических векторов речи сетью, они представляются в виде кодов пониженной размерности с фонетической привязкой.

В практической части работы предлагаются результаты экспериментов применения голосовой модели к задаче конверсии голоса с текстонезависимым обучением.

## SUMMARY

*Azarov E., Petrovsky A.* **Training personal voice model of a speaker with unified phonetic space of features using artificial neural network.**

The paper investigates possibility of training a personal voice model for given speaker using transcribed speech samples. Solution to the problem is required in many speech processing applications such as voice conversion, accent correction, text-to-speech synthesis, coding and other.

The proposed voice modeling scheme is based on neural network with rectified-linear units designed as deep autoencoder. Characteristic speech vectors are transformed by the network into low-dimensional codes with phonetic alignment.

In the practical part of the work some experimental results are given, that show applicability of the model to voice conversion using non-parallel training sets.

Е.В. ГЕНИАТУЛИНА, М.Г. Гриф  
**МЕТОДЫ ПРОЕКТИРОВАНИЯ И МОДЕЛИРОВАНИЯ В  
ЗАДАЧАХ ОПТИМИЗАЦИИ ПРОЦЕССОВ  
ФУНКЦИОНИРОВАНИЯ ЧЕЛОВЕКО-МАШИННЫХ СИСТЕМ**

*Гениатулина Е.В., Гриф М.Г. Методы проектирования и моделирования в задачах оптимизации процессов функционирования человеко-машинных систем.*

**Аннотация.** В статье рассмотрены модели и методы автоматизации проектирования процессов функционирования человеко-машинных систем (ПФ ЧМС) на основе функционально-структурной теории (ФСТ) человеко-машинных систем и обобщенного структурного метода проф. А.И. Губинского. Приведены основные понятия и определения ФСТ и методов последовательной оптимизации ПФ ЧМС. В статье предлагаются новые методы и алгоритмы, направленные на расширение области применения ФСТ ЧМС А.И. Губинского в оптимальном проектировании ПФ ЧМС. Предложен способ расширения области использования ФСТ ЧМС за счет интеграции оптимизационных моделей ПФ ЧМС с методом имитационного моделирования, так как ФСТ можно применять только для процессов без последействия и при отсутствии зависимых операций. Предлагается способ устранения данного ограничения путем интеграции технологии проектирования ПФ ЧМС основе ФСТ с методом имитационного моделирования тех участков процесса, для которых не выполняются указанные выше требования ФСТ. Описан алгоритм генерации последовательно-параллельных соединений операций с учетом дополнительных ограничений, алгоритм генерации альтернативных вариантов ПФ ЧМС на основе совпадения целей операций, алгоритм генерации параметрических альтернатив на основе шаблона. Приведены основные понятия и определения, необходимые для алгоритма генерации фрагментов процесса с учетом обязательных сочетаний операций. Предложено использование матрицы обязательных сочетаний операций, в которой ненулевые элементы строк имеют смысл единственно возможных сочетаний способов выполнения соответствующих типовых функциональных единиц в альтернативах. Введено понятие состав и понятие парной несовместимости составов, на основе которых происходит распределение выполняемых функций составами.

**Ключевые слова:** функционально-структурная теория, человеко-машинная система, множество альтернатив, альтернативный граф, функциональная сеть, цель операции, обязательные сочетания операций, парная несовместимость, автоматизация проектирования, имитационное моделирование.

*Geniatulina E.V., Grif M.G. Man-machine Systems Functioning Design and Modeling Methods in Optimization Processes Problems.*

**Abstract.** The article is dedicated to the models and methods of automation man-machine systems functioning processes design based on functional-structural theory of man-machine systems and generalized structural method of prof. A.I. Gubinsky. The basic concepts and definitions of the functional-structural theory and sequential optimization method of man-machine systems functioning processes are presented. They are intended to expand the scope of functional-structural theory of prof. A.I. Gubinsky in man-machine systems functioning processes optimal design. The functional-structural theory is applicable only for processes without aftereffect and in the case of dependent operations absence. That is why the integration of man-machine systems functioning processes optimization models with simulation method is proposed. Also the man-machine systems functioning processes integration design technology method based on the functional-structural theory with simulation method for the areas that are

not fulfilled the above requirements of the functional-structural theory is researched. The proposed methods allow to bypass limitations of functional-structural theory. The algorithm for generating a series-parallel connection operations with the additional constraints, the algorithm generating alternatives process of man-machine systems functioning processes based on the coincidence of the objectives of operations, the algorithm generating parametric alternatives based on a template are described. The basic concepts and definitions necessary for the generation algorithm of process fragments with the mandatory combinations of operations are represented. The using of binding combinations operations matrix, in which the nonzero elements of rows have meaning of the only possible method combinations for performing the relevant functional units in alternatives, is proposed. The performed functions distribution is based on the concept of the composition and the concept of compositions pair incompatibility.

**Keywords:** functional-structural theory, man-machine system, the set of alternatives, the alternative graph, functional network, the operation target binding combination of operations, pair incompatibility, design automation.

---

**1. Введение.** Разработка технологий проектирования процессов функционирования человеко-машинных систем по показателям эффективности, качества и надежности (ЭКН) является одним из доминирующих направлений в исследовании и автоматизации проектных работ, управления объектами и принятия решений. Существует множество различных подходов к оптимальному проектированию процессов функционирования ЧМС, которые обеспечивают возможность генерации и быстрого анализа достаточно большого числа альтернатив [1,2]. Одним из хорошо зарекомендовавших себя на практике является функционально-структурная теория и обобщенный структурный метод проф. Губинского А.И. [3,4], получившие творческое развитие в работах В.Г. Евграфова, П.И. Падерно, А.П. Ротштейна, П.П. Чабаненко, Е.А. Лаврова и др. [7-11].

В работах Е.Б. Цоя и М.Г. Грифа [5,6] разработаны модели, методы и технологии последовательной оптимизации процессов функционирования (ПФ) ЧМС по показателям эффективности, качества и надежности на основе ФСТ. Однако применяемый в них способ задания множества альтернатив ПФ ЧМС требует явного описания процессов и достаточно трудоемок. Кроме того, механизм задания ограничений на сочетании способов выполнения отдельных операций в ПФ представляет для проектировщика большую сложность.

В настоящей работе предлагается способ расширения области использования ФСТ ЧМС за счет интеграции оптимизационных моделей ПФ ЧМС с методом имитационного моделирования, так как ФСТ можно применять только для процессов без последствия и при отсутствии зависимых операций. Описывается способ устранения данного ограничения путем интеграции технологии проектирования ПФ ЧМС основе ФСТ с методом имитационного моделирования тех участков процесса, для которых не выполняются указанные выше требования ФСТ.

Предлагаются также новые методы оптимального проектирования ПФ ЧМС. Описываются и обосновываются подходы к заданию множества альтернатив ПФ ЧМС с использованием модели данных на основе объектно-ориентированного представления операций, которое строится путем синтеза функционально-структурной теории и объектно-ориентированного проектирования. Рассматриваются алгоритм генерации последовательно-параллельных процессов, алгоритм генерации альтернативных фрагментов ПФ ЧМС на базе совпадения целей процессов, а также алгоритм с учетом обязательных сочетаний операций, которые реализованы в гибридной экспертной системе (ГЭС) проектирования человеко-машинных систем и принятия решений – ИНТЕЛЛЕКТ-3.

**2. Постановка задачи оптимального проектирования ПФ ЧМС.** Использование при проектировании ПФ ЧМС ФСТ А.И. Губинского предполагает, что каждый альтернативный процесс функционирования ЧМС задается в виде функциональной сети (ФС) и представляется состоящим из ряда формализованных единиц – типовых функциональных единиц (ТФЕ) и их типовых комбинации – типовых функциональных структур (ТФС). Причем множество альтернатив может содержать как структурные, так и параметрические альтернативы (рисунк 1) [3-6].

Непосредственно в основе способа оценки вероятностных показателей эффективности, качества и надежности процесса функционирования ЧМС – вероятности правильного (безошибочного) выполнения  $B$ , среднего времени  $T$  и средних затрат (дохода)  $V$  от выполнения, вероятности своевременного выполнения  $P(t < T_d)$ , лежит вероятностный граф, а также правила его редукции (укрупнения).

Рассмотрим способы задания оптимизационной модели ПФ ЧМС на основе функциональной сети с использованием множеств элементов ЧМС, выполняемых функций и операций. Множество альтернативных процессов ЧМС, пользователь описывает в виде альтернативного графа (рисунк 2).

Отдельный процесс функционирования должен быть представлен в виде суперпозиции ТФС, в которые входят наборы операций. Под операцией  $O \equiv O(F, E, Q)$  понимается процесс выполнения функции  $F$  элементом  $E$  в состоянии ЧМС  $S$ ,  $Q$  - показатели эффективности, качества и надежности. Отдельный процесс функционирования ЧМС (функциональная сеть) представляется в виде суперпозиции ТФС:

$$O_z = TFC_i(O_{i_1}, O_{i_2}, \dots, O_{i_k}), \quad (1)$$

где  $T\Phi C_i \in M_{T\Phi C}$ ,  $O_{i_j}$  – простая или составная операция. Две операции с совпадающей функцией  $F - O(F, E_1, Q_1)$  и  $O(F, E_2, Q_2)$  являются альтернативными ("параметрическими") способами выполнения операции  $O$ , так же как и составные операции  $O = T\Phi C_i(O_{i_1}, O_{i_2}, \dots)$  и  $O = T\Phi C_s(O_{s_1}, O_{s_2}, \dots)$ ,  $i \neq s$  – "структурными".

Задача оптимизации (обобщенная задача динамического программирования) ставится следующим образом:

$$\begin{aligned} K_{ЭКН}(A) &\rightarrow \text{extr}, \\ A &\in M_d \subseteq M_a \end{aligned} \quad (2)$$

где  $K_{ЭКН}(A)$  – критерий оптимальности для сочетаний критериев ЭКН;  $M_d$  – множество допустимых альтернатив, альтернативные варианты процесса –  $M_a$ .

Вероятностные и нечеткие показатели эффективности, качества и надежности процесса (алгоритма) функционирования:  $B(A), T(A), V(A), P(t < T_d)(A), \tilde{B}(A), \tilde{T}(A), \tilde{V}(A)$ .

$B(A)$  – вероятность правильного (безошибочного) выполнения,  $T(A)$  – среднее время и  $V(A)$  – средние затраты или доход в зависимости от поставленной задачи.

В табл. 1 приведены некоторые из возможных постановок задач оптимизации с показателями  $B(A), T(A)$  и  $V(A)$ . Здесь  $P_{ogr}(A)$  – ограничения на совместимость способов выполнения компонентов альтернатив в виде предиката – "Если  $P_{ogr}(A)$  есть "Истина", то  $A$  удовлетворяет ограничениям задачи".

В [5] разработан метод последовательной оптимизации ПФ ЧМС на модели ФС в рамках общей схемы метода последовательного анализа вариантов с пошаговым конструированием частичных решений. Конкретный алгоритм пошагового конструирования определяется правилом выбора частичных решений (подсетей)  $\vartheta$ , подлежащих развитию на каждом шаге, и набор тестов  $\xi$ , осуществляющих отсев тех из них, которые не могут быть построены до оптимальных. Вариация параметров  $\vartheta$  и  $\xi$  приводит к различным алгоритмам метода последовательного анализа вариантов применительно к задачам оптимизации ПФ ЧМС на функциональных сетях.

Таблица 1. Скалярные и векторные задачи оптимизации с показателями  $B(A)$ ,  $T(A)$  и  $V(A)$

Критерий	Ограничения	Примечания
$B(A) \rightarrow \max,$ $A \in M_d$	$B(A) \leq B_d,$ $T(A) \leq T_d,$ $P_{ogr}(A)$	$V_d \geq 0,$ $T_d \geq 0$
$T(A) \rightarrow \min,$ $A \in M_d$	$V(A) \leq V_d,$ $B(A) \geq B_d,$ $P_{ogr}(A)$	$V_d \geq 0,$ $B_d \in [0.1]$
$V(A) \rightarrow \min,$ $A \in M_d$	$T(A) \leq T_d,$ $B(A) \geq B_d,$ $P_{ogr}(A)$	$T_d \geq 0,$ $B_d \in [0.1]$
$F_1 = c_1 B(A) - c_2 T(A)$ $-c_3 V(A) \rightarrow \max$	$P_{ogr}(A)$	$c_i \geq 0, i=1...3$

Технология проектирования ПФ по показателям ЭКН наиболее полно реализована в гибридной экспертной системе ГЭС Интеллект-3, которая предназначена для автоматизации проектирования ПФ ЧМС на основе вероятностных и нечетких показателей ЭКН методом последовательной оптимизации.

**3. Построение множества альтернатив на основе объектно-ориентированной модели данных.** Ранее пользователю было необходимо задавать процесс функционирования в явном виде. Это достаточно трудоемкий процесс, который требует от пользователя обладания определенными навыками. Предлагается автоматическая генерация всего ПФ за счет алгоритма генерации последовательно-параллельного ПФ ЧМС, где  $M = \{O_1, O_2, \dots, O_n\}$  – множество абстрактных операций, отличающихся друг от друга уникальным именем (или номером),  $n$  – общее количество операций. В основе данного алгоритма лежат такие понятия как множество абстрактных операций, а также бинарные отношения следования операций.

$R = \{ \langle O_i, O_j \rangle / O_i R O_j \Leftrightarrow O_i \neq O_j; O_i, O_j \in M \}, R \subseteq M^2$  – бинарное отношение на множестве  $M^2$  «следует за» (за операцией  $O_i$  следует операция  $O_j$ ), задает отношение последовательности на множестве  $M$ .

Для определения последовательно – параллельной структуры (алгоритм) выполнения операций процесса функционирования человеко-машинных систем, необходимо разбить множество  $M$  на  $m$  попарно не пересекающихся подмножеств  $P_i$  таких, что каждый элемент из  $M$  принадлежит только одному из этих подмножеств[12-14]:

$$P = P_1, P_2, \dots, P_m, \quad (3)$$

где  $m$  – количество подмножеств множества  $P$ .

Условия распараллеливания, в соответствии с которыми необходимо максимально распараллелить алгоритм (3), но с минимально возможным значением  $m$  (удовлетворяющим от 1) до б):

1)  $\forall i P_i \subseteq M$  – все подмножества  $P$  входят во множество абстрактных операций  $M$ ;

2)  $\forall i \forall j i \neq j P_i \cap P_j = \emptyset$  – одна и та же операция не должна входить одновременно в два различных подмножества  $P$ ;

3)  $\bigcup_{i=1, m} P_i = M$  – все операции из множества  $M$  должны быть учтены;

4) Для мощности  $\forall O_{s_1} \in P_j$  и  $\forall O_{s_2} \in P_i$ , и  $j < i, i \neq j \Rightarrow \exists O_{s_1} R O_{s_2}$  – должна обязательно существовать начальная операция (та, которая не следует ни за какой другой);

5) Условие параллельности 1:

Для мощности  $|P_i| > 1: \forall O_{s_1}$  и  $\forall O_{s_2} \in P_i \Rightarrow \exists O_{s_1} R O_{s_2}$  – для любых двух пар операций, не должно существовать отношений следования друг за другом;

6) Условие последовательности:

$\forall O_{s_1} \in P_j$  и  $\forall O_{s_2} \in P_i$  и  $j < i, i \neq j \Rightarrow \exists O_{s_1} R O_{s_2}$  – операция  $O_{s_1}$  следует за операцией  $O_{s_2}$ , только в том случае, если не существует отношения «следования за»  $- O_{s_1} R O_{s_2}$ .

В результате получаем последовательно-параллельную структуру вида (рисунок 1).

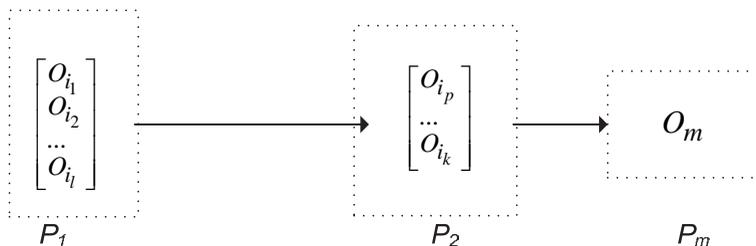


Рис. 1. Последовательно-параллельная структура

Отношения следования «следует за»  $R$  задаются посредством матрицы смежности:

$$A = \|a_{ij}\|, \text{ где } i, j = \overline{1, n} \text{ и } a_{ij} = 1 \Leftrightarrow \exists O_i R O_j.$$

Исходя из которой, находится матрица достижимости:

$$A' = \|a'_{ij}\|, \text{ где } i, j = \overline{1, n}.$$

Вводимые пользователем отношения следования, в конечном итоге, представляются в виде списков:  $S_i = \{O_{s_1}, \dots, O_{s_m} \mid \forall O_s \in M / a'_{is} = 1\}$

или  $S_i = \{O_{s_1}, \dots, O_{s_m} \mid \forall O_s \in M \exists Put(O_i, O_s)\}$ .

$Put(O_i, O_j)$  – существование достижимости из  $O_i$  в  $O_j$ , определяется путем учета явных или неявных отношений «следования за» –  $O_i R O_j$ .

Расширением алгоритма последовательно-параллельных ПФ является алгоритм автоматической генерации параметрических способов выполнения операций на базе объектно-ориентированной модели представления данных, элементами которой являются классы, объекты, отношения (наследования, агрегации, ассоциации, метакласс), свойства, методы и т.п. В рамках данного способа происходит описание операций, множества альтернатив и методов его модификации [13-17]. В качестве способа описания введено понятие «состав».

Состав – совокупность объектов разных классов, объединенных для достижения поставленных задач и обладающий критериями ЭКН.

$St_i \equiv (O_1, O_2, \dots, O_n, B, T, V)$ , где  $St_i$  – состав,  $O_1, O_2, \dots, O_n$  – операции.

В общем случае во множество параметрических способов выполнения операции входит совокупность составов, целью каждого из которых является результат выполнения необходимой операции:  $M_p(O) \equiv \{St_1, St_2, \dots, St_n\}$ , где  $M_p(O)$  – множества параметрических способов выполнения операции.

Разработан алгоритм генерации параметрических альтернатив на основе меры близости составов. Он позволяет автоматически подобрать наиболее близкие способы выполнения операций в рамках той или иной ТФЕ на основе меры близости составов.

Мера близости  $\nu$  – значение в диапазоне от 0 до 1, по которому определяется степень "похожести", близости между объектами (или группами объектов) и между их атрибутами [19].

*Определение.* Состав  $St_i$  является близким составу  $St_j$  тогда, когда мера близости между этими составами  $\nu(St_i, St_j) > 0$ :

$$\nu(St_i, St_j) = \frac{N_s}{N_o},$$

где  $N_s$  – количество схожих объектов состава  $St_i$  по отношению к составу  $St_j - K_{O_s}$ , а также соответствующих им характеристик (атрибутов) –  $K_{C_s}$ :  $N_s = K_{O_s} + K_{C_s}$ ,

$N_o$  – общее количество объектов  $K_{O_o}$  и характеристик (со значениями, если они есть)  $K_{C_o}$  состава  $St_j$ :  $N_o = K_{O_o} + K_{C_o}$ .

$N_s$  – определяется в соответствии с объектами и их характеристиками, т.е. проверяется соответствие конкретных объектов одного состава с объектами другого состава.

$N_o$  – количество объектов одного состава, значения характеристик которых совпадают со значениями характеристик другого состава.

Использование меры близости позволяет определить уровень сходства составов объектов, для более полного понимания пользователем схожести тех или иных составов, помогает в выборе близких составов в качестве альтернатив в рамках ТФЕ.

Разработан метод генерации структурных альтернатив на основе деления на фрагменты ПФ. Две операции являются структурными альтернативами  $O = TFC_i(O_{i_1}, O_{i_2}, \dots)$  и  $O = TFC_s(O_{s_1}, O_{s_2}, \dots)$ ,  $i \neq s$  если при выполнении операций в рамках этих ТФС достигается необходи-

мый результат или, иначе говоря, операции направлены на достижении одной и той же заданной цели.

Цель – это желаемое состояние процесса, которое достигается путем перехода из одного состояния системы в другое посредством операций, входящих в конкретные фрагменты процесса.

В рамках ПФ существует множество целей:  $Z = \{Z_1, Z_2, \dots, Z_n\}$ , где  $Z_i = Z_i(TФE_i, \dots, TФE_j)$  – множество ТФЕ, объединенных для достижения единой цели.

$Z_i(TФE_i, \dots, TФE_j)$  – будем называть цель от точки  $TФE_i$  до точки  $TФE_j$ .

Ранее цель задавалась только для конкретной ТФЕ исходя из выполняемой операции, и не давала пользователю никакой возможности сгенерировать структурную альтернативу. Разработанный метод позволяет пользователю создавать цели на любых участках ПФ и сохранять их, что позволяет пользователю автоматически получить представление процесса или отдельной его части - фрагмента в виде альтернативного графа на основе ранее сохраненных целей фрагментов. Необходимо лишь поставить цель для новой части ПФ, система сама сгенерирует возможные варианты решения.

Алгоритм генерации параметрических альтернатив может расширить алгоритм последовательно-параллельного алгоритма и дать возможность пользователю не только сформировать весь алгоритм с учетом внешних факторов, а также подобрать возможные параметрические альтернативы для каждой ТФЕ в рамках графа всего процесса.

**4. Оптимизация с учетом обязательных и недопустимых сочетаний операций.** Ранее была предложена только матрица недопустимых сочетаний операций. Она заполнялась вручную пользователем, что сопровождалось поиском номеров способов выполнения операций в рамках ТФЕ, причем некоторые из них не были исходными.

Предложено автоматизировать заполнение матрицы недопустимых сочетаний операций за счет ввода понятия списка парной несовместимости составов. Все данные по несовместимости хранятся в системе и не требуют повторного ввода.

Рассмотрим пример понятия ограничений на совместимость способов выполнения ТФЕ с учетом недопустимых сочетаний операций. Пусть набором текста на ЭВМ могут заниматься два оператора - А и В (рисунок 2.). В данной ФС только три операции, следовательно, имеем восемь альтернатив.

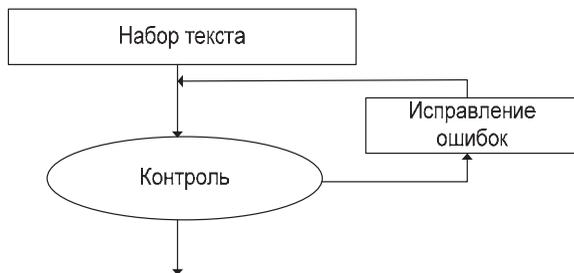


Рис. 2. Функциональная сеть набора текста оператором ЭВМ

Однако реальных альтернатив только две – назначить на все три операции либо  $A$ , либо  $B$ . Следовательно, нужно задать матрицу:

$$M_{SOV} = \begin{Bmatrix} 1 & 2 & 0 \\ 1 & 0 & 2 \\ 2 & 1 & 0 \\ 2 & 0 & 1 \end{Bmatrix} \text{ или } M_{ZSOV} = \begin{Bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \end{Bmatrix}.$$

Для упрощения ввода и хранения информации, используемой ранее, а также для автоматического формирования матрицы  $M_{SOV}$ , введем понятие списка по парной несовместимости составов:

$$Sp \equiv (St_i, St_j), \text{ где } St_i, St_j \equiv (O_1, O_2, \dots, O_n, B, T, V).$$

На основе приведенного примера, достаточно сказать что  $A$  и  $B$  не совместимы, чтобы получить автоматическую матрицу  $M_{SOV}$  приведенную выше, тогда как, вводя матрицу самостоятельно, пользователю необходимо помнить какой номер у способа выполнения операции в определенной ТФС, что заставляет пользователя тратить больше времени на ввод матрицы ограничений на совместимость.

Таким образом, получаем:

$Sp \equiv (A, B)$  - означает, что состав  $A$  и состав  $B$  – несовместимы.

В результате, на основе примера получается матрица несовместимых сочетаний способов выполнения операций:

$$M_{SOV} = \begin{Bmatrix} 1 & 2 & 0 \\ 1 & 0 & 2 \\ 2 & 1 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{Bmatrix}.$$

Поскольку просмотр графа последовательный, и сравнение попарно, добавляется в матрицу еще 2 строки  $\{0,2,1\}$  и  $\{0,1,2\}$ . Они являются лишними, поскольку сочетания  $\{1,2,0\}$  и  $\{1,0,2\}$  уже учитывают  $\{0,2,1\}$  и  $\{0,2,1\}$ .

В связи этим встает необходимость поиска взаимозаменяющих строк, который предполагает поиск лишних строк и удаление их из матрицы несовместимых способов сочетания операций для более быстрого его анализа, что сокращает время работы всего алгоритма оптимизации в целом.

Бывают ситуации, когда пользователю проще задать те сочетания способов выполнения операций, которые обязательно должны быть. В связи с этим, предложено использование матрицы обязательных сочетаний операций  $M_{ZSOV}$ , в которой ненулевые элементы строк имеют смысл единственно возможных сочетаний способов выполнения соответствующих ТФЕ в альтернативах.

В той части алгоритма оптимизации, где происходит отсев частичных решений исходя их введенных матриц недопустимых и обязательных сочетаний операций, время нахождения решения увеличивается в зависимости от количества введенных в них строк, поскольку решение об отсеке принимается на основе последовательного анализа каждой матрицы.

Таким образом, можно сделать вывод о том, что алгоритм оптимального проектирования процессов функционирования ЧМС с учетом недопустимых и обязательных сочетаний операций, позволяет снизить время решения задачи по сравнению с методом поиска на основе недопустимых сочетаний операций при большем их количестве [17].

**5. Интеграция оптимизационных моделей ПФ ЧМС с методом имитационного моделирования.** В [6] была впервые предложена схема интеграции метода оптимального проектирования ПФ ЧМС на основе ФСТ с методом имитационного моделирования показателей эффективности, качества и надежности отдельных участков процесса, которые нельзя описать функциональной сетью.

Суть предлагаемого подхода состоит в следующем [18, 19].

1. Множество альтернатив ПФ ЧМС описывается одной из популярных и наглядных нотаций описания бизнес-процессов – EPC. Данная нотация дополняется специальным блоком для задания альтернативных участков ПФ ЧМС.

2. В каждом блоке, аналогичном ТФЕ в ФС, задаются вероятности исходов выполнения (правильно, ошибка) как функции от различных факторов, например, освещенность рабочего места, квалифика-

ция работника и т.п. Кроме того, задаются и формулы для вычисления (изменения) значений факторов в процессе выполнения данного блока.

3. Проводится анализ множества альтернатив ПФ ЧМС на предмет выделения в нем участков, которые нельзя описывать моделью ФС. Так, например, если для линейного участка процесса находится фактор, влияющий на исходы первой и последней операции, то весь участок нельзя описывать ФС. Или для циклического участка процесса выясняется, что фактор на входе в цикл может изменяться внутри цикла.

4. Для выделенных участков процесса, не представимых ФС, проводится имитационное моделирование показателей эквивалентного блока –  $B(A)$ ,  $T(A)$  и  $V(A)$ , который и заменит данный участок. Необходимо заметить, что данный блок, как правило, является альтернативным.

5. После того как во множестве альтернатив ПФ ЧМС, описанного в системе нотаций ЕРС, не останется участков непредставимых ФС, производится трансляция нотаций ЕРС в ФС.

Решения задач оптимального проектирования ПФ ЧМС на модели ФС осуществляется в ГЭС ИНТЕЛЛЕКТ-3. Программный комплекс имитационного моделирования ПФ ЧМС разработан для операционной системы Windows, написан на языке C# с использованием библиотеки Microsoft .NETFramework 2.0 и компонента построения диаграмм MicrosoftVisio. Применяется совместно с ГЭС ИНТЕЛЛЕКТ-3.

**6. Заключение.** Все приведенные методы и алгоритмы направлены на расширение области применения ФСТ ЧМС А.И. Губинского в оптимальном проектировании ПФ ЧМС. Так, алгоритм генерации последовательно-параллельных процессов функционирования ЧМС позволяет автоматически сформировать ПФ с учетом максимального распараллеливания операций. Это уменьшает время, затрачиваемое на рассматриваемый пользователем процесс, позволяет учесть влияние факторов внешней среды, сформировать возможные альтернативы в рамках той или иной операции, исключая ситуации распараллеливания операций, в том случае, если операции выполняются в различных условиях. Кроме того, возможна автоматическая генерация частей ПФ на основе алгоритма деления на фрагменты, который позволяет учитывать и хранить фрагменты ПФ на основе заданной цели.

Описанный метод генерации процессов функционирования на основе обязательных сочетаний операций позволяет пользователю задавать быстро и удобно альтернативы, имеющие ограничения на совместимость или же являющиеся обязательными для того или иного процесса, что ускоряет время работы алгоритма в целом. Данные ме-

тоды программно реализованы в ГЭС ИНТЕЛЛЕКТ-3 с учетом всех описанных особенностей.

Интеграция ФСТ с методом имитационного моделирования позволяет проводить оптимизацию ПФ ЧМС в случае, когда имеют место процессы с последействием, что также расширяет возможности модели ФСТ ЧМС.

### Литература

1. *MacGarry K., Wermter S., MacIntyre J.* Hybrid neural system: from simply coupling to fully integrated neural network // *Neural computing surveys*. 1999. vol. 2. pp. 62-93.
2. *Поспелов Г.С.* Искусственный интеллект – основа новой информационной технологии // М.: Наука. 1988. 280с.
3. *Адаменко А.Н.* Информационно-управляющие человеко-машинные системы: Исследование, проектирование, испытания: Справочник // М.: Машиностроение, 1993. 528 с.
4. *Губинский А.И.* Надежность и качество функционирования эргатических систем // Л.: Наука, 1982. 270 с.
5. *Гриф М.Г. Цой Е.Б.* Автоматизация проектирования процессов функционирования человеко-машинных систем на основе метода последовательной оптимизации: монография // Новосибирск. НГТУ. 2005. 264 с.
6. *Гриф М.Г., Никитюк А.А.* Проектирование и оптимизация бизнес-процессов на основе аппарата функциональных сетей // *Информатика: проблемы, методология, технологии* : материалы 11 междунар. науч.-метод. конф. Воронеж: ИПЦ ВГУ. 2011. Т. 1. С. 209–212.
7. *Падерно П.И.* Функционально-структурная теория – направления развития // *Человеческий фактор: проблемы психологии и эргономики*. 2011. № 3-1. С. 89–90.
8. *Падерно П.И., Смирнов А.В.* Оценка безошибочности выполнения фрагментов алгоритмов при различных видах ошибок // *Известия СПбГЭТУ "ЛЭТИ"*. 2012. № 2. С. 38–45.
9. *Ротштейн А. П.* Алгебра алгоритмов и нечеткая логика в анализе надежности систем // *Известия РАН. Теория и системы управления*. 2010. № 2. С. 91–103.
10. *Лавров Е. А., Пасько Н. Б.* Подход к формированию банка оптимизационных моделей для распределения функций между операторами АСУ // *Восточно-Европейский журнал передовых технологий*. Серия «Математика и кибернетика – фундаментальные и прикладные аспекты». 2011. № 1/4 (49). С. 46–50.
11. *Чабаненко П. П.* Исследование безопасности и эффективности функционирования систем человек-техника эргосетями // Севастополь: Изд-во Акад. воен.-мор. сил им. П. С. Нахимова. 2012. 160 с.
12. *Гениатулина Е.В., Гриф М.Г.* Метод генерации процессов функционирования человеко-машинных систем // *Сборник научных трудов НГТУ*. Новосибирск: Изд-во НГТУ. 2009. №2(56). С. 71–76.
13. *Гениатулина Е.В., Гриф М.Г.* Методы генерации множества альтернатив в задачах оптимизации человеко-машинных систем // *Науч. вест. НГТУ*. Новосибирск: Изд-во НГТУ. 2010. №4(41). С. 41–50.
14. *Гениатулина Е.В.* Проектирование процессов функционирования в человеко-машинных системах // *Перспективы развития информационных технологий*. Сборник материалов II Ежегодной Всероссийской научно-практической конференции с международным участием г. Новосибирск: Изд-во «Сибпринт». 2010. С.205-210.

15. *Гениатулина Е.В.* Метод генерации процессов функционирования человеко-машинных систем в интеллектуальных системах // «Молодой ученый». Чита. ООО «Издательство Молодой ученый». 2009. №10. С.83 – 85.
16. *Гениатулина Е.В.* Представление данных в интеллектуальных системах, усеченно-естественный язык // Материалы девятой международной научно-методической конференции. Информатика: проблемы, методология, технологии. Воронеж. Воронежский государственный университет. 2009. Т. 1. С. 202-210.
17. *Гениатулина Е.В., Гриф М.Г.* Метод генерации фрагментов процесса функционирования с учетом обязательных сочетаний операций // Сборник научных трудов Новосибирского государственного технического университета. 2014. №3 (77). С. 53-62.
18. *Grif M.G., Sundui O., Tsoy E.B.* Methods of designing and modeling of man-machine systems // Proc. of International Summerworkshop Computer Science 2014. Germany. Chemnitz. 2014. pp. 38-40.
19. *Grif M.G., Sundui O., Tsoy E.B.* Selecting the method of maintenance and elimination of network failures in the National University of Mongolia // News of Science and Education. 2014. vol. 14(14). pp. 116-123.

## References

1. MacGarry K., Wermter S., MacIntyre J. Hybrid neural system: from simply coupling to fully integrated neural network. Neural computing surveys. 1999. vol. 2. pp. 62–93.
2. Pospelov G.S. *Iskusstvennyi intellekt – osnova novoi informatsionnoi tekhnologii* [Artificial Intelligence - the basis of new information technology]. M.:Nauka. 1988. 280 p. (In Russ.).
3. Adamenko A.N. *Informatsionno-upravliaiushchie cheloveko-mashinnye sistemy: Issledovanie, proektirovanie, ispytaniya: Spravochnik* [Information and Control Man-Machine Systems: Research, design, testing; test handbook] Moscow: Mashinostroenie. 1993. 528 p. (In Russ.).
4. Gubinskii A.I. *Nadezhnost' i kachestvo funktsionirovaniya ergaticheskikh sistem* [Reliability and quality of operation ergonomics systems]. L.: Nauka, 1982. 270 p. (In Russ.).
5. Grif M.G. Tsoi E.B. *Avtomatizatsiya proektirovaniya protsessov funktsionirovaniya cheloveko-mashinnykh sistem na osnove metoda posledovatel'noi optimizatsii: monografiya* [Computer-aided design processes of man-machine systems on the basis of serial optimization: a monograph]. Novosibirsk, NSTU. 2005. 264 p. (In Russ.).
6. Grif M.G., Nikitiuk A.A. [Design and optimization of business processes on the basis of the device functional networks]. *Informatika: problemy, metodologiya, tekhnologii: materialy 11 mezhdunar. nauch.-metod. konf* [Materials of 11th International Scientific Conference]. Voronezh: IPTS VGU. 2011. vol. 1. pp. 209–212. (In Russ.).
7. Paderno P.I. [Functional- structural theory - development directions]. *Chelovecheskii faktor: problemy psikhologii i ergonomiki – Human Factors: Problems of Psychology and Ergonomics*. 2011. vol. 3-1. pp. 89–90. (In Russ.).
8. Paderno P.I., Smirnov A.V. [Evaluation of infallibility perform fragments of algorithms for different types of errors]. *Izvestiia SPbGETU "LETI" – Proceedings of Saint-Petersburg Electrotechnical University "LETI"*. 2012. vol. 2. pp. 38–45. (In Russ.).
9. Rotshtein A. P. [Algebra algorithms and fuzzy logic to analyze the reliability of systems]. *Izvestiia RAN. Teoriia i sistemy upravleniia – Journal of Computer and Systems Sciences International*. 2010. vol. 2. pp. 91–103. (In Russ.).
10. Lavrov E. A., Pas'ko N. B. [Approach to the formation of the bank's optimization models for the distribution of functions between operators ACS]. *Vostochno-Evropeskii zhurnal peredovykh tekhnologii. Seriya «Matematika i kibernetika - fundamental'nye i prikladnye aspekty» – East European Journal of advanced technologies*. 2011. vol. 1/4 (49). pp. 46–50. (In Russ.).
11. Chabanenko P. P. *Issledovanie bezopasnosti i effektivnosti funktsionirovaniia sistem chelovek-tekhnika ergosetiami* [Study of the safety and effectiveness of the systems of

- human- technology ergonomic networks]. Sevastopol': Izd-vo Akad. voen.-mor. sil im. P. S. Nakhimova. 2012. 160 p. (In Russ.).
12. Geniatulina E.V., Grif M.G. [The method of generating processes of man-machine systems]. *Sbornik nauchnykh trudov – Collection of scientific works NSTU*. NGTU–Novosibirsk: Izd-vo NGTU, 2009. №2(56). pp. 71–76. (In Russ.).
  13. Geniatulina E.V., Grif M.G. [Methods for generating a plurality of alternatives in optimization problems of man-machine systems]. *Nauch. vest. NGTU – Scientific Bulletin NSTU*. Novosibirsk: Izd-vo NGTU. 2010. vol. 4(41). pp. 41–50. (In Russ.).
  14. Geniatulina E.V. [Designing processes functioning in man-machine systems. Prospects of development of information technologies]. *Perspektivy razvitiya informatsionnykh tekhnologii. Sbornik materialov II Ezhгодnoi Vserossiiskoi nauchno-prakticheskoi konferentsii s mezhduarodnym uchastiem* [Collection of materials II Annual All-Russian scientific-practical conference with international participation]. Novosibirsk: Izd-vo «Sibprint». 2010. pp. 205–210. (In Russ.).
  15. Geniatulina E.V., Grif M.G. [The generating fragments of the process operation method with the mandatory operations combination]. *Molodoi uchenyi – Young researcher*. Chita. OOO «Izdateľ stvo Molodoi uchenyi. 2009. vol. 10. pp. 83–85. (In Russ.).
  16. Geniatulina E.V. [Data representation in intelligent systems, truncated natural language]. *Materialy devyatoi mezhduarodnoi nauchno-metodicheskoi konferentsii. Informatika: problemy, metodologiya, tekhnologii* [Proceedings of the Ninth International Scientific Conference]. Voronezh. Voronezhskii gosudarstvennyi universitet. 2009. vol. 1. pp. 202–210. (In Russ.).
  17. Geniatulina E.V., Grif M.G. [The method of generating of fragments of the process of functioning in view of the mandatory combinations of operations]. *Sbornik nauchnykh trudov NGTU – Proceedings of NSTU*. 2014. № 3 (77). pp. 53-62.
  18. Grif M.G., Sundui O., Tsoy E.B. Methods of designing and modeling of man-machine systems. Proc. of International Summerworkshop Computer Science 2014. Summerworkshop. Germany. Chemnitz. 2014. pp. 38–40.
  19. Grif M.G., Sundui O., Tsoy E.B. Selecting the method of maintenance and elimination of network failures in the National University of Mongolia. *News of Science and Education*. 2014. vol. 14(14). pp. 116–123.

**Гениатулина Елена Владимировна** — к-т техн. наук, доцент кафедры Автоматизированных систем управления Новосибирского государственного технического университета (НГТУ). Область научных интересов: методы проектирования и оптимизации человеко-машинных систем. Число научных публикаций — 18. genilen@gmail.com; пр. К.Маркса, 20, Новосибирск, 630092; р.т. +7(383) 346-15-59, факс +7(383)346-11-00.

**Geniatulina Elena Vladimirovna** — Ph.D., department of automated control systems, Novosibirsk State Technical University. Research interests: methods of designing and optimization of man-machine systems, systems of sign language machine translation. Number of publications — 18. genilen@gmail.com; prospect Karla Marksa 20, Novosibirsk; office phone. +7(383) 346-15-59, fax +7(383)346-11-00.

**Гриф Михаил Геннадьевич** — д-р техн. наук, профессор, заведующий кафедрой автоматизированных систем управления Новосибирского государственного технического университета (НГТУ). Область научных интересов: методы проектирования и оптимизации человеко-машинных систем, системы компьютерного сурдоперевода. Число научных публикаций — 220. grifmg@mail.ru; пр. К.Маркса, 20, Новосибирск, 630092; р.т. +7(383) 346-15-59, факс +7(383)346-11-00.

**Grif Mikhail Gennadyevich** — Ph.D., Dr. Sci., professor, head of department of automated control systems, Novosibirsk State Technical University. Research interests: methods of designing and optimization of man-machine systems, systems of sign language machine translation. Number of publications — 220. grifmg@mail.ru; prospect Karla Marksa 20, Novosibirsk; office phone +7(383) 346-15-59, fax +7(383)346-11-00.

## РЕФЕРАТ

*Гениатулина Е.В., Гриф М.Г.* **Методы проектирования и моделирования в задачах оптимизации процессов функционирования человеко-машинных систем.**

Целью данной статьи является описание способов расширения области использования ФСТ ЧМС за счет интеграции оптимизационных моделей ПФ ЧМС с методом имитационного моделирования, а также рядом алгоритмов, описанных в статье.

Статья состоит из 6 основных частей: введение, постановка задачи оптимального проектирования ПФ ЧМС, построение множества альтернатив на основе объектно-ориентированной модели данных, оптимизация с учетом обязательных и недопустимых сочетаний операций, интеграция оптимизационных моделей ПФ ЧМС с методом имитационного моделирования и заключение. Во введении описываются основные проблемы в проектировании ПФ ЧМС на основе ФСТ, ставится актуальность задачи. Во второй части ставится задача оптимального проектирования ПФ ЧМС, описывается оптимизационная модель и показатели эффективности. В третьей части описываются и обосновываются подходы к заданию множества альтернатив ПФ ЧМС с использованием модели данных на основе объектно-ориентированного представления операций, которое строится путем синтеза функционально-структурной теории и объектно-ориентированного проектирования. Рассматриваются алгоритм генерации последовательно-параллельных процессов и алгоритм генерации альтернативных фрагментов ПФ ЧМС на базе совпадения целей процессов. В четвертой части представлен алгоритм с учетом обязательных и недопустимых сочетаний операций. Рассматривается конкретный пример. В пятой части описывается способ устранения ограничения ФСТ путем интеграции технологии проектирования ПФ ЧМС основе ФСТ с методом имитационного моделирования, так как ФСТ можно применять только для процессов без последствия и при отсутствии зависимых операций. Предлагаются также новые методы оптимального проектирования ПФ ЧМС.

Данные алгоритмы позволяют сократить время на описание, постановку и решение задач оптимизации в ПФ ЧМС на основе ФСТ в рамках использования гибридной экспертной системы ГЭС ИНТЕЛЛЕКТ-3. Интеграция ФСТ с методом имитационного моделирования позволяет проводить оптимизацию ПФ ЧМС в случае, когда имеют место процессы с последствием. Это также расширяет возможности модели ФСТ ЧМС.

## SUMMARY

*Geniatulina E.V., Grif M.G. Man-machine systems functioning design and modeling methods in optimization processes problems.*

The article considers design of HMS functioning process (FP) based on functional structural theory of HMS (FST HMS) and suggests how to expand application of FST HMS by integrating FP HMS optimization models and simulation method. A number of algorithms are described in the paper.

The article consists of six parts: introduction, FP HMS optimal design problem statement, alternatives set construction based on an object-oriented data model, optimization based on mandatory and invalid operations combinations, FP HMS optimization models and simulation method integration and conclusion. The introduction describes main problems in the FP HMS design based on FST, states the urgency of the subject. The second part includes the definition of the FP HMS optimal design problem, optimization model and efficiency indicators information. The third part describes and justifies the approaches to defining the FP HMS alternatives set using a data model based on pre-operations object-oriented representation, which is constructed by synthesizing FST and object-oriented design. The algorithm for generating series-parallel processes and the algorithm used to generate FP HMS alternative fragments based on process goals matching are considered. The fourth section presents an algorithm based on combination of binding and invalid operations. The concrete example is referring. A particular example is considered. In the fifth section a way to eliminate FST restrictions by integrating the design of FP HMS technology based on FST and simulation method is stated, since FST can be applied only to processes with no aftereffect and dependent operations. New methods for optimal design of FP HMS were offered.

These algorithms are used to reduce the time for description, formulation and optimization problems resolving of FP HMS based on the FST in a hybrid expert system framework HES INTELLEKT-3. Integrating FST and simulation method allows to optimize FP HMS having processes with aftereffects. It also enlarges scope of the FDT HMS model application.

Б.А. Кулик, В.Г. Курбанов, А.Я. Фридман  
**ПАРАЛЛЕЛЬНАЯ ОБРАБОТКА ДАННЫХ И ЗНАНИЙ  
МЕТОДАМИ АЛГЕБРЫ КОРТЕЖЕЙ**

---

*Кулик Б.А., Курбанов В.Г., Фридман А.Я. Параллельная обработка данных и знаний методами алгебры кортежей.*

**Аннотация.** Алгебра кортежей – математическая система для формализации многоместных отношений. С ее помощью можно моделировать в одних и тех же структурах как данные (графы, многоместные отношения), так и знания (семантические сети, модели рассуждений, формулы исчисления высказываний и предикатов, продукционные системы, онтологии и т.д.). В то же время сами эти структуры имеют матрицеподобную форму, а все алгоритмы их обработки легко распараллеливаются.

**Ключевые слова:** параллельные вычисления, алгебра кортежей, структуры данных, базы знаний.

*Kulik B.A., Kurbanov V.G., Fridman A.Ya. Parallel Processing of Data and Knowledge by Means of N-tuple Algebra.*

**Abstract.** *N*-tuple algebra is a mathematical system to formalize *n*-ary relations. This algebra provides for modeling both data (graphs, *n*-ary relations, etc.) and knowledge (semantic networks, reasoning models, formulas of propositional and predicate calculi, production systems, ontologies and so on) by the same structures. These structures look like matrices and can be easily processed by parallel algorithms.

**Keywords:** parallel computation, *n*-tuple algebra, data structure, knowledge base.

---

**1. Введение.** Распараллеливание операций наиболее эффективно осуществляется в тех случаях, когда структуры данных представлены в форме матриц или таблиц. Однако во многих задачах и алгоритмах обработки данных и знаний используются структуры, для которых матричная форма представления не предусматривается, к ним относятся и различные структуры искусственного интеллекта (предикаты, правила, логические формулы и т.д.) Для выполнения параллельных вычислений в этих структурах необходимо предварительно составить граф программы; только потом выбираются его независимые ветви, допускающие параллельную обработку. В частности, в современных системах искусственного интеллекта параллелизм достигается в основном за счет построения графа пар резольвирующихся дизъюнктов Ковальского [1]) и выбора в нем тех связей, которые допустимо обрабатывать одновременно [2]. Этот метод очень сложен в использовании (необходимо учитывать много ограничений и условий) и обладает малой эффективностью.

Поэтому целесообразно найти такую обобщенную форму представления структур данных и знаний, которая позволяла бы относительно легко реализовать вычислительный процесс в виде большого числа независимых операций. По мнению авторов, подобные задачи

можно решать, если выразить многие типы данных и знаний в структурах алгебры кортежей [3, 4], одним из свойств которых является матрицеподобная форма представления и возможность эффективного распараллеливания операций.

**2. Краткие сведения об алгебре кортежей.** *Алгебра кортежей* (АК) математически определена как алгебраическая система, в которой должны быть заданы носитель, совокупность операций, совокупность отношений, а также основные свойства операций и отношений. В некоторых случаях свойства однозначно определяются с помощью доказательств изоморфизма данной системы и какой-либо известной алгебраической системы. В частности, доказано, что АК изоморфна алгебре множеств и относится к классу булевых алгебр [5].

*Носителем* в АК служит произвольное множество *многочестных отношений*, для представления которых предложено четыре типа структур (о них далее), называемых *АК-объектами*. Каждый АК-объект погружен в определенное пространство *атрибутов*, множество значений атрибута называется *доменом*. Имена АК-объектов содержат идентификатор и заключенную в прямые скобки последовательность имен атрибутов, входящих в *схему отношения*, в которой задан этот АК-объект. Например, запись  $R[XYZ]$  говорит, что АК-объект  $R$  определен в пространстве атрибутов  $X$ ,  $Y$  и  $Z$ .

«Произвольное множество *многочестных отношений*», определенное выше как носитель АК, подразумевает, что для этих отношений нет никаких ограничений на состав атрибутов, в которых они заданы – пространства атрибутов разных отношений могут совпадать, а могут существенно отличаться, и эта «произвольность» не нарушает свойств АК как алгебраической системы.

*Основные операции АК* включают *операции алгебры множеств* (пересечение, объединение, дополнение) и *операции с атрибутами* (переименование, перестановка, элиминация, добавление фиктивного атрибута). Комбинации перечисленных операций позволяют ввести *производные операции* с отношениями: соединение, композиция, транзитивное замыкание и т.д. Для сравнения АК-объектов используются два *основных отношения* – *включения* и *равенства*. По своим аналитическим возможностям АК сопоставима с исчислением предикатов, причем АК-объекты моделируют области истинности предикатов и логических формул.

АК-объекты обеспечивают сжатое отображение *многочестных отношений*. При необходимости можно с помощью стандартных алгоритмов преобразовать АК-объекты в обычные *многочестные отношения*.

ния, состоящие из множеств кортежей элементов (в АК эти кортежи называются *элементарными кортежами*).

*Однотипные* АК-объекты – это структуры, заданные в одном пространстве атрибутов. В АК можно выполнять все теоретико-множественные операции не только с однотипными отношениями, но и с отношениями, имеющими разные схемы.

АК-объекты (*C*-кортеж, *C*-система, *D*-кортеж, *D*-система) формируются в виде матриц из подмножеств доменов атрибутов, называемых *компонентами*. В их число входят две *фиктивные компоненты*:

– *полная компонента* (обозначается “\*”) есть множество, равное домену соответствующего (по месту ее расположения в кортеже) атрибута; например, если задан АК-объект  $R[XYZ] = [A * C]$ , то символ “\*” здесь обозначает множество, равное домену атрибута *Y*.

– *пустое множество* –  $\emptyset$ .

Перейдем к описанию основных структур АК – *C*-систем и *D*-систем.

*C*-система записывается в виде матрицы, состоящей из компонент-множеств и ограниченной прямыми скобками. Например,  $R[XYZ] = \begin{bmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{bmatrix}$  есть *C*-система, ее можно преобразовать в обычное отношение (т.е. во множество кортежей) следующим образом:  $R[XYZ] = (A_1 \times A_2 \times A_3) \cup (B_1 \times B_2 \times B_3)$ . При этом  $A_1, B_1 \subseteq X$ ;  $A_2, B_2 \subseteq Y$ ;  $A_3, B_3 \subseteq Z$ .

С помощью *C*-систем удобно представлять дизъюнктивные нормальные формы одноместных предикатов. *C*-система, состоящая из одной строки, называется *C*-кортежем. В логике *C*-кортежу соответствует отдельный конъюнкт.

*D*-системы в АК моделируют конъюнктивные нормальные формы одноместных предикатов. *D*-система обозначается матрицей компонент-множеств, ограниченной перевернутыми прямыми скобками.

Используя *D*-системы, легко вычислить дополнения *C*-систем.

Так, *D*-система  $\overline{T} [XYZ] = \begin{bmatrix} \overline{A} & \emptyset & \overline{C} \\ \overline{D} & \overline{E} & \emptyset \end{bmatrix}$  есть дополнение *C*-системы

$$T[XYZ] = \begin{bmatrix} A & * & C \\ D & E & * \end{bmatrix}.$$

Аналогично *C*-системе, *D*-система из одной строки называется *D*-кортежем, ему в логике соответствует отдельный дизъюнкт.

Правила выполнения операций объединения и пересечения для *C*- и *D*-структур имеют свою специфику, они подробно описаны

в [3,4]. Отметим, что в АК для выполнения всех теоретико-множественных операций и проверок отношений (равенства, включения и т.д.) нет необходимости преобразовывать АК-объекты во множество элементарных кортежей – все операции выполняются с матричными формами.

Для выполнения операций с АК-объектами, имеющими разные схемы отношений, вводятся *операции с атрибутами*, в частности, *добавление фиктивного атрибута (+Atr)* и *элиминация атрибута (-Atr)*. Операция +Atr соответствует *правилу обобщения* в исчислении предикатов, поэтому семантика отношений при ее выполнении не нарушается. Операция производится добавлением имени нового атрибута в схему отношения АК-объекта и нового столбца с фиктивными компонентами – в матричное представление. Например, пусть  $R_k[XZ] = \begin{bmatrix} A_1 & A_3 \\ B_1 & B_3 \end{bmatrix}$ .

Если это отношение моделирует предикат  $R_k(x, z)$ , то добавление фиктивного атрибута  $Y$  в  $R_k[XZ]$  соответствует формуле  $\forall y(R_k(x, z))$ . Для АК-объекта  $R_k[XZ]$  добавление фиктивного атрибута  $Y$  выполняется так:  $+Y(R_k[XZ]) = \begin{bmatrix} A_1 & * & A_3 \\ B_1 & * & B_3 \end{bmatrix}$ .

Операция +Atr часто используется для приведения разнотипных АК-объектов к одной схеме отношения, после чего можно, применяя стандартные алгоритмы АК, выполнять все необходимые операции и проверки. С учетом этого введены *обобщенные операции* ( $\cap_G, \cup_G$ ), в которых операции  $\cap$  и  $\cup$  выполняются после приведения разнотипных АК-объектов к одной схеме отношения. Операции  $\cap_G$  и  $\cup_G$  семантически соответствуют логическим связкам конъюнкции и дизъюнкции. Доказано, что алгебра отношений с обобщенными операциями изоморфна обычной алгебре множеств. Тем самым в теории отношений было снято ограничение, что законы алгебры множеств выполняются только для отношений, определенных на одном и том же декартовом произведении.

Операция элиминации атрибута (например,  $X$ ) при применении к  $C$ -структурам соответствует навешиванию квантора  $\exists x$ , а к  $D$ -структурам – навешиванию квантора  $\forall x$  на соответствующие логические формулы. В АК эта операция выполняется как удаление атрибута из схемы отношения и соответствующего ему столбца из матричного представления АК-объекта. Например, для  $D$ -системы  $R[XYZ] = \begin{bmatrix} A & \emptyset & B \\ C & D & \emptyset \end{bmatrix}$  при вычислении  $-Y(R[XYZ])$  получим

$Q[XZ] = \begin{bmatrix} A & B \\ C & \emptyset \end{bmatrix}$ . При этом, если  $D$ -системе  $R[XYZ]$  соответствует логическая формула  $F(x, y, z)$ , то  $D$ -системе  $Q[XZ]$  – логическая формула  $\forall yF(x, y, z)$ .

Расширенный вариант АК предусматривает применение некоторых совокупностей атрибутов в качестве отдельных (составных) атрибутов. Тогда компонентами АК-объектов становятся не обычные множества, а многоместные отношения, выраженные АК-объектами. Другим расширением является использование в структурах АК атрибутов, домены которых представлены структурами, отличающимися по свойствам от обычных множеств (например, нечеткими множествами). Для атрибутов, представляющих систему интервалов на числовой оси, в АК разработан метод квантования интервалов, после применения которого такие атрибуты можно обрабатывать всеми методами АК [5].

Значительная часть алгоритмов выполнения операций АК (пересечение, объединение, дополнение, проверка включения АК-объектов и т.д.) имеет полиномиальную сложность. Экспоненциальной сложностью характеризуется, в частности, алгоритмы преобразования  $C$ -систем в  $D$ -системы и обратно. Во многих процедурах обработки информации подобные преобразования не нужны. Если все же возникает необходимость применения таких алгоритмов, их трудоемкость удастся снизить до полиномиальной для некоторых частных случаев структур АК-объектов [6].

В настоящее время исследована и обоснована возможность применения АК в следующих структурах данных и знаний [5]:

- 1) графы и сети;
- 2) модели исчисления высказываний и предикатов;
- 3) системы искусственного интеллекта (семантические сети, экспертные системы, фреймы, онтологии);
- 4) модели дедуктивных и абдуктивных рассуждений;
- 5) логико-вероятностные методы, включая вероятностную логику;
- 6) дискретные автоматы.

Продолжаются исследования, связанные с возможностью использования АК в других областях, в частности, в структурах с неопределенностями и в динамических системах.

**3. Методы и средства распараллеливания операций в структурах АК.** В основе операций АК лежат теоретико-множественные операции с компонентами (подмножествами доменов атрибутов). Например, для вычисления пересечения двух однотипных  $C$ -кортежей требуется найти пересечение пар компонент в соответствующих атри-

бутах, причем в результате получится пустое множество, если пересечение хотя бы одной пары компонент пусто. Пересечение  $C$ -кортежа  $R$  с однотипной  $C$ -системой  $Q$  вычисляется как результат пересечения  $R$  с каждым  $C$ -кортежем из  $Q$ . Для двух однотипных  $C$ -кортежей  $R$  и  $R_1$  имеет место включение  $R \subseteq R_1$ , если все компоненты  $R$  включены в соответствующие компоненты  $R_1$ . Некоторые операции в АК не требуют операций с компонентами. Например, объединение двух однотипных  $C$ -систем получается просто присписыванием всех  $C$ -кортежей одной из этих  $C$ -систем другой  $C$ -системе.

Приведенные примеры достаточно полно характеризуют специфику выполнения операций в АК. Для обоснования наиболее подходящей архитектуры вычислительного комплекса, предназначенного для распараллеливания операций в АК, целесообразно за основу взять часто встречающуюся операцию пересечения двух  $C$ -систем. Пусть заданы  $C$ -системы:

$$R_1[XYZ] = \left[ \begin{array}{ccc} \{a, b, d\} & \{f, h\} & \{b\} \\ \{b, c\} & * & \{a, c\} \end{array} \right] \text{ и}$$

$$R_2[XYZ] = \left[ \begin{array}{ccc} \{a, d\} & * & \{b, c\} \\ \{b, d\} & \{f, h\} & \{a, c\} \\ \{b, c\} & \{g\} & \{b\} \end{array} \right].$$

Их пересечение вычисляется как пересечение каждого  $C$ -кортежа первой  $C$ -системы с каждым  $C$ -кортежем второй:

$$[\{a, b, d\} \{f, h\} \{b\}] \cap [\{a, d\} * \{b, c\}] = [\{a, d\} \{f, h\} \{b\}];$$

$$[\{a, b, d\} \{f, h\} \{b\}] \cap [\{b, d\} \{f, h\} \{a, c\}] = \emptyset;$$

$$[\{a, b, d\} \{f, h\} \{b\}] \cap [\{b, c\} \{g\} \{b\}] = \emptyset;$$

$$[\{b, c\} * \{a, c\}] \cap [\{a, d\} * \{b, c\}] = \emptyset;$$

$$[\{b, c\} * \{a, c\}] \cap [\{b, d\} \{f, h\} \{a, c\}] = [\{b\} \{f, h\} \{a, c\}];$$

$$[\{b, c\} * \{a, c\}] \cap [\{b, c\} \{g\} \{b\}] = \emptyset.$$

Затем из непустых  $C$ -кортежей формируем  $C$ -систему:

$$R_1 \cap R_2 = \left[ \begin{array}{ccc} \{a, d\} & \{f, h\} & \{b\} \\ \{b\} & \{f, h\} & \{a, c\} \end{array} \right].$$

На этом типичном примере нетрудно сформулировать основные закономерности распараллеливания операций в АК. Во-первых, можно одновременно выполнять операции с компонентами при вычислении пересечения пар кортежей. Во-вторых, параллельное выполнение допускают операции пересечения одного из кортежей первой  $C$ -системы

со всеми  $C$ -кортежами второй. И, в-третьих, при определенной организации вычислительного процесса можно выполнить параллельно все вычисления с компонентами, необходимые для получения окончательного результата.

Выигрыш во времени расчета за счет распараллеливания вычислений в АК будет достаточно велик. Например, при вычислении пересечения двух  $C$ -систем в пространстве из  $K$  атрибутов по третьему из перечисленных вариантов распараллеливания, если первая  $C$ -система содержит  $M$  строк, а вторая –  $N$ , то время расчета уменьшится в  $K \times M \times N$  раз. Той же сложностью и такой же возможностью распараллеливания обладает и операция проверки включения одного АК-объекта в другой, если первый выражен как  $C$ -система, а второй – как  $D$ -система. Здесь в качестве элементарной операции с компонентами используется операция проверки включения множеств.

Рассмотренную операцию пересечения  $C$ -систем можно считать типовой операцией, поскольку она соответствует следующим операциям в традиционных системах: 1) конъюнкция предикатов или формул, выраженных  $C$ -системами; 2) дизъюнкция предикатов или формул, представленных как  $D$ -системы; 3) операции соединения и композиции отношений, которые, в частности, используются при вычислении транзитивного замыкания отношения; 4) операция поиска ответа на вопрос в тех случаях, когда структура данных и структура вопроса выражены  $C$ -системами. Ясно, что для выполнения операций с компонентами АК-объектов достаточно использовать в качестве процессорного элемента (ПЭ) вычислительное устройство, способное обрабатывать булевы векторы. Для каждого атрибута целесообразно предусмотреть множество однородных ПЭ с общей памятью, в которой хранятся заданные и получаемые в процессе вычислений компоненты данного атрибута. Такая структура соответствует SMP архитектуре (symmetric multiprocessing) [7].

Работу со множеством атрибутов предпочтительно организовать с помощью модулей SMP в качестве элементов вычислительных систем с распределенной памятью MPP (massive parallel processing) [8]. Поскольку атрибуты могут существенно различаться по числу значений, в вычислительный комплекс следует включить разные SMP-кластеры, отличающиеся разрядностью процессорных элементов и соответствующих им модулей памяти. Таким образом, в общем случае для параллельной реализации операций со структурами АК подходит гибридная архитектура, сочетающая в себе особенности систем с MPP и SMP архитектурой. Схема такого комплекса показана на рисунке 1.

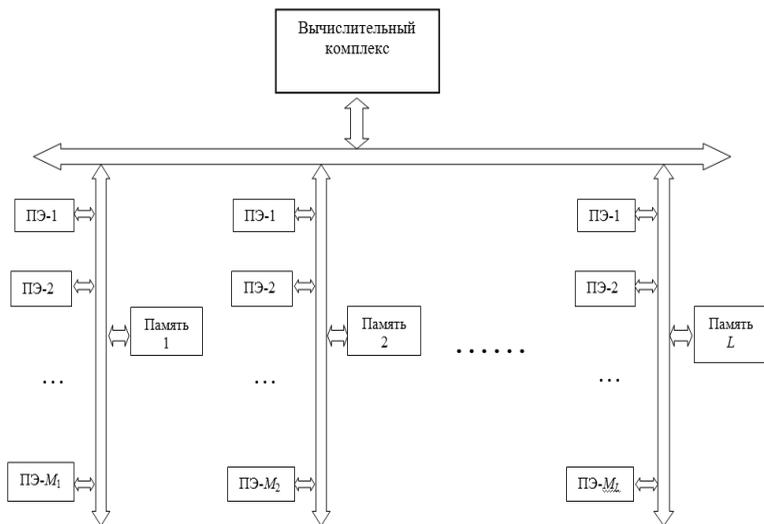


Рис. 1. Схема гибридной MPP-SMP архитектуры

Рассмотрим, как выполняется операция пересечения двух  $S$ -систем  $A$  и  $B$ , представленных в виде матриц  $\|a_{iv}\|$  и  $\|b_{jv}\|$  с размерностью  $M \times K$  и  $N \times K$  соответственно. Пусть из множества SMP-кластеров вычислительного устройства выбраны те, которые соответствуют атрибутам пространства, где заданы  $A$  и  $B$ . Совокупность всех компонент каждого  $v$ -го атрибута размещается в памяти соответствующего SMP-кластера (на рис. 1, например, Память 1, Память 2, ..., Память  $K$ ). Для вычисления пересечения векторов  $\|a_{1v}\|$  и  $\|b_{1v}\|$  в соответствующие процессорные элементы ПЭ-1 в каждом из SMP-кластеров загружаются компоненты этих векторов и выполняется их пересечение. Если все результаты непустые, то они записываются в соответствующие ячейки Памяти  $v$ . Если требуется выполнить одновременно пересечение вектора  $\|a_{1v}\|$  со всеми векторами матрицы  $\|b_{jv}\|$ , то в соответствующих SMP-кластерах компоненты векторов одновременно передаются в процессорные элементы ПЭ-1, ПЭ-2, ..., ПЭ- $N$  и выполняются требуемые операции. Для параллельного вычисления пересечений всех векторов матриц  $\|a_{iv}\|$  и  $\|b_{jv}\|$  необходимо использовать дополнительные ресурсы вычислительного комплекса. Воз-

можно два варианта: 1) использовать процессорные элементы ПЭ- $N+1$ , ПЭ- $N+2$  и т.д. в используемых SMP-кластерах; 2) использовать другие SMP-кластеры.

**4. Заключение.** Алгебра кортежей позволяет представлять и обрабатывать данные и знания разнообразных форматов. В настоящей работе предложена гибридная архитектура вычислительного комплекса, обеспечивающая эффективное распараллеливание операций на структурах АК.

### Литература

1. Kowalski R. A. Proof Procedure Using Connection Graphs // J. of the ACM. 1975. vol. 22(4). pp. 572–599.
2. Loganathanaraj R., Mueller R. A. Parallel Theorem Proving with Connection Graph // 8th Int. Conf. on Autom. Deduc. LNCS 230. 1986. pp. 337–352.
3. Кулик Б.А. Вероятностная логика на основе алгебры кортежей // Известия РАН. Теория и системы управления. 2007. № 1. С. 118–127.
4. Kulik B., Fridman A., Zuenko A. Logical Inference and Defeasible Reasoning in N-tuple Algebra // In: “Diagnostic Test Approaches to Machine Learning and Commonsense Reasoning Systems”. IGI Global. 2013. pp. 102–128.
5. Кулик Б.А., Зуенко А.А., Фридман А.Я. Алгебраический подход к интеллектуальной обработке данных и знаний // СПб.: Изд-во Политехн. ун-та. 2010. 235 с.
6. Кулик Б.А. Новые классы КНФ с полиномиально распознаваемым свойством выполнимости // Автоматика и телемеханика. 1995. № 2. С. 111–124.
7. Severance C., Dowd K. High Performance Computing (RISK Architectures, Optimization & Benchmarks). 2nd Edition // O’Reily Media. 1998. 466 p.
8. Batcher K. E. Design of a Massively Parallel Processor // IEEE Transactions on Computers. 1980. vol. C–29 (9). pp. 836–840.

### References

1. Kowalski R. A. Proof Procedure Using Connection Graphs. J. of the ACM. 1975. vol. 22(4). pp. 572–599.
2. Loganathanaraj R., Mueller R. A. Parallel Theorem Proving with Connection Graph. 8th Int. Conf. on Autom. Deduc. LNCS 230. 1986. pp. 337–352.
3. Kulik B. A. [Probabilistic logic based on the N-tuple Algebra]. *Izvestija RAN. Teorija i sistemy upravlenija – Journal of Computer and Systems Sciences International*. 2007. vol. 1. pp. 118–127. (In Russ.).
4. Kulik B., Fridman A., Zuenko A. Logical Inference and Defeasible Reasoning in N-tuple Algebra. In: “Diagnostic Test Approaches to Machine Learning and Commonsense Reasoning Systems”. IGI Global. 2013. pp. 102–128.
5. Kulik B. A., Zuenko A. A. Fridman A. Ya. *Algebraicheskiy podhod k intellektual'noj obrabotke dannyh i znaniy* [Algebraic Approach to Intelligent Processing of Data and Knowledge]. Saint-Petersburg. Polytechnic Univ. Publ. 2010. 235 p. (In Russ.).
6. Kulik B. A. [New classes of CNF with polynomially recognizable property of satisfiability]. *Avtomatika i telemekhanika – Automation and Remote Control*. 1995. vol. 2. pp. 111–124. (In Russ.).
7. Severance C., Dowd K. High Performance Computing (RISK Architectures, Optimization & Benchmarks). 2nd Edition. O’Reily Media. 1998. 466 p.
8. Batcher K. E. Design of a Massively Parallel Processor. IEEE Transactions on Computers. 1980. vol. C–29 (9). pp. 836–840.

**Поддержка исследований.** Работа выполнена при финансовой поддержке РФФИ (проекты №№ 12-07-00302-а, 12-07-00550-а, 12-07-00689-а, 13-07-00318-а, 14-07-00256-а, 14-07-00257-а, 14-07-00205-а) и Президиума РАН (проект 4.3 Программы № 16).

**Acknowledgements.** This research is supported by RFBR (grants 12-07-00302-a, 12-07-00550-a, 12-07-00689-a, 13-07-00318-a, 14-07-00256-a, 14-07-00257-a, 14-07-00205-a), presidium RAS (grant 4.3 of the program №16).

**Кулик Борис Александрович** — д-р физ.-мат. наук, доцент, ведущий научный сотрудник лаборатории интеллектуальных электромеханических систем, Институт проблем машиноведения РАН. Область научных интересов: искусственный интеллект, методы логического анализа систем, логико-вероятностный анализ. Число научных публикаций — 90. ba-kulik@yandex.ru; 199178 Санкт-Петербург, В.О., Большой проспект, д.61; р.т. +7(812) 321 90 07.

**Kulik Boris Alexandrovich** — Ph.D., Dr. Sci., associate professor, leading researcher of Laboratory of Methods and Systems of Automation, Institute of Problems in Mechanical Engineering (IPME RAS). Scientific interests: artificial intelligence techniques of logical analysis systems, logic and probabilistic analysis. The number of publications — 90. ba-kulik@yandex.ru; 61, V.O., Bol'shoi av., 199178, Saint-Petersburg, Russian Federation; office phone +7(812) 321 90 07.

**Курбанов Вугар Гариб оглы** — к-т физ.-мат. наук, доцент, старший научный сотрудник, Институт проблем машиноведения РАН. Область научных интересов: общие вопросы математики, математическая кибернетика, математическое моделирование, оптимальные системы. Число научных публикаций — 60. vugar\_borchali@yahoo.com; 196135, Санкт - Петербург, Большой пр., 61, В.О.; р.т. +7(812)3219007

**Kurbanov Vugar Garib ogli** — Ph.D., associate professor, senior researcher Institute of Problems in Mechanical Engineering (IPME RAS). Scientific interests: general questions of mathematical cybernetics, mathematical modeling, optimum systems. The number of publications — 60. vugar\_borchali@yahoo.com; 61, V.O., Bol'shoi av., 199178, Saint-Petersburg, Russian Federation; office phone +7(812)3219007.

**Фридман Александр Яковлевич** — д-р техн. наук, профессор, ведущий научный сотрудник лаборатории информационных технологий управления промышленно-природными системами, Институт информатики и математического моделирования Кольского научного центра РАН. Область научных интересов: моделирование комплексных технологий и их воздействия на окружающую среду, прикладные интеллектуализированные системы. Число научных публикаций — 230. fridman@iimm.kolasc.net.ru; 184209, Мурманская обл., г.Апатиты, ул.Ферсмана, 24а; р.т. +7(81555) 79782.

**Fridman Alexander Jakovlevich** — Ph.D., Dr. Sci., professor, leading researcher of Laboratory of Information Technologies for Nature-Industrial Systems Control, Institute for Informatics and Mathematical Modelling of Technological Processes of RAS. Scientific interests: modeling of complex technologies and their impact on the environment, application intellectualized system. The number of publications — 230. fridman@iimm.kolasc.net.ru; 24A, Fersman st., 184209, Apatity, Murmansk Region, Russian Federation; office phone +7(81555) 79782.

## РЕФЕРАТ

### *Кулик Б.А., Курбанов В.Г., Фридман А.Я.* **Параллельная обработка данных и знаний методами алгебры кортежей.**

При формализации и компьютерной реализации данных и знаний часто используются структуры, которые не представлены в матричной форме и поэтому для них трудно осуществить параллельную обработку данных. К таким трудным для ускорения вычислений структурам относятся различные формы искусственного интеллекта (предикаты, правила, логические формулы, семантические сети и т.д.) Для выполнения параллельных вычислений в этих структурах необходимо предварительно составить граф программы; только потом выбираются его независимые ветви, допускающие параллельную обработку. Этот метод очень сложен в использовании (необходимо учитывать много ограничений и условий) и обладает малой эффективностью.

Поэтому целесообразно найти такую обобщенную форму представления структур данных и знаний, которая позволяла бы относительно легко реализовать вычислительный процесс в виде большого числа независимых операций. По мнению авторов, подобные задачи можно решать, если выразить многие типы данных и знаний в структурах алгебры кортежей одним из свойств которых является матрицеподобная форма представления и возможность эффективного распараллеливания операций. С помощью алгебры кортежей отображаются практически все структуры искусственного интеллекта и в то же время структуры алгебры кортежей имеют матричную форму и тем самым позволяют легко распараллеливать алгоритмы обработки данных и знаний. Для этого предложена гибридная архитектура вычислительного комплекса, обеспечивающая эффективное распараллеливание операций на структурах алгебры кортежей.

## SUMMARY

### *Kulik B.A., Kurbanov V.G., Fridman A.Ya.* **Parallel Processing of Data and Knowledge by Means of N-tuple Algebra.**

The formalization and implementation of computer data and knowledge are often used structures that are not represented in matrix form and therefore difficult for them to implement parallel processing. Such difficult to speed up the calculations structures include various forms of artificial intelligence (predicates, rules, logical formulas, semantic network, etc.) to perform parallel computations in these structures must first create a graph of the program; Only then chooses its independent branches, allowing parallel processing. This technique is very complicated to use and low efficient due to the necessity of considering many conditions and restrictions.

Efficient paralleling of algorithms is achievable in cases when initial data are represented in a matrix form. Conversely, most knowledge processing and analyzing systems use structures not similar to matrices. This is why it looks reasonable to propose such a generalized structure form for data and knowledge representation that would allow for a comparatively easy transformation of a computational process into a large number of independent operations. In our opinion, N-tuple algebra (NTA) can solve this kind of problems by expressing many formats of data and knowledge as NTA structures. With NTA displayed virtually all structures of artificial intelligence, and at the same time, the structure of the N-tuple algebra have a matrix form, and thus make it easy to parallelize algorithms for processing data and knowledge. For this proposed hybrid architecture computing system that provides effective parallelization of operations on structures N-tuple algebra.

А.А. МУСАЕВ

## ЧИСЛЕННОЕ ИССЛЕДОВАНИЕ УПРАВЛЯЮЩИХ СТРАТЕГИЙ, ОСНОВАННЫХ НА АНАЛИЗЕ ТРЕНДОВ

---

*Мусаев А.А.* Численное исследование управляющих стратегий, основанных на анализе трендов.

**Аннотация.** На основе серии вычислительных экспериментов рассмотрен вопрос об эффективности управляющих стратегий технического анализа, основанного на анализе локальных трендов в квазихаотических процессах. В качестве полигона данных используются длительные интервалы наблюдений за котировками валютных инструментов на электронном рынке Forex. Установлено, что наличие локального тренда не является достаточным условием для построения эффективной управляющей стратегии.

**Ключевые слова:** хаотические процессы, управляющая стратегия, тренд.

*Musaev A.A.* Numerical Research of the Operating Strategy Based on the Trends Analysis.

**Abstract.** On the basis of a series of computing experiments the problem of efficiency of technical analysis operating strategy based on the analysis of local trends in quasischaotic processes is considered. As the range of data long intervals of observations over quotations of currency tools in the electronic market Forex are used. It is established that existence of a local trend isn't a sufficient condition for creation of efficiency operating strategy.

**Keywords:** the chaotic processes, operating strategy, trend.

---

**1. Введение.** В настоящее время существует большое число управляющих стратегий, явно или косвенно основанных на предположении о наличии памяти у квазихаотического процесса ценообразования на электронных рынках капитала [1-11]. Получить подтверждение данной гипотезе на основе аналитических технологий крайне сложно в связи с бесконечным разнообразием реализации хаотической динамики. Тем не менее, имеется возможность численного анализа данной проблемы на основе полигонов данных большой размерности. В частности, такой подход использовался в [12], в рамках вычислительных экспериментов, посвященных инерционности наблюдаемого процесса изменения котировок валютных инструментов. Однако количественные характеристики наблюдаемых тенденций в этой работе никак не учитывались. При этом инерционность процесса рассматривается как обоснование серии управляющих стратегий, основанных на обнаружении выявленных тенденций.

Таким образом, представляют интерес исследования, в которых оцениваются варианты анализа инерционности на основе количественных характеристик наблюдений котировок. В качестве количественных характеристик тенденций будем рассматривать

параметры степенных полиномов, аппроксимирующих динамику изучаемых хаотических процессов.

Приведенные исследования следует рассматривать как продолжение и развитие анализа инерционности хаотических процессов [12].

**2. Математическая постановка задачи.** Имеем хаотический процесс  $Y(t)$ . Формируется скользящее окно наблюдений:

$$W(Y, l) = [Y(t-l+1), Y(t-l+2), \dots, Y(t)].$$

Здесь  $l$  - размер скользящего окна. Далее, осуществляется полиномиальная аппроксимация наблюдений по *методу наименьших квадратов* (МНК). В результате получаем оценки коэффициентов аппроксимирующего полинома  $a = (a_0, a_1, \dots, a_q)$ , где величина  $q \leq 3$ .

Более высокая степень аппроксимации, как правило, приводит к вырожденности матрицы нормальных уравнений.

Гипотеза об отсутствии инерционности обнаруженного тренда означает, что процесс, после обнаружения тренда, достигает пороговых значений  $Y_0(t) + dL$  или  $Y_0(t) - dL$  с одинаковой вероятностью  $H_0 : p = p_0 = 0.5$ . Здесь  $Y_0(t)$  - значение наблюдаемого процесса на момент обнаружения тренда.

Альтернативная гипотеза, указывающая на инерционность тренда, а, следовательно, на принципиальную возможность построения выигрышной управляющей стратегии, основанной на трендах, будет иметь вид  $H_1 : p \neq p_0 = 0.5$ .

Помимо основного хаотического процесса  $Y(t)$  в дальнейшем будем использовать его сглаженную версию  $Y_s(t)$ :  $Y_s(t) = \alpha Y(t) + \beta Y_s(t-1)$ , где  $\beta = 1 - \alpha$ ,  $\alpha = 0.01 - 0.02$ . Предложенный коэффициент сглаживания позволяет сформировать процесс  $Y_s(t)$ , имитирующий системную составляющую хаотической динамики. Остаток  $v(t) = Y(t) - Y_s(t)$  носит характер централизованного случайного процесса с распределением, близким к гауссовскому. Значение дисперсии остаточного процесса  $D(v(t))$  позволяет оценить нижнюю границу параметра  $dL$ , определяющего уровень подтверждения или отрицания гипотезы  $H_0 : p = 0.5$  об отсутствии инерционности тренда  $H_0$ . Степень близости частоты  $m/n$  к уровню  $p_0 = 0.5$  определяется

на основе традиционной методики проверки статистических гипотез, в частности, на основе применения статистики:

$$U = \left( \frac{m}{n} - p_0 \right) \sqrt{n(p_0 q_0)} \frac{1}{2},$$

где  $q_0 = 1 - p_0 = 0.5$ . Уровень доверия, используемый для определения критического значения  $U^*$ , выбран равным  $\alpha = 0.99$ . Критическая область для симметричной конкурирующей гипотезы  $H_1: p \neq p_0$  определяется на основе выбранного уровня  $\alpha$ . Для двухсторонней критической области  $u^*$  определяется по таблице значений функции Лапласа из условия  $\Phi(u^*) = (1 - \alpha)/2 = 0.005$ . Из таблиц распределения функции Лапласа определяем  $u^* = 2,85$ .

**3. Общее описание серии вычислительных экспериментов.** В качестве полигона хаотических данных рассматривается временной ряд, образованный котировками валютного инструмента EURUSD на различных непересекающихся 100-дневных интервалах наблюдений. Формируется скользящее окно наблюдений  $W(Y, l)$  размером  $l$ , на котором вычисляются аппроксимирующие полиномы  $P(q, l)$ . На основе сопоставления оценок коэффициентов полиномов  $a$  с критическими значениями  $a^*$  вырабатывается решение о наличии тренда. Для реализации статистической проверки инерционности определяется число исходов, соответствующих достижению процессом заранее установленного уровня  $dL$  (в терминологии трейдеров уровни  $TP$  (take profit)). В силу симметричности задачи отрицательный результат состоит в развороте тренда и достижении уровня  $-dL$  (уровня  $SL$  (stop loss)).

Заметим, что в трейдерской практике обычно  $TP \neq SL$ . Однако это мало влияет на результативность, поскольку увеличение уровня  $SL$  (и уменьшение вероятности его достижения) пропорционально увеличивает размер проигрыша.

В случае, если отношение числа достижения уровня  $dL$  к общему числу открытий позиций  $m/n$  (частота события) близка к 0.5, то это означает подтверждение гипотезы  $H_0$  об отсутствии инерционности обнаруженных тенденций. Это, по сути, будет означать невозможность построения выигрышных стратегий на основе анализа трендов.

Параметрами вычислительного эксперимента являются величины окна наблюдений  $l$ , степень сглаживающего полинома  $q$ , пороговые значения трендов  $a^*$ , уровень подтверждения тренда (или уровень закрытия позиции)  $dL$ .

**4. Вычислительный эксперимент 1.** Используется простейшая схема линейной аппроксимации  $Y_M(t) = a_0 + a_1 t$  на скользящем окне наблюдения  $W(Y, l)$ . Фиксация тренда осуществляется при выполнении условия превышения коэффициентом  $a_1$  заранее установленное значение  $a^*$ :  $a_1 \geq a^*$ . Подтверждение или отрицание факта наличия тренда осуществляется при выполнении условия  $Y(t) = Y_0 \pm dL$  где  $Y_0$  - значение процесса в момент фиксации тренда,  $dL = 30, 50, 100$ , - уровни подтверждения тренда. Размер окна наблюдения  $W(Y, l)$  варьируется в диапазоне (0,1-0,5) суток.

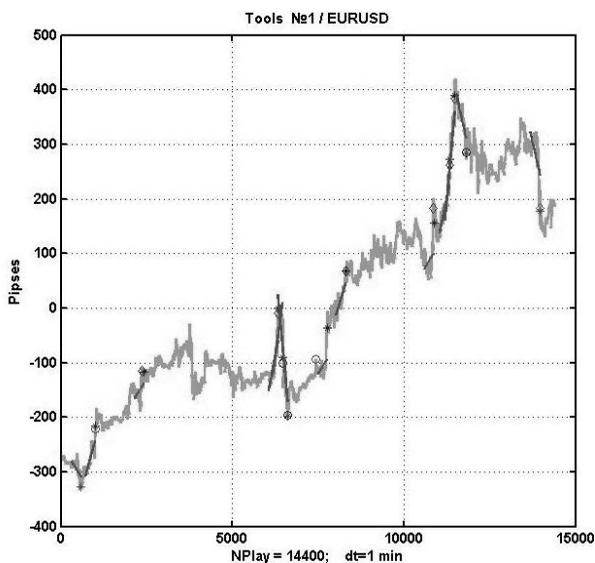


Рис. 1. Пример реализации простейшей игровой стратегии на основе динамики линейных трендов на интервале 10 суток

В качестве примера на рисунке 1 приведен график  $Y(t)$  и выявленных трендов на момент их фиксации с критическим уровнем принятия решения  $a^* = 0.1$  на интервале 10 суток.

При этом рассматривается простейший вариант игровой стратегии, когда позиция открывается вверх или вниз при условии  $a_1 \geq a^*$ , и закрывается, соответственно, при достижении уровней  $Y(t) = Y_0 \pm dL$ , где  $dL = 75$ .

Результаты оценки вероятности (частоты) достижения уровня подтверждения тренда для его различных значений  $dL = 25, 50, 75, 100$ , для окна наблюдения  $l = 0.1, 0.25, 0.5$  суток и для пороговых значений фиксации тренда  $a^* = 0.05, 0.075, 0.1$  на интервале наблюдения в 100 игровых дней приведены в таблице 1.

Таблица 1. Частота подтверждения тренда для различных значений параметров управляющей стратегии

Частота подтверждения тренда					
$a^*$	$dL, n. \setminus l$ , дн.	0.025	0.05	0.075	0.1
0.025	25	0.50	0.48	0.49	0.49
0.025	50	0.51	0.50	0.51	0.50
0.025	75	0.50	0.50	0.50	0.51
0.025	100	0.50	0.51	0.51	0.51
0.05	25	0.50	0.48	0.48	0.50
0.05	50	0.50	0.50	0.51	0.50
0.05	75	0.50	0.50	0.51	0.50
0.05	100	0.50	0.51	0.51	0.51
0.075	25	0.50	0.49	0.49	0.49
0.075	50	0.50	0.50	0.50	0.50
0.075	75	0.50	0.50	0.51	0.51
0.075	100	0.51	0.51	0.51	0.51
0.1	25	0.50	0.48	0.48	0.49
0.1	50	0.50	0.51	0.51	0.50
0.1	75	0.50	0.50	0.50	0.50
0.1	100	0.50	0.50	0.50	0.51

Представленные данные с полной очевидностью показывают полную отсутствие инерционности процесса  $Y(t)$  в широком диапазоне изменения значений интенсивности, уровней фиксации и уровней подтверждения тренда.

Недостатком данного эксперимента является фиксированный уровень скользящего окна наблюдения  $l$ .

Большое окно приводит к задержке времени обнаружения тренда, что приводит к запаздывающему решению и, как следствие, к

открытию позиции на завершающем этапе развития тенденции. Соответствующий этому случаю пример приведен на рисунке 2.

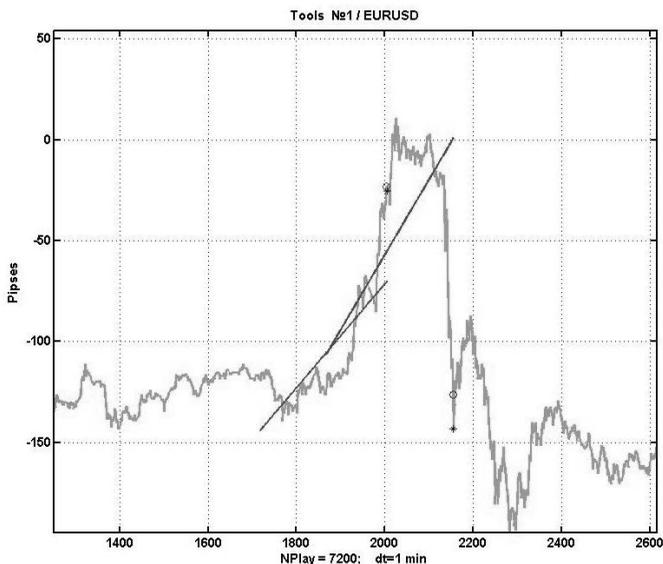


Рис 2. Примеры неправильного решения о наличии тренда, обусловленного запаздыванием принятия решения

Небольшое окно приводит к повышенной чувствительности процедуры обнаружению тренда к случайной составляющей, что приводит к статистическим ошибкам 2-го рода («ложные тревоги»), то есть к обнаружению несуществующего тренда.

В связи с этим перейдем ко второму эксперименту, когда используются два скользящих окна наблюдения разных размеров.

**5. Вычислительный эксперимент 2.** В отличие от первого эксперимента, для формирования управляющего решения рассматриваются два тренда. При этом используются линейные аппроксимации  $q = 1$  для двух скользящих окон наблюдения размером  $l_1$  и  $l_2$ , причем  $l_1 > l_2$ . Очевидно, что процедура обнаружения первого тренда более инерционна, второго – более чувствительна как к системным изменениям процесса, так и к «ложным тревогам».

Пусть  $l_1 = 300$ ,  $l_2 = 90$  минутных отсчетов, критические значения коэффициента линейной регрессии  $\alpha_1^* = 0.05$ ,  $\alpha_2^* = 0.1$ , уровень подтверждения тренда  $dL = 75$ . Решение о наличии тренда

принимается, если коэффициенты линейной регрессии  $\alpha_1, \alpha_2$  превышают свои критические значения (для отрицательного тренда – меньше отрицательных значений критических величин).

Пример реализации такой схемы приведен на рисунке 3. Более длинные тренды соответствуют большим окнам наблюдения.

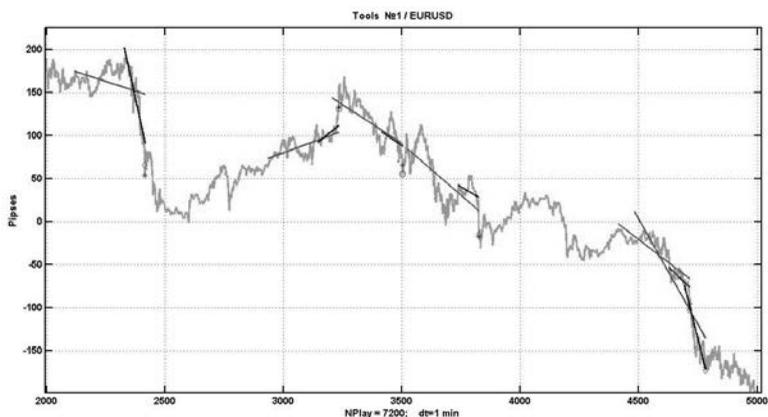


Рис. 3. Пример реализации схемы принятия решения с двумя трендами

Рассмотрим результат применения данной программы для 4-х дневных интервалов с разными уровнями подтверждения тренда. Соответствующие данные приведены в таблице 2. Нетрудно видеть, что организованная модификация не дала положительного эффекта.

Таблица 2. Частота подтверждения тренда для различных временных интервалов реализации стратегии

Частота подтверждения тренда				
$dL, n, \Delta T$	1-100дн.	101-200дн.	201-300дн.	301-400дн.
25	0.48	0.49	0.48	0.47
50	0.53	0.46	0.48	0.48
75	0.53	0.50	0.49	0.45
100	0.55	0.50	0.52	0.47

Очевидным недостатком данной стратегии (как и предыдущей) является то, что программа обнаруживает тренд, как правило, в момент подтверждения (или отрицания) предыдущего тренда. В то же время в течение времени подтверждения нового обнаружения тренда не происходит, для этого нужна программа с одновременным анализом

нескольких тенденций. Однако сделать вывод о наличии инерционности приведенный подход вполне позволяет.

Приведенные данные достаточно наглядно иллюстрируют крайне незначительные флуктуации частоты подтверждения тренда относительно величины 0.5. Указанный вывод легко подтверждается на основе традиционной технологии проверки статистической гипотезы  $H_0 : p = 0.5$  об отсутствии инерционности тренда с использованием статистики  $U$  и уровнем доверия  $\alpha = 0.99$ .

**6. Вычислительный эксперимент 3.** Дополнительно используются нелинейные аппроксимации порядка  $q = 2, 3$ .

Заметим, что при использовании подгонки по МНК полиномами 2-го порядка, в качестве параметров, пригодных для построения решающих статистик можно предложить значение разницы, между экстремумом аппроксимирующей параболы  $\hat{Y}^{(2)}(t^*)$  и отвечающим ей значением линейной аппроксимации  $\hat{Y}^{(1)}(t^*)$ . Примеры соответствующих графиков для восходящего и нисходящего трендов представлены на рисунках 4, 5.

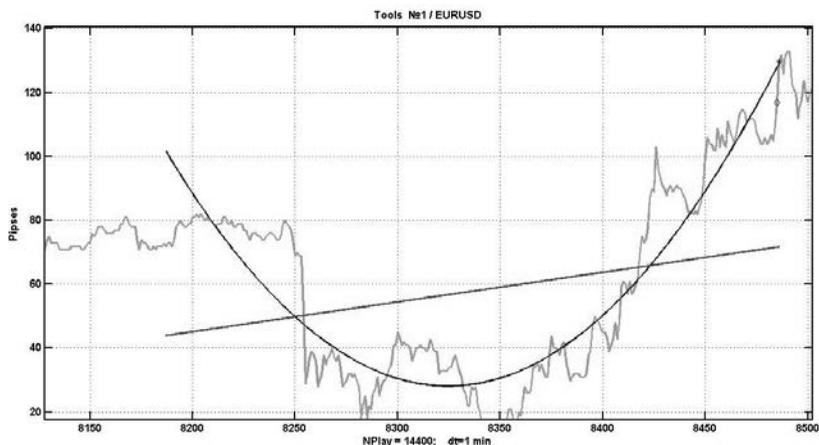


Рис. 4. Фиксация восходящего тренда с аппроксимациями 1-го и 2-го порядка

Еще более эффективным средством отображения перегиба тренда является статистика  $\int_{t \in W} |\hat{Y}^{(2)}(t) - \hat{Y}^{(1)}(t)|$ , где  $W$  - окно наблюдения.

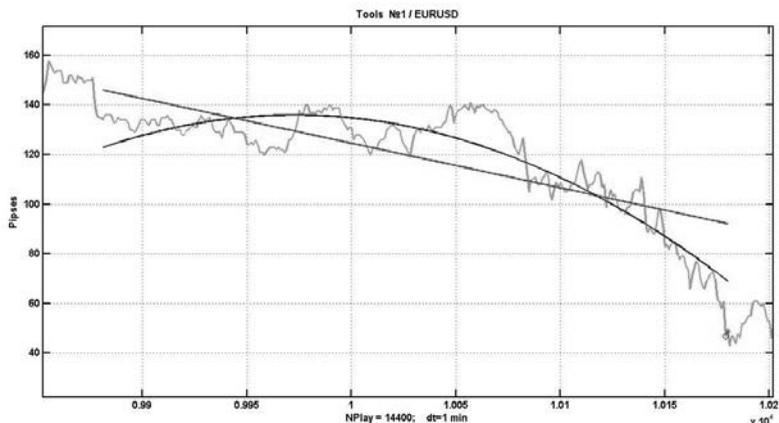


Рис. 5. Фиксация нисходящего тренда с аппроксимациями 1-го и 2-го порядка

Очевидно, что на участках с монотонным трендом эти статистики будут иметь минимальные значения, на участках перегибов тренда, крайне интересных для построения игровых стратегий, они будут максимальны.

К сожалению, построить эффективные решающие правила пока что не удастся. Это связано с тем, что перегибы тренда часто оказываются «ложными тревогами» (т.е. статистическими ошибками 2-го рода), а обнаружение сильного тренда осуществляется на его завершающем участке. Иными словами, данный индикатор также находится в прокрустовом ложе «Опоздание – ложная тревога».

В качестве примера рассмотрен 100-дневный интервал наблюдения котировок валютной пары EURUSD, для которого вычислены изменения первого  $a_1$  и второго  $a_2$  коэффициентов МНК-аппроксимации на скользящем окне наблюдения размером  $w = 300$  минутных отсчетов. Соответствующий график для процесса и оценок скорости ( $a_1$ ) и ускорения ( $a_2$ ) представлен на рисунке 6.

Очевидно, что критические значения указанных коэффициентов, в случае их использования в задаче анализа инерционности хаотической динамики, необходимо выбирать с учетом динамического диапазона их изменений и значений *среднеквадратического отклонения (ско)*. Соответствующий диапазон для первого коэффициента  $a_1$  составляет для выбранного интервала наблюдения равен  $[-0.32919, 0.34136]$ , а для  $a_2$  -  $[-0.0025116,$

0.0022677]. Значения *ско* для этих же величин, соответственно равны (0.11397; 0.00078).

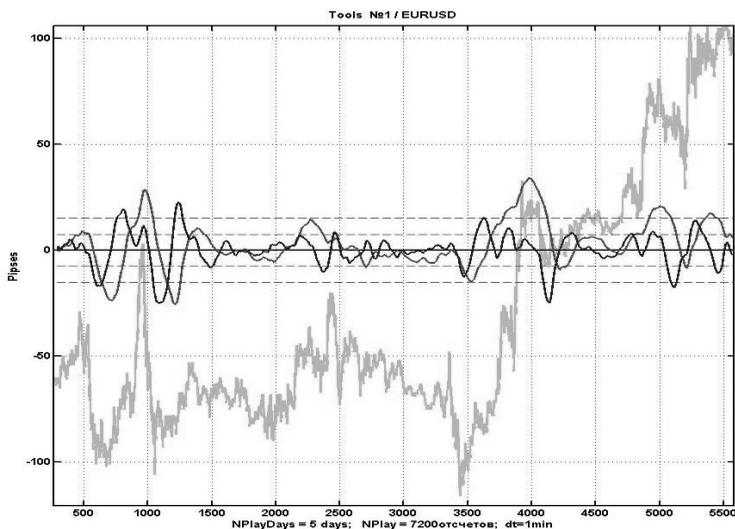


Рис. 6. Графики процесса, оценок скорости и ускорения

Для фиксации тренда значения параметра  $a_1$ , играющего роль скорости, в первом приближении можно выбрать на уровне  $0.5 \cdot \text{ско}$ , а именно,  $a_1^* = 0.057$ . Значение ускорения  $a_2$  должно быть положительным и желательно на восходящем интервале изменения. В качестве уровня подтверждения тренда выбрана величина  $dL = \pm 75n$ .

Результаты оценки вероятности подтверждения тренда на выбранном уровне  $dL$  для 4-х различных непересекающихся 100-дневных интервалов наблюдения и трех значений скользящего окна  $w = 100, 200, 300$  приведены в таблице 3.

Таблица 3. Частота подтверждения тренда для различных значений скользящего окна наблюдения

Частота подтверждения тренда				
w, мин	1-100дн.	101-200дн.	201-300дн.	301-400дн.
100	0.54	0.46	0.54	0.54
200	0.48	0.51	0.50	0.50
300	0.49	0.48	0.52	0.52

Из приведенных в таблице оценок видно практически полное отсутствие статистического подтверждения инерционности тренда, способного сохранить свою направленность на уровне  $dL$ .

**7. Заключение.** Из представленных выше результатов статистических экспериментов можно сделать вывод о том, что в условиях хаотической динамики наличие локального тренда является чисто случайным событием и процесс равновероятно развивается как в направлении развития выявленной тенденции, так и в противоположном направлении.

Обнаружения факта изменения параметров тренда всегда запаздывает по отношению к текущему состоянию хаотического процесса, что связано с использованием ретроспективных наблюдений. В условиях хаотической динамики такое запаздывание неизбежно приводит к проигрышу любую игровую стратегию, основанную на трендах.

Запаздывание процесса идентификации тренда можно уменьшить путем соответствующего уменьшения размера окна наблюдения. Однако такое сокращение неизбежно приводит к росту ложных тревог, т.е. «обнаружению» ложных трендов, что также ведет к проигрышной игровой стратегии.

Остается открытым вопрос о принципиальной возможности достижения положительного результата при использовании управляющей стратегии, основанной на анализе трендов. Можно предположить, что существует комбинация параметров игровой стратегии  $(w, dL, a_1^*, a_2^*, a_3^*)$ , обеспечивающее выигрыш при условной игре на апостериорных (ретроспективных) данных. По-видимому, данный результат не будет обладать устойчивостью и на новых данных даст явный проигрыш. Тем не менее, такой анализ необходим для получения окончательного вывода об инерционности трендов в хаотических процессах. Данному вопросу будут посвящены дальнейшие исследования.

### Литература

1. *Чежулаев М.* Торгуем в тренде // Валютный спекулянт. 2006. № 2. URL: [http://www.spekulant.ru/archive/Torguem\\_v\\_trende.html](http://www.spekulant.ru/archive/Torguem_v_trende.html).
2. *Гринблат Д.* Из прошлого в будущее // Forex magazine. 2014. №523/15. С. 22-24.
3. *Bosse T., Treur J., Jonker C.M.* Formal analysis of design process dynamics // AI EDAM. 2010. vol. 24. no. 3. pp. 397–423.
4. *Lien K.* Day Trading the Currency Market: Technical and Fundamental Strategies to Profit from Market Swings // NY: Wiley & Sons, Inc. 2005. 256 p.
5. *Свиргун О.* Торговля против тренда // Forex magazine. 2013. №470/13. С. 17–21.
6. *Лемон А.* В поисках Грааля // Forex magazine. 2014. №533/25. С. 25–26.
7. *Тайлер И. Т.* Поддержанное движение // Forex magazine. 2014. №511/3. С. 23–24.

8. *Ионина Н.* Технический анализ: линии канала, линии Фибоначчи, три линии тренда // *Forex magazine*. 2014. №533/25. С. 17–20.
9. *Фулер Н.* Как переиграть рынок // *Forex magazine*. 2014. №517/9. С. 23–24.
10. *Рафли С.* Техника торговли // *Forex magazine*. 2014. №511/3. С. 25–27.
11. *Чеботарев Ю.* Спор трейдеров о том, есть ли у рынка память // Валютный спекулянт. 2006. №11–12. С. 130–132.
12. *Мусаев А.А.* Статистический анализ инерционности хаотических процессов // Труды СПИИРАН. 2014. Вып. 2(33). С. 48–59.

## References

1. Tchekulaev M. [Trading in trend]. *Valjutnyj spekuljant – Currency speculator*. 2006. № 2. URL: [http://www.spekulant.ru/archive/Torguem\\_v\\_trende.html](http://www.spekulant.ru/archive/Torguem_v_trende.html). (In Russ.).
2. Greenblut D. [From past to future]. *Forex magazine*. 2014. no. 523/15. pp. 22–24. (In Russ.).
3. Bosse T., Treur J., Jonker C.M. Formal analysis of design process dynamics. *AI EDAM*. 2010. vol.24. no. 3. pp. 397–423.
4. Lien K. *Day Trading the Currency Market: Technical and Fundamental Strategies to Profit from Market Swings*. NY: Wiley & Sons, Inc. 2005. 256 p.
5. Svirgun O. [A trade against the trend]. *Forex magazine*. 2013. №470/13. pp. 17–21. (In Russ.).
6. Lemon A. [In the search of Grail]. *Forex magazine*. 2014. no. 533/25. pp. 25–26. (In Russ.).
7. Tailer I. [Supported movement]. *Forex magazine*. 2014. no. 511/3. pp. 23–24. (In Russ.).
8. Ionina N. [Technical analysis: channel lines, Fibonacci lines, three trends lines]. *Forex magazine*. 2014. no. 533/25. pp. 17–20. (In Russ.).
9. Fuller N. [How to over win the market]. *Forex magazine*. 2014. no. 517/9. pp. 23–24. (In Russ.).
10. Raflı C. [A trade technology]. *Forex magazine*. 2014. no. 511/3. pp. 25–27. (In Russ.).
11. Tcheboterev Jr. [Traders bet about market memory existence]. *Valjutnyj spekuljant – Currency speculant*. 2006. no. 11-12. pp. 130–132. (In Russ.).
12. Musaev A.A. [Numerical Analysis of Chaotic Processes Persistence]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2014. vol. 2(33). pp. 48–59. (In Russ.).

**Мусаев Александр Азерович** — д-р техн. наук, профессор, ведущий научный сотрудник лаборатории информационных технологий в системном анализе и моделировании СПИИРАН, декан факультета информационных технологий и управления Санкт-Петербургского государственного университета, научный консультант ОАО Специализированная инжиниринговая компания «Севзапмонтажавтоматика». Область научных интересов: анализ данных, управление и прогнозирование в сложных динамических системах, стохастические и хаотические системы. Число научных публикаций — 214. [amusaev@technolog.edu.ru](mailto:amusaev@technolog.edu.ru), [www.szma.com](http://www.szma.com); СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)350-5885, факс +7(812)350-1113.

**Musaev Alexander Azerovich** — Ph.D., Dr. Sci., professor, leading researcher, Laboratory of IT in System Analysis and Modeling, SPIIRAS, dean of IT and control systems department, St. Petersburg State Institute of Technology, expert, public corporation Specialized Engineering Company "Sevzapmontageautomatıca". Research interests: data analysis, complicated dynamic processes prognosis and control, stochastic and chaos systems. The number of publications — 215. [amusaev@szma.com](mailto:amusaev@szma.com), [www.szma.com](http://www.szma.com); SPIIRAS, 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)350-5885, fax +7(812)350-1113.

## РЕФЕРАТ

*Мусаев А.А.* **Численное исследование управляющих стратегий, основанных на анализе трендов.**

В рамках вычислительных экспериментов рассмотрим ряд возможных управляющих стратегий, в соответствии с которой обнаружение тренда одновременно является рекомендацией для открытия позиции в направлении выявленного тренда. Данная задача рассматривалась ранее с точки зрения инерционности наблюдаемого процесса. Однако количественные характеристики наблюдаемых тенденций никак не учитывались. Таким образом, приведенные в статье материалы следует рассматривать как продолжение и развитие исследований инерционности хаотических процессов. При этом инерционность процесса рассматривается как обоснование серии игровых стратегий, основанных на обнаружении и следовании выявленных тенденций.

В настоящей работе в качестве количественных характеристик тенденций рассматриваются параметры степенных полиномов, аппроксимирующих динамику изучаемых хаотических процессов.

Из представленных результатов статистических экспериментов можно сделать вывод о том, что в условиях хаотической динамики наличие локального тренда является чисто случайным событием и процесс равновероятно развивается как в направлении развития выявленной тенденции, так и в противоположном направлении.

Обнаружения факта изменения параметров тренда всегда запаздывает по отношению к текущему состоянию хаотического процесса, что связано с использованием ретроспективных наблюдений. В условиях хаотической динамики такое запаздывание неизбежно приводит к проигрышу любую игровую стратегию, основанную на трендах.

Запаздывание процесса идентификации тренда можно уменьшить путем соответствующего уменьшения размера окна наблюдения. Однако такое сокращение неизбежно приводит к росту ложных тревог, т.е. «обнаружению» ложных трендов, что также ведет к проигрышной игровой стратегии.

Остается открытым вопрос о принципиальной возможности достижения положительного результата при использовании управляющей стратегии, основанной на анализе трендов. По-видимому, такой результат не будет обладать устойчивостью и на новых данных даст явный проигрыш. Тем не менее, такой анализ необходим для получения окончательного вывода об инерционности трендов в хаотических процессах. Данному вопросу будут посвящены дальнейшие исследования.

## SUMMARY

### ***Musaev A.A.* Numerical research of the operating strategy based on the trends analysis.**

Within computing experiments are considered a number of possible control strategies according to which trend detection at the same time is the recommendation for position opening in the direction of the revealed trend. This task was considered earlier from the point of view of observed process persistence. However quantitative characteristics of observed tendencies weren't considered in any way. So, the given researches it is necessary to consider as prolongation and development research of chaotic processes lag effect. Thus the process persistence effect is considered as justification of the game strategy series based on revealed tendencies detection and its following.

In the present article as tendencies numerical characteristics are considered parameters of the power polynoms approximating of studied chaotic processes.

From the statistical experiments presented results it is possible to draw a conclusion that in the conditions of chaotic dynamics existence of a local trend is purely casual event. Process equiprobable develops both in the direction of the revealed tendency, and in an opposite direction.

Detection of the fact of trend parameters change always is late in relation to current state of chaotic process. That is connected with use of retrospective observations. In the chaotic dynamics conditions such delay inevitably leads any based on trends control strategy to loss.

Delay of a trend identification process can be reduced by the corresponding reduction of the watch window size. However such reduction inevitably leads to growth of false alarms, i.e. "detection" of false trends that also conducts to the losing control strategy.

There is open a question of basic possibility of advantageous strategy creation on the basis of trends. It is possible to assume that there is a combination of parameters of the game strategy, providing a winning at conditional game on the retrospective data. Apparently, this result won't possess stability and on new data will give obvious loss. Nevertheless, this analysis is necessary for receiving a final conclusion about a lag effect of trends in chaotic processes. This subject will be considered in the next research.

В.В. ВИХМАН, М.А. ПАНКОВ  
**ПОВЫШЕНИЕ СТОЙКОСТИ ХЕШ-ФУНКЦИЙ В  
ИНФОРМАЦИОННЫХ СИСТЕМАХ НА ОСНОВЕ АЛГОРИТМА  
МНОГОИТЕРАЦИОННОГО ХЕШИРОВАНИЯ С  
НЕСКОЛЬКИМИ МОДИФИКАТОРАМИ**

---

*Вихман В.В., Панков М.А. Повышение стойкости хеш-функций в информационных системах на основе алгоритма многоитерационного хеширования с несколькими модификаторами.*

**Аннотация.** В данной работе рассматривается влияние параметров алгоритма многоитерационного хеширования с несколькими модификаторами на его криптостойкость. Обоснована актуальность применения алгоритма многоитерационного хеширования с несколькими модификаторами и необходимость исследования его параметров, приводится описание алгоритма. Стойкость хеш-функции к атакам, не зависящим от алгоритма, обуславливается ее разрядностью, т.е. фактически – количеством уникальных значений, которое способна генерировать данная хеш-функция. Для оценки стойкости алгоритма к атакам методами «грубой силы», «дней рождения» и словарным атакам алгоритм многоитерационного хеширования с несколькими модификаторами рассматривается как самостоятельная хеш-функция. Оценку стойкости алгоритма при заданном количестве итераций предлагается производить путем вычисления средней разрядности эквивалентно стойкой хеш-функции для алгоритма. Приводится описание метода оценки стойкости алгоритма. Эксперименты производятся с использованием усеченной криптостойкой хеш-функции. Приводятся результаты экспериментов, позволяющие сравнить между собой показатели стойкости алгоритма при различных значениях его параметров. Кроме того, результаты экспериментов позволяют понять, как значения тех или иных параметров, а также сочетания значений этих параметров влияют на криптостойкость алгоритма к атакам методами «грубой силы», «дней рождения» и словарным атакам. На основании полученных результатов можно сделать выводы о значениях параметров, рекомендуемых для практического применения данного алгоритма. В заключении представлены основные результаты работы. Авторы статьи полагают, что алгоритм может найти применение в подсистемах аутентификации информационных систем, а также в системах, в которых наиболее важным требованием является стойкость в течение длительного времени.

**Ключевые слова:** хеш-функция, хеш-значение, хеш, аутентификация, атака, усеченная хеш-функция, алгоритм, многоитерационное хеширование, модификатор, эквивалентно стойкая хеш-функция, итерация, криптостойкость.

*Vikhman V.V., Pankov M.A. Security increasing of hash functions in information systems on the basis of multi-iterative hashing algorithm with several modifiers*

**Abstract.** In this paper influence of multi-iterative hashing with several modifiers algorithm's parameters on its cryptographic persistence is considered. Relevance of multi-iterative hashing with several modifiers algorithm's application and need of research of its parameters are justified, the description of algorithm is provided. Cryptographic persistence of hash function to attacks which are not depend on algorithm is caused by its bitness, i.e. actually on the amount of unique hash values that hash function is able to generate. For an estimation of algorithm's persistence to dictionary attacks and attacks by methods of "brute force" and "birthdays" the algorithm of multi-iterative hashing with several modifiers is considered as independent hash function. Estimation of the algorithm's persistence for a given number of

iterations is offered to produce by calculating the average bitness of equivalently persistent hash function for the algorithm. The description of estimation method of algorithm's persistence is provided. The experiments are performed using a truncated cryptographically persistent hash function. The results of experiments allow to compare the algorithm's persistence metrics of under different values of its parameters. Besides, the results of the experiments allow to understand how the values of certain parameters, and combinations of values for these parameters affect for the algorithm's cryptographic persistence to dictionary attacks and attacks by methods of "brute force" and "birthdays". On the basis of the received results it is possible to draw conclusions about the values of the parameters recommended for practical application of this algorithm. In conclusion, the paper presents the main results of the work. Authors of the article believe that the algorithm can find application in authentication subsystems of information systems, and also in systems where the most important requirement is persistence for a long time.

**Keywords:** hash function, hash value, hash, authentication, attack, truncated hash function, algorithm, multi-iterative hashing, modifier, bitness, equivalently persistent hash function, iteration, cryptographic persistence.

---

**Введение.** В настоящее время задача защита информации является одной из наиболее важных для информационных систем. Безопасность хранения паролей в подсистемах аутентификации ИС достигается с помощью использования криптографических хеш-функций. Криптографическая хеш-функция должна удовлетворять ряду требований [1 – 5]:

– *стойкость к вычислению прообраза* – невозможность нахождения неизвестного прообраза для любых предварительно заданных хеш-значений, т.е. для заданной хеш-функции  $h$  вычислительно невозможно найти неизвестный прообраз  $x$  при предварительно заданном хеш-значении  $y=h(x)$  для любого значения  $y$  (под термином «вычислительно невозможно» здесь и далее будем понимать, что алгоритм, выполняющий данное преобразование, обладает не менее чем экспоненциальной сложностью);

– *стойкость к вычислению второго прообраза* – невозможность нахождения любого другого прообраза, который давал бы такое же хеш-значение, как и заданный, т.е. для заданной хеш-функции  $h$  и прообраза  $x$  вычислительно невозможно найти другой прообраз  $x' \neq x$ , для которого выполнялось бы условие  $h(x)=h(x')$ ;

– *стойкость к коллизиям* – невозможность нахождения двух прообразов, для которых выработывалось бы одинаковое значение, т.е. для заданной хеш-функции  $h$  вычислительно невозможно найти два прообраза  $x$  и  $x'$ ,  $x' \neq x$ , для которых выполнялось бы условие  $h(x)=h(x')$ .

На криптографические хеш-функции возможны следующие атаки:

- нахождение прообраза  $x$  по заданному значению  $y=h(x)$ ;
- нахождение прообраза  $x'$  по заданному прообразу  $x$ , для которого выполняется условие  $h(x)=h(x')$ ;

– нахождение двух прообразов  $x$  и  $x'$ ,  $x \neq x'$ , для которых выполнялось бы условие  $h(x)=h(x')$ .

Атака “грубой силой” [1] может быть выполнена для нахождения прообраза по заданному хеш-значению или для нахождения прообраза, дающего заданное хеш-значение. Суть атаки заключается в последовательном или случайном переборе входных сообщений и сравнения результата выполнения хеш-функции с заданным. Сложность такой атаки оценивается  $2^{l-1}$  операций вычисления хеш-значений, где  $l$  – длина хеш-значения в битах.

Атака методом «дней рождения» [1, 6] выполняется для нахождения двух различных сообщений с одинаковыми хеш-значениями. Эта атака основана на парадоксе «дней рождения» и заключается в том, что в двух сгенерированных множествах хеш-значений, содержащих  $n_1$  и  $n_2$  элементов соответственно, вероятность нахождения совпадающих элементов между этими множествами оценивается следующей формулой:

$$P \approx 1 - e^{-\frac{n_1 n_2}{2^l}}. \quad (1)$$

В частности, при  $n_1 = n_2 = 2^{\frac{l}{2}}$  сложность атаки оценивается как  $2^{\frac{l}{2}+1}$  операций вычисления хеш-значений, а вероятность успеха равна:

$$P \approx 1 - \frac{1}{e} \approx 0,63. \quad (2)$$

Также существуют атаки, основанные на принципах атаки методом «грубой силы». Разновидностями этой атаки являются словарные атаки [7]. Название этого вида атак произошло благодаря тому, что основу множества перебираемых паролей составляют слова какого-либо языка. Словарные атаки используют присущую большинству пользователей тенденцию к использованию легко запоминаемых паролей, к которым часто относятся различные слова и их варианты (например, замена части букв слова на похожие по написанию цифры или спецсимволы). По сравнению с методом «грубой силы» словарные атаки осуществляют перебор по существенно меньшему множеству возможных значений [7, 8].

**2. Постановка задачи.** Для эффективного противостояния атакам методами «грубой силы» и «дней рождения» можно применять многокритериальное хеширование, которое позволяет повысить количество операций, необходимое для генерации хеш-значения, и, следовательно, усложнить указанные атаки [9]. Наиболее действенным

методом защиты от различных видов словарных атак [7, 8] является дополнение (например, путем конкатенации) хешируемого пароля какой-либо случайной величиной (модификатором), которая впоследствии хранится вместе с хешированным паролем. Следовательно, для определения пароля атакующий должен рассмотреть также все множество значений данной величины [7]. В качестве примеров функций, реализующих оба этих подхода (многоитерационное хеширование, использование модификатора), можно привести функции `bcrypt` и `scrypt` [10, 11].

Алгоритм многоитерационного хеширования с несколькими модификаторами [12], предлагаемый авторами, комбинирует в себе указанные выше методы защиты, и позволяет противостоять как атакам методами «грубой силы» и «дней рождения», так и словарным атакам. Данный алгоритм позволяет использовать несколько модификаторов, что следует из его названия. Алгоритм также может использоваться для повышения стойкости к рассмотренным атакам произвольной функции хеширования. Ввиду последнего обстоятельства актуальной задачей является анализ параметров алгоритма многоитерационного хеширования с несколькими модификаторами и их влияния на криптостойкость данного алгоритма.

**3. Алгоритм многоитерационного хеширования с несколькими модификаторами.** Алгоритм многоитерационного хеширования с несколькими модификаторами, предназначен для повышения стойкости применяемой в нем криптографической хеш-функции к атакам, не зависящим от алгоритма (атака «грубой силой», атака методом «дней рождения», словарные атаки) [12].

Рассмотрим алгоритм. Сначала с помощью криптостойкого генератора случайных чисел (ГСЧ) инициализируются  $k$  модификаторов разрядностью  $d$  бит каждый. Каждому модификатору присваивается индекс от 0 до  $k-1$  ( $k \in [1; 255]$ ). Затем вычисляется Z-хеш [13 – 17], который будет задавать порядок выбора модификаторов:

$$Z = h(x + h(x)), \quad (3)$$

где  $x$  – входное значение,  $h$  – функция хеширования, «+» – операция конкатенации [12]. Хеш-значение  $Z$  представим в виде массива байтов:

$$Z = \begin{pmatrix} z_0 \\ z_1 \\ \dots \\ z_{t-1} \end{pmatrix}. \quad (4)$$

На каждой итерации хеширования будет применяться один из модификаторов, выбираемый по индексу  $j$ :

$$j = z_i \bmod t \bmod k, \quad (5)$$

где  $i$  – номер итерации,  $t$  – размерность массива  $Z$  [12]. На первой итерации хеш вычисляется от конкатенации входного значения и модификатора, на последующих итерациях в качестве входного значения используется хеш, полученный на предыдущей итерации. Таким образом, входное значение будет косвенным образом определять порядок выбора модификаторов при многоитерационном хешировании. Так как инициализация модификаторов осуществляется при помощи криптостойкого ГСЧ, для повторного вычисления хеша необходимо сохранять модификаторы и порядок их следования.

Кроме количества итераций, алгоритм имеет три параметра: длина  $Z$ -хеша в битах, количество модификаторов, разрядность (битность) модификаторов. Определим, как значения этих параметров влияют на стойкость алгоритма к атакам методами «грубой силы» и «дней рождения».

Стойкость хеш-функции к таким атакам зависит от ее разрядности, т.е. фактически – от количества уникальных хеш-значений, которые способна генерировать данная хеш-функция [1, 12]. Для оценки стойкости алгоритма подадим на его вход множество, содержащее  $2^b$  уникальных значений, где  $b$  – разрядность оцениваемой хеш-функции, а затем подсчитаем количество уникальных хешей  $n$  во множестве хеш-значений. Тогда стойкость алгоритма можно оценить через разрядность  $s$  абстрактной эквивалентно стойкой хеш-функции:

$$s = \log_2 n. \quad (6)$$

В экспериментах использовалась усеченная 16-битная хеш-функция SHA-1. Использование усеченной хеш-функции возможно ввиду того, что криптографическая хеш-функция (в данном случае это криптографическая хеш-функция SHA-1) позволяет выбирать любой набор битов результата для формирования усеченной хеш-функции с аналогичными свойствами, хотя и, естественно, меньшей стойкости – соответственно своей новой разрядности [18 – 20]. Входные данные – беззнаковые двоичные 16-разрядные числа в диапазоне от 0 до 65535. Параметры алгоритма обозначаются в формате  $t/k/d$ , где  $t$  – разрядность  $Z$ -хеша,  $k$  – количество модификаторов,  $d$  – разрядность модификаторов. Так как значения модификаторов являются случайными, для каждого количества итераций будем проводить серию из пяти экспериментов, а среднюю разрядность эквивалентно

стойкой хеш-функции  $\bar{s}$  для текущего количества итераций будем рассчитывать следующим образом:

$$\bar{s} = \frac{1}{5} \sum_{q=1}^5 \log_2 n_q. \quad (7)$$

Первая серия экспериментов проводилась при использовании восьми 8-битных модификаторов при изменяемом значении разрядности Z-хеша (из 160-битного хеша SHA-1 выбиралось необходимое количество младших двоичных разрядов). Результаты представлены на рисунке 1.

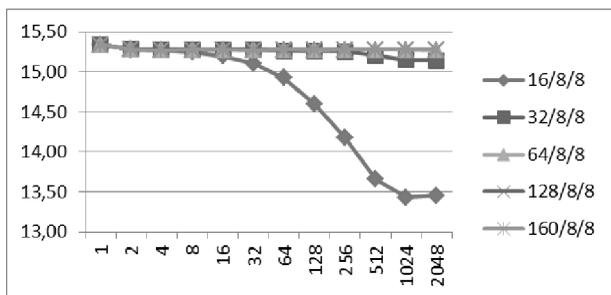


Рис. 1. Диаграммы зависимостей средних разрядностей эквивалентно стойких хеш-функций от количества итераций алгоритма (изменяемая разрядность Z-хеша, восемь 8-битных модификаторов)

Затем две серии экспериментов были проведены с использованием 16-битных и 32-битных модификаторов (рисунки 2, 3). Из полученных диаграмм видно, что с увеличением разрядности Z-хеша, при неизменном количестве и разрядности модификаторов, растут и средние разрядности эквивалентно стойких хеш-функций. Таким образом, можно сделать вывод о том, что разрядность Z-хеша напрямую влияет на стойкость алгоритма.

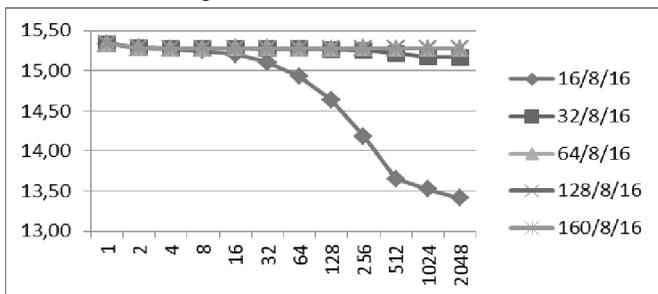


Рис. 2. Диаграммы зависимостей средних разрядностей эквивалентно стойких хеш-функций от количества итераций алгоритма (изменяемая разрядность Z-хеша, восемь 16-битных модификаторов)

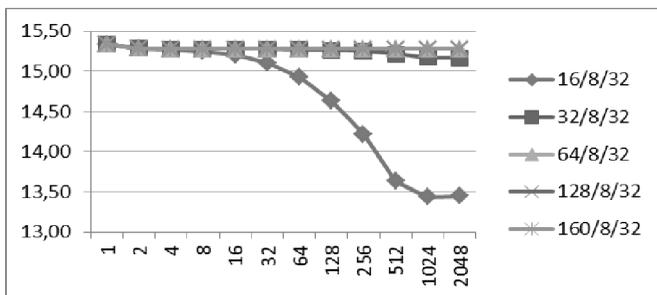


Рис. 3. Диаграммы зависимостей средних разрядностей эквивалентно стойких хеш-функций от количества итераций алгоритма (изменяемая разрядность Z-хеша, восемь 32-битных модификаторов)

Следует отметить, что разрядность модификаторов, при прочих равных параметрах алгоритма, не влияет на среднюю разрядность эквивалентно стойкой хеш-функции. При этом количество всех возможных значений набора из  $k$  модификаторов оценивается как  $2^{dk}$ .

Также было исследовано влияние количества используемых модификаторов на стойкость алгоритма. В экспериментах использовались полные значения Z-хешей при изменяемом количестве модификаторов. Проведено три серии экспериментов с использованием 8-битных, 16-битных и 32-битных модификаторов соответственно. Полученные диаграммы представлены на рисунках 4, 5 и 6.

Из полученных результатов можно сделать вывод, что чем меньше модификаторов применяется в алгоритме, тем быстрее, с увеличением количества итераций, растет количество коллизий, и, соответственно, падает стойкость эквивалентно стойкой хеш-функции алгоритма. Для того чтобы разрядность эквивалентно стойкой хеш-функции для алгоритма равнялась разрядности используемой в нем хеш-функции, необходимо применять не менее четырех модификаторов.

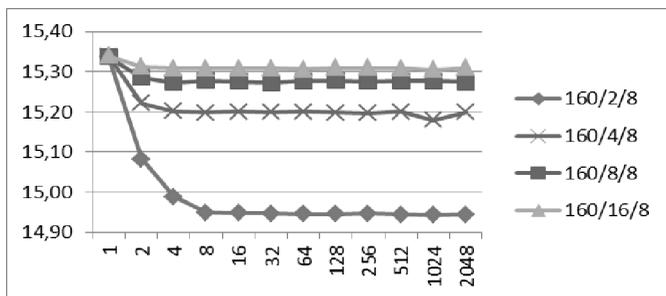


Рис. 4. Диаграммы зависимостей средних разрядностей эквивалентно стойких хеш-функций от количества итераций алгоритма (изменяемое количество 8-битных модификаторов, 160-битный Z-хеш)

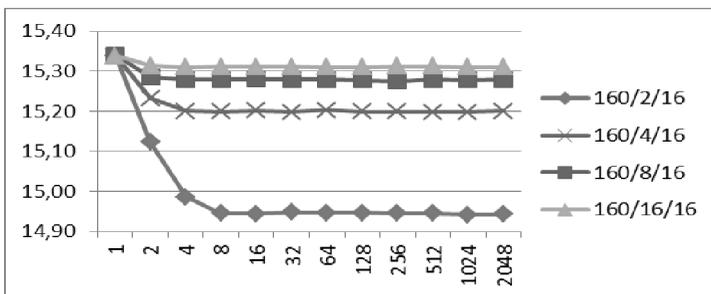


Рис. 5. Диаграммы зависимостей средних разрядностей эквивалентно стойких хеш-функций от количества итераций алгоритма (изменяемое количество 16-битных модификаторов, 160-битный Z-хеш)

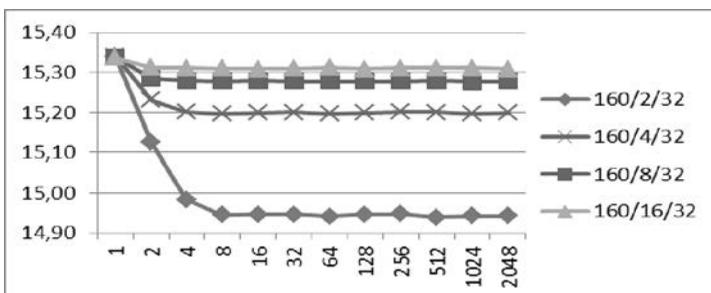


Рис. 6. Диаграммы зависимостей средних разрядностей эквивалентно стойких хеш-функций от количества итераций алгоритма (изменяемое количество 32-битных модификаторов, 160-битный Z-хеш)

Из данных диаграмм также видно, что разрядность модификаторов, при прочих равных параметрах алгоритма, не влияет на его стойкость.

**4. Заключение.** Алгоритм многоитерационного хеширования с несколькими модификаторами позволяет повысить трудоемкость вычисления хеш-значений с помощью существующих криптостойких хеш-функций. На стойкость алгоритма напрямую влияет разрядность Z-хеша: чем больше байтов содержит Z-хеш, тем выше средняя разрядность эквивалентно стойкой хеш-функции при одном и том же количестве итераций. Таким образом, в данном алгоритме рекомендуется использовать Z-хеш с разрядностью, равной полной разрядности применяемой в нем хеш-функции.

Количество модификаторов, применяемых в алгоритме, также влияет на его стойкость. При одинаковом количестве итераций средняя разрядность эквивалентно стойкой хеш-функции выше для алгоритма с большим количеством модификаторов. Для того чтобы не

происходило потери стойкости применяемой хеш-функции, в алгоритме необходимо использовать не менее четырех модификаторов. Разрядность модификаторов, при прочих равных параметрах алгоритма, не влияет на среднюю разрядность эквивалентно стойкой хеш-функции, но при этом влияет на сложность атак методами, не зависящими от алгоритма. Количество и разрядность модификаторов следует задавать исходя из требований к конкретной системе аутентификации.

Направление дальнейших исследований будет связано с особенностями внедрения разработанного алгоритма в подсистемы аутентификации пользователей информационных систем.

### Литература

1. *Вервейко В.Н., Пушкарев А.И., Ценурит Т.В.* Функции хэширования: классификация, характеристика и сравнительный анализ. URL: <http://bezopasnik.org/article/book/94.pdf>.
2. *Schneier B.* One-Way Hash Functions // *Dr. Dobbs' journal*. 1991. vol. 16. no. 9. pp. 148–151.
3. *Biham E., Shamir A.* Differential cryptanalysis of FEAL and NHash // In *Advances in Cryptology (Eurocrypt '91)*. 1990. pp. 1–16.
4. *Biham E.* On the Applicability of Differential Cryptanalysis to Hash Functions // In *E.I.S.S Workshop on Cryptographic Hash Functions*. 1992. pp. 25–27.
5. *Quisquater J.-J., Delescaille J.-P.* How Easy is Collision Search. New results and applications to DES // In *Advances in Cryptology (CRYPTO'89)*. 1990. vol. 435. pp. 408–415.
6. *Ohta K., Koyama K.* Meet-in-the-Middle Attack on Digital Signature Schemes // In *Abstract of AUSCRYPT '90*. 1990. pp. 110–121.
7. *Панасенко С.П.* Словарные атаки на хэш-функции // *Мир и безопасность*. 2009. № 4. С. 24–31.
8. *Коржик В.И., Пантелеева З.А.* Исследование метода радужных таблиц для восстановления паролей // *Актуальные проблемы инфотелекоммуникаций в науке и образовании. II-я Международная научно-техническая и научно-методическая конференция: сб. научных статей*. СПб.: Санкт-Петербургский государственный университет телекоммуникаций им. проф. М. А. Бонч-Бруевича. 2013. С. 824–829.
9. *Лёвин В.Ю.* О повышении криптостойкости однонаправленных хеш-функций // *Фундаментальная и прикладная математика*. 2009. Т. 15:5. С. 171–179.
10. *Provos N., Mazières D.* A Future-Adaptable Password Scheme // *The OpenBSD Project*. URL: <http://www.openbsd.org/papers/bcrypt-paper.pdf>.
11. *Percival C., Josefsson S.* The scrypt Password-Based Key Derivation Function // *IETF*. 2012.
12. *Вихман В.В., Панков М.А.* Исследование криптостойкости алгоритмов многонтерационного хэширования в подсистемах аутентификации МИС // *Актуальные проблемы электронного приборостроения (АПЭП–2014): тр. 12 междунар. конф. Новосибирск : Изд-во НГТУ*. 2014. Т. 2. С. 193–199.
13. *Шнайер Б.* Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке Си // М.: Триумф. 2002. 610 с.
14. *Фергюсон Н., Шнайер Б.* Практическая криптография // М.: Издательский дом «Вильямс». 2005. С. 101–114.

15. *Алферов А.П., Зубов А.Ю., Кузьмин А.С. и др.* Основы криптографии: учебное пособие / 2-е изд., испр. и доп. // М.: Гелиос АРВ. 2002. 480 с.
16. *Menezes A.J., Van Oorschot C., Vanstone S.A.* Handbook of applied cryptography // CRC Press. Boca Raton. New York. 1997.
17. Введение в криптографию / Под общ. ред. В. В. Ященко. 4-е изд., доп. // М.: МЦНМО. 2012. 348 с.
18. *Rivest R.* The MD5 Message Digest Algorithm // RFC 1321. 1992.
19. National Institute of Standards and Technology. Secure Hash Standard // FIPS PUB 180-1. 1995.
20. National Institute of Standards and Technology. Secure Hash Standard (draft) // DRAFT FIPS PUB 180-2. 2001.

## References

1. Vervejko V.N., Pushkarev A.I., Cepurit T.V. *Funksii kshirovaniia: klassifikatsiia, kharakteristika i sravnitel'nyi analiz* [Hashing functions: classification, characterization and comparative analysis] Available at <http://bezopasnik.org/article/book/94.pdf> (accessed 12.12.2014) (In Russ.).
2. Schneier B. One-Way Hash Functions. Dr. Dobbs' journal. 1991. vol. 16. no. 9. pp. 148–151.
3. Biham E., Shamir A. Differential cryptoanalysis of FEAL and NHash. In Advances in Cryptology (Eurocrypt '91). 1990. pp. 1–16.
4. Biham E. On the Applicability of Differential Cryptoanalysis to Hash Functions. In E.I.S.S Workshop on Cryptographic Hash Functions. 1992. pp. 25–27.
5. How Easy is Collision Search. New results and applications to DES. In Advances in Cryptology (CRYPTO'89). 1990. vol. 435. pp. 408–415.
6. Ohta K., Koyama K. Meet-in-the-Middle Attack on Digital Signature Schemes. In Abstract of AUSCRYPT '90. 1990. pp.110–121.
7. Panasenko S. [Dictionary attacks on hash functions]. *Mir i bezopasnost' – World and security*. 2009. no. 4, pp. 24–31. (In Russ.).
8. Korzhik V., Panteleeva Z. [Rainbow Tables Method for Cryptanalytic Time Memory Trade-Offs]. *Aktual'nye problemy infotelekkommunikacij v nauke i obrazovani. II-ja Mezhdunarodnaja nauchno-tehnicheskaja i nauchno-metodicheskaja konferencija: sb. nauchnyh statej* [Actual problems of telecommunications in science and education. II international scientific-technical and scientific-methodical conference: collection of scientific articles]. St. Petersburg. 2013. pp. 824–829. (In Russ.).
9. Ljovin V. [About increasing of cryptographic security unidirectional hash functions]. *Fundamental'naja i prikladnaja matematika – Fundamental and applied mathematics*. 2009. vol. 15:5. pp. 171–179. (In Russ.).
10. Provos N., Mazieres D. A Future-Adaptable Password Scheme. Available at <http://www.openbsd.org/papers/bcrypt-paper.pdf> (accessed 12.12.2014)
11. Percival C., Josefsson S. The scrypt Password-Based Key Derivation Function. IETF. 2012.
12. Vikhman V., Pankov M. [Research of multi-iterative hashing algorithms' cryptographic persistence in authentication subsystems of MIS]. *Trudy 12 mezhdunarodnoi konferentsii «Aktual'nye problemy elektronnoho priborostroeniia (APEP-2014)»* [Proc. 12th Int. Conf. “Actual problems of electronic instrument engineering (APEIE-2014)”]. Novosibirsk. 2014. pp. 193–199. (In Russ.).
13. Schneier B. *Prikladnaja kriptografija. Protokoly, algoritmy, ishodnye teksty na jazyke Si* [Applied Cryptography. Protocols, algorithms, and Source Code in C]. Wiley. 1995. 662 p. (Russ. ed.: Shnaier B. *Prikladnaia kriptografiia. Protokoly, algoritmy, iskhodnye teksty na iazyke Si*. Moskow. Triumf Publ. 2002. 610 p.).

14. Ferguson N., Schneier B. *Prakticheskaja kriptografija* [Practical Cryptography]. Wiley. 2003. 432 p. (Russ. ed.: Ferguson N., Shnaier B. *Prakticheskaja kriptografiia*. Moscow. Izdatel'skii dom «Vil'iams» Publ. 2005. 424 p.).
15. Alferov A., Zubov A., Kuz'min A., Cheremushkin A. *Osnovy kriptografii* [Foundations of cryptography]. Moscow. Gelios ARV Publ. 2002. 480 p. (In Russ.).
16. Menezes A.J., Van Oorschot P.C., Vanstone S.A. *Handbook of applied cryptography*. CRC Press. Boca Raton. New York. 1997. 816 p.
17. Iashchenko V. *Vvedenie v kriptografiu* [Introduction to cryptography]. Moscow. MTsNMO. 2012. 348 p. (In Russ.).
18. Rivest R. The MD5 Message Digest Algorithm. RFC 1321. 1992.
19. National Institute of Standards and Technology. *Secure Hash Standard*. FIPS PUB 180-1. 1995.
20. National Institute of Standards and Technology. *Secure Hash Standard (draft)*. DRAFT FIPS PUB 180-2. 2001.

**Вихман Виктория Викторовна** — к-т техн. наук, доцент, заведующий кафедрой интеграционных информационных систем Новосибирского государственного технического университета (НГТУ), доцент кафедры вычислительной техники факультета автоматки и вычислительной техники Новосибирского государственного технического университета (НГТУ). Область научных интересов: автоматизированные системы управления, системы информационной безопасности. Число научных публикаций — 58. vvv@vt.cs.nstu.ru, www.nstu.ru; НГТУ, 630073, г. Новосибирск, пр. Карла Маркса, 20, РФ; п.т. +7(383)3460492, +7(383)346-0219.

**Vikhman Victoria Victorovna** — Ph.D, associate professor, Head of Department of information systems' integration, Novosibirsk State Technical University (NSTU), associate professor, Department of Computer Engineering, Faculty of Automation and Computer Engineering, Novosibirsk State Technical University (NSTU). Research interests: automated control systems, information security systems. The number of publications — 58. vvv@vt.cs.nstu.ru, www.nstu.ru; NSTU, 20, prospect Karla Marksa, Novosibirsk, 630073, Russian Federation; office phones +7(383)346-0492, +7(383)346-0219.

**Панков Максим Александрович** — аспирант кафедры вычислительной техники факультета автоматки и вычислительной техники Новосибирского государственного технического университета. Область научных интересов: системы информационной безопасности, алгоритмы и модели управления персоналом в интегрированной информационной среде. Число научных публикаций — 4. hm.mobile@mail.ru; НГТУ, 630073, г. Новосибирск, пр. Карла Маркса, 20, РФ; п.т. +7(383)3460492, +7(383)346-0219.

**Pankov Maksim Aleksandrovich** — Ph.D. student, Department of Computer Engineering, Faculty of Automation and Computer Engineering, Novosibirsk State Technical University (NSTU). Research interests: information security systems, algorithms and models of human resource management in integrated information environment. The number of publications — 4. hm.mobile@mail.ru; NSTU, 20, prospect Karla Marksa, Novosibirsk, 630073, Russian Federation; office phones +7(383)346-0492, +7(383)346-0219.

## РЕФЕРАТ

### ***Вихман В.В., Панков М.А. Повышение стойкости хеш-функций в информационных системах на основе алгоритма многоитерационного хеширования с несколькими модификаторами.***

В статье рассматривается анализ параметров алгоритма многоитерационного хеширования с несколькими модификаторами, разработанного авторами. Главной особенностью алгоритма является возможность его применения для повышения стойкости различных криптографических хеш-функций к атакам, не зависящим от алгоритма. Следующие параметры задают настройки алгоритма: количество итераций, количество байт Z-хеша, количество модификаторов, разрядность (битность) модификаторов. Эксперименты производятся для оценки влияния значений различных параметров алгоритма на стойкость применяемой в нем хеш-функции. Оценка производится путем вычисления средних разрядностей эквивалентно стойких хеш-функций. Из результатов экспериментов следует, что количество байт Z-хеша напрямую влияет на стойкость хеш-функции. Разрядность Z-хеша следует выбирать равной полной разрядности применяемой хеш-функции. Также влияние на стойкость хеш-функции оказывает количество применяемых в алгоритме модификаторов. Увеличение количества модификаторов, при прочих равных значениях параметров алгоритма, приводит к повышению стойкости хеш-функции. Разрядность модификаторов на среднюю разрядность эквивалентно стойкой хеш-функции не влияет. Параметры алгоритма необходимо подбирать, исходя из требований к конкретной системе аутентификации.

## SUMMARY

### ***Vikhman V.V., Pankov M.A. Security increasing of hash functions in information systems on the basis of multi-iterative hashing algorithm with several modifiers.***

The article is considers the analysis of multi-iterative hashing with several modifiers algorithm's parameters, developed by the authors. The main feature of the algorithm is the possibility of its application to increase the security of different cryptographic hash functions to attacks that do not depend on the algorithm. The following parameters specify the settings of the algorithm: the number of iterations, the number of bytes in the Z-hash, the number of modifiers, the bitness of modifiers. The experiments are carried out to estimate the impact of different values for the algorithm's parameters on the security of hash function applied in it. The estimation is carried out by calculating the average bitnesses of equivalently persistent hash functions. From the experimental results it follows that the number of bytes in the Z-hash directly affects for security of the hash function. Bitness of Z-hash should be chosen equal to the full bitness of the used hash function. Also the impact to security of the hash function provides by the number of modifiers used in the algorithm. The increasing in the number of modifiers, under other equal values of algorithm's parameters, increases the security of the hash function. The bitness of the modifiers on the average bitness of equivalently persistent hash function has no effect. The parameters of algorithm should be selected based on requirements to specific authentication system.

М.С. ЩЕКотов  
**АНАЛИЗ ПОДХОДОВ К ПОЗИЦИОНИРОВАНИЮ ВНУТРИ ПОМЕЩЕНИЙ С ИСПОЛЬЗОВАНИЕМ ТРИЛАТЕРАЦИИ СИГНАЛОВ WI-FI**

---

*Щекотов М.С. Анализ подходов к позиционированию внутри помещений с использованием трилатерации сигналов Wi-Fi.*

**Аннотация.** В статье рассмотрены два основных метода позиционирования внутри помещений с помощью мобильного устройства: метод трилатерации сигналов Wi-Fi и метод радиотпечатков (fingerprinting). Кроме того, рассмотрены подходы к позиционированию внутри помещений, различающиеся определением расстояния до источника сигнала, которые основываются на вышеуказанных методах и открывают перспективу создания комбинированного метода позиционирования.

**Ключевые слова:** определение местоположения внутри помещений, метод радиотпечатков, триангуляция, Wi-Fi.

*Shchekotov M.S. Indoor Localization Method Based on Wi-Fi Trilateration Technique.*

**Abstract.** This paper describes two main indoor positioning methods for mobile devices based on Wi-Fi trilateration and fingerprinting. Two indoor positioning approaches based on these methods and having different schemes of distance between transmitter and receiver calculation are considered. These approaches open a possibility of hybrid indoor positioning method implementation.

**Keywords:** indoor localization, fingerprinting, triangulation, Wi-Fi.

---

**1. Введение.** Разработка систем для решения проблем позиционирования и навигации внутри зданий является перспективной, актуальной и наукоемкой задачей. Данная проблема требует создания карт на основе поэтажных планов помещений, выбора эффективной технологии позиционирования и алгоритмов, а также развертывания соответствующей инфраструктуры для позиционирования внутри зданий. Современные системы определения местоположения внутри помещений, например, как Navizon [1] или Wi-FiSLAM [2], могут обеспечить точность позиционирования до 3 метров. Системы определения местоположения внутри помещений могут использовать различные беспроводные технологии передачи данных для определения местоположения, например, Bluetooth, Wi-Fi, ZigBee [3] и сигналы сетей сотовой связи. Использование Wi-Fi более предпочтительно для решения задачи позиционирования мобильного устройства пользователя (смартфона, планшета и т.п.) в зданиях, так как технология Wi-Fi имеет в настоящий момент широкое распространение как на современных мобильных устройствах, так и в большинстве зданий, что позволяет действовать для этих целей уже существующую инфраструктуру [4].

Подходы к определению местоположения внутри помещения с помощью беспроводных технологий передачи данных могут основыв-

ваться на геометрии распространения сигнала, вероятностных моделях или построении карт сигналов (метод радиопечатков). Трилатерация – метод определения абсолютных или относительных местоположений геодезических пунктов, с помощью построения системы смежных треугольников [5]. Данный метод широко используется в геодезии. Для того чтобы определить координаты точки на местности, необходимо знать координаты не менее трех точек, а также расстояния до них. Далее, используя систему уравнений трех сфер (1), можно найти точку их пересечения, которая и будет соответствовать искомой точке:

$$\begin{aligned}r_1^2 &= (x-x_1)^2 + (y-y_1)^2 + (z-z_1)^2, \\r_2^2 &= (x-x_2)^2 + (y-y_2)^2 + (z-z_2)^2, \\r_3^2 &= (x-x_3)^2 + (y-y_3)^2 + (z-z_3)^2,\end{aligned}\tag{1}$$

где  $x_1, x_2, x_3, y_1, y_2, y_3, z_1, z_2, z_3$  – координаты точки доступа,  $r_1, r_2, r_3$  – радиусы сфер. При позиционировании в радионавигации, используется тот же способ, но вместо опорных точек на местности используются координаты базовых станций связи или координаты точек доступа Wi-Fi, как в нашем случае. Однако встает вопрос о том, как определить расстояния до источников сигнала, на основе которых рассчитывается область местоположения (область локализации). Расстояния до точек доступа могут быть определены с помощью измерения уровня принятого сигнала (RSS) [6], времени прибытия радиосигналов от передатчиков (ToA) [7] или определения разницы моментов времени прихода радиосигналов (TDoA) [8].

Метод радиопечатков [9], в свою очередь, основывается на построенной карте уровней сигналов в помещении, и не требует использования моделей распространения радиосигнала внутри помещения. Для того чтобы определить местоположение этим методом, создается база данных, содержащая информацию об уровнях сигнала в различных участках помещения, а также план помещения. Карта уровня сигналов должна состоять из реальных координат и связанных с этими значениями уровней сигналов от доступных точек доступа Wi-Fi. Местоположение пользователя может быть оценено с помощью измерения уровня сигналов точек доступа, и поисков точных совпадений в базе данных или ближайшей к приемнику точки, для которой измерены уровни сигнала [10, 11]. Данный метод обеспечивает высокую точность определения местоположения, однако требует поддержки карты сигналов в актуальном состоянии.

Далее будут рассмотрены подходы к позиционированию внутри помещений с помощью трилатерации сигналов Wi-Fi, и комбинации трилатерации и построения карты зон уровней сигналов Wi-Fi.

**2. Подходы к определению местоположения.** Метод трилатерации используется в соответствии с тем же принципом, что и в геодезии,

но с той разницей, что он использует сигналы беспроводных сетей передачи данных. Основной проблемой является разработка способа определения расстояния между источником сигнала и приемником. В качестве исходных данных для трилатерации используются такие параметры известных сетей Wi-Fi, как частота Wi-Fi сигнала, его уровень, сетевой MAC-адрес и реальные координаты точек доступа Wi-Fi. Полученные с помощью мобильного устройства уровни сигнала могут быть использованы для оценки расстояния между точкой доступа и мобильным устройством. При использовании этого метода рассматриваются три или более точек доступа, находящихся в здании на одном этаже. Уровни сигналов этих точек понижаются экспоненциально в зависимости от расстояния между передатчиком и приемником, и от фактора шума. Таким образом, эту зависимость можно рассматривать, как функцию от расстояния. Расстояние по оценкам силы сигнала представлено в виде круга с радиусом вокруг точки доступа. Пересечение трех точек доступа радиусов определяют точку или зону приемника. Эта модель может быть представлена в виде следующей системы уравнений:

$$\begin{aligned} r_1^2 &= (x-x_1)^2 + (y-y_1)^2, \\ r_2^2 &= (x-x_2)^2 + (y-y_2)^2 \\ r_3^2 &= (x-x_3)^2 + (y-y_3)^2 \end{aligned} \quad (2)$$

где  $x_1, x_2, x_3, y_1, y_2, y_3$  – координаты точки доступа,  $r_1, r_2, r_3$  – оцененные расстояния. Решение этой системы уравнений дает точки пересечения окружностей, определяющих область локализации (рисунок 1).

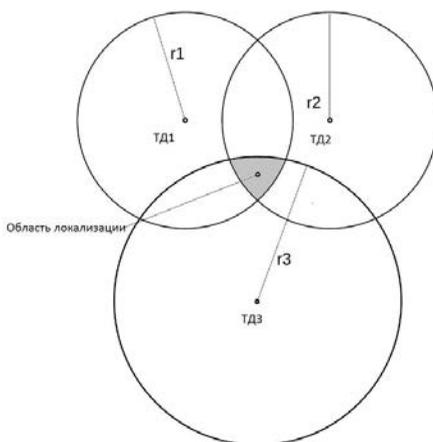


Рис. 1. Определение области местоположения объекта методом трилатерации

Однако для этого необходимо определить расстояния между приемником и каждой точкой доступа, используя данные об уровнях сигнала. На рисунке 1 каждый из трех радиусов соответствует уровню сигнала внутри окружностей. Эта задача может быть решена с помощью создания модели распространения сигнала внутри здания. Отражение и поглощение препятствиями при распространении внутри помещения являются основными причинами ошибок определения местоположения. Затухание зависит от частоты сигнала, расположения приемника относительно препятствий и отражающих поверхностей и многих других факторов [10]. Один из возможных подходов для оценки расстояния между приемником и передатчиком является использование модели распространения сигнала в свободном пространстве, для чего может использоваться формула потерь мощности сигнала в свободном пространстве [11, 12]:

$$FSPL = 20\log_{10}(d) + 20\log_{10}(f) - 27.55, \quad (3)$$

где,  $d$  – расстояние между источником и приемником сигнала,  $f$  – частота сигнала, FSPL – потеря мощности сигнала в дБ.

В ходе исследования, с помощью приложения, написанного для Android-смартфона, были измерены средние значения уровня сигнала Wi-fi для одной точки доступа. В таблице 1 измерения, произведенные с помощью Android-устройства, сравниваются с измерениями, рассчитанными на основе уравнения 1, при которых из начального уровня сигнала, измеренного Android-устройством, вычитаются потери мощности сигнала.

Таблица 1. Сравнение измеренного и оцененного уровней сигнала

Расстояние, м	Реальный уровень сигнала, дБм	Оцененный уровень сигнала, дБм
1	-33.3	-51.0
2	-45.7	-58.0
3	-50.9	-61.5
4	-51.7	-64.0
5	-51.8	-65.9
6	-53.4	-67.5

Измерения сделаны на расстоянии от 1 до 6 метров в 6 точках по одной линии в пределах помещения, в котором расположена точка доступа Wi-Fi. Уровень сигнала измеряется 10 раз в каждой из этих 6 точек. Площадь помещения составляет 25 квадратных метров. Измерение производится для сети с частотой сигнала 2412 МГц.

Представленное в таблице 1 сравнение показывает, что модель распространения сигнала в свободном пространстве является не эф-

фективной для использования в пределах одной комнаты. Этот подход требует создания модели, учитывающей отражение и преломление сигнала, которая будет обеспечивать более высокую точность. Кроме того, одним из решений проблемы повышения точности является использование вероятностной оценки параметров распространения сигнала.

Другим подходом может служить создание карты радиосигналов точек доступа Wi-Fi. Предлагается, однако, делать не полную карту радиосигналов, а определить средний сигнал на заданном расстоянии от источника. В данной работе уровни сигнала измерялись по расстоянию приемника от трех точек доступа, распределенных в трех комнатах в пределах одного этажа. Необходимые данные об уровнях сигнала на различном удалении от точек доступа были предварительно собраны, для оценки расстояния. Измерения были произведены в 15 точках в интервале 1 метр для каждой точки доступа с использованием разработанной программы для Android-устройства. Это приложение обнаруживает три различных точки доступа по заранее заданным MAC-адресам и измеряет уровни сигнала 10 раз для каждой из 15 точек от каждой точки доступа. Так как величина уровня сигнала изменяется во времени, то было использовано ее среднее значение. Уровни сигналов для всех точек доступа отражены в таблице 2.

Таблица 2. Уровни сигналов трех рассмотренных точек доступа (ТД)

Расстояние, м	Уровень сигнала ТД 1, дБм	Уровень сигнала ТД 2, дБм	Уровень сигнала ТД 3, дБм
1	-33.3	-38.8	-55.3
2	-45.7	-43.1	-50.3
3	-50.9	-48.9	-65.7
4	-51.7	-55.2	-61.2
5	-51.8	-75.1	-62,5
6	-53.4	-75.5	-66.4
7	-57.8	-76.4	-70.5
8	-62.4	-80.8	-72.3
9	-65.7	-80.8	-74.7
10	-62.9	-76.0	-78.0
11	-72.9	-88.6	-76.07
12	-72.7	-88.2	-86.02
13	-63.9	-91.0	-79.03
14	-74.0	-91.9	-85.08
15	-76.7	-92.1	-82.05

Полученные измерения на заданных расстояниях могут быть выбраны для оценки расстояния между источником и приемником в

качестве опорных точек. Опорные точки являются точками с разницей уровня сигнала большей, чем ошибка измерения, рассчитанная для каждой из 15 точек измерения. Таким образом, можно определить расстояние по уровню сигнала, действующему в области между двумя концентрическими окружностями (рисунок 2).

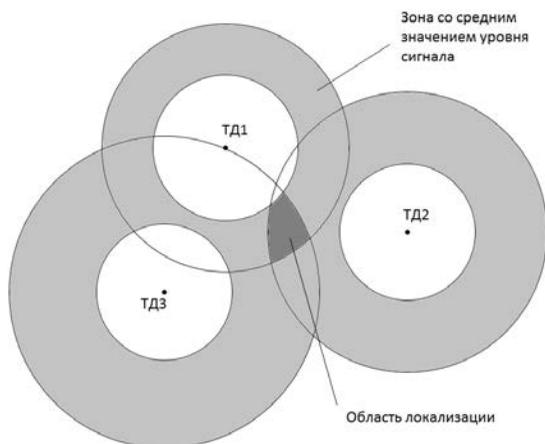


Рис. 2. Определение области вероятного местоположения приемника

Уровни сигнала измеряются с помощью написанного для этой цели приложения для Android-смартфона, которое определяет уровень сигнала сети Wi-Fi с заданным MAC-адресом, используя системные функции Android. Для определения области местоположения приемника рассчитывается пересечение областей между концентрическими окружностями, которые, в свою очередь, соответствуют заранее измеренным уровням сигналов Wi-Fi. Данное приложение использует java.awt библиотеку для построения геометрических фигур.

**3. Заключение.** В статье были проанализированы два подхода к позиционированию внутри помещений. Подход, основанный на трилатерации сигналов Wi-Fi для определения местоположения внутри помещений, описанный в данной статье, и использующий модель распространения сигнала, основанную на формуле потерь мощности в свободном пространстве, не способен обеспечить приемлемую точность. Использование подхода, основанного на методе радиопечатков, является приемлемым и может значительно увеличить точность. Для ее улучшения необходимо расширить выборку измерений уровней сигнала за счет увеличения количества опорных точек или точек доступа. Более того, дальнейшая работа может быть продолжена на основе использования метода радиопечатков, так как подход к определению

нию местоположения, описанный выше, можно рассматривать как частный случай метода радиопечатков. Реализация метода радиопечатков требует большего числа измерений уровня сигнала и создание карты радиопечатков, что может обеспечить высокую точность определения местоположения.

### Литература

1. Официальный сайт NAVIZON. URL: [www.navizon.com](http://www.navizon.com) (дата обращения: 20.11.2014).
2. Ferris B., Fox D., Lawrence N. WiFi-SLAM Using Gaussian Process Latent Variable Models // Proceedings of International Joint Conference on Artificial Intelligence (IJCAI). 2007.
3. Benkic K., Malajner M., Planinšic P., Cucej Ž. Using RSSI value for distance estimation in Wireless sensor networks based on ZigBee // Proceedings of IWSSIP Conference. 2008. pp. 303–306.
4. Kshevnik A., Shchekotov M. Comparative Analysis of Indoor Positioning Systems Based on Communications Supported by Smartphones // Proceedings of FRUCT Conference. 2012. pp. 43–48.
5. Ахромеев С. Ф. Институт военной истории Военный энциклопедический словарь // Воен. изд-во. 1986. 863 с.
6. Kaemarungsi K., Krishnamurthy P. Modeling of Indoor Positioning Systems Based on Location Fingerprinting // Proceedings of INFOCOM Conference. 2004. pp. 1012–1022.
7. Ali A.A., Omar A S. Time of Arrival Estimation for WLAN Indoor Positioning Systems using Matrix Pencil Super Resolution Algorithm // Proceedings of the 2nd Workshop on Positioning, Navigation and Communication. 2005.
8. Aatique M. Evaluation of TDOA Techniques for Position Location in CDMA Systems // Thesis submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of Master of Science. 1997.
9. Oguejiofor O., Okorogu V., Adewale A., Osuesu B. Outdoor Localization System Using RSSI Measurement of Wireless Sensor Network // Proceedings of International Journal of Innovative Technology and Exploring Engineering. 2013. pp. 1–6.
10. Henniges R. Current approaches of Wi-Fi Positioning // TU-Berlin. 2012. pp. 1–8.
11. Aleshy F., Mohd Sabri R. Sevak Z., Arslan T. Improving Indoor Positioning Accuracy through a Wi-Fi Handover Algorithm // Proceedings of International Technical Meeting of the Institute of Navigation. 2010. pp. 822–829.
12. Debus W. RF Path Loss & Transmission Distance Calculations // Axonn. Technical Memorandum. 2006. pp. 1–13.

### References

1. Official'nyj sajt NAVIZON [Official web site of Navizon, Inc.]. Available at: [www.navizon.com](http://www.navizon.com). (accessed 20.11.2014). (In Russ.).
2. Ferris B., Fox D., Lawrence N. WiFi-SLAM Using Gaussian Process Latent Variable Models. Proceedings of International Joint Conference on Artificial Intelligence (IJCAI). 2007.
3. Benkic K., Malajner M., Planinšic P., Cucej Ž. Using RSSI value for distance estimation in Wireless sensor networks based on ZigBee. Proceedings of IWSSIP Conference. 2008. pp. 303–306.

4. Kashevnik A., Shchekotov M. Comparative Analysis of Indoor Positioning Systems Based on Communications Supported by Smartphones. Proceedings of FRUCT Conference. 2012. pp. 43–48.
5. Akhromeev S. F. *Voennyj jenciklopedicheskij slovar'* [Military Encyclopaedic Dictionary]. Moscow: Voenizdat. 1986. p.863 (In Russ.).
6. Kaemarungsi K., Krishnamurthy P. Modeling of Indoor Positioning Systems Based on Location Fingerprinting. Proceedings of INFOCOM Conference. 2004. pp. 1012–1022.
7. Ali A.A., Omar A.S. Time of Arrival Estimation for WLAN Indoor Positioning Systems using Matrix Pencil Super Resolution Algorithm. Proceedings of the 2nd Workshop on Positioning, Navigation and Communication. 2005.
8. Aatique M. Evaluation of TDOA Techniques for Position Location in CDMA Systems. Thesis submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of Master of Science. 1997.
9. Oguejiofor O., Okorogu V., Adewale A., Osuesu B. Outdoor Localization System Using RSSI Measurement of Wireless Sensor Network. Proceedings of International Journal of Innovative Technology and Exploring Engineering, 2013. pp. 1–6.
10. Henniges R. Current approaches of Wi-Fi Positioning. TU-Berlin. 2012. pp. 1–8.
11. Aleshly F., Mohd Sabri R. Sevak Z., Arslan T., Improving Indoor Positioning Accuracy through a Wi-Fi Handover Algorithm. Proceedings of International Technical Meeting of the Institute of Navigation, 2010. pp. 822–829. (In Russ.).
12. Debus W. RF Path Loss & Transmission Distance Calculations. Axonn. Technical Memorandum. 2006. pp. 1-13.

**Щекотов Максим Сергеевич** — младший научный сотрудник лаборатории интегрированных систем автоматизации федерального государственного бюджетного учреждения Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: системы поддержки принятия решений, системы навигации. Число научных публикаций — 5. shekotov@iias.spb.su, www.spiiras.nw.ru; 14 линия, 39, Санкт-Петербург, 199178, РФ; р.т. +7(812)328-8071, факс +7(812)328-0685.

**Shchekotov Maxim Sergeevich** — junior researcher of the laboratory of computer aided integrated systems institution of the St.Petersburg Institute for Informatics and Automation of Russian Academy of Sciences (SPIIRAS). Research interests: knowledge management, smart-spaces, navigation systems. The number of publications — 5. shekotov@iias.spb.su; SPIIRAS, 14th Line V.O., 39, Saint-Petersburg, 199178, Russia; office phone +7(812)328-8071, fax +7(812)328-0685.

**Поддержка исследований.** Работа выполнена при финансовой поддержке РФФИ (проект 13-07-00336-а).

**Acknowledgements.** This research is supported by RFBR (grant # 13-07-00336-a).

## РЕФЕРАТ

### *Щекотов М.С.* **Анализ подходов к позиционированию внутри помещений с использованием трилатерации сигналов Wi-Fi.**

Создание систем для решения проблем позиционирования и навигации внутри зданий является перспективной, актуальной и сложной задачей. Данная проблема требует создания карт на основе поэтажных планов в помещении, выбор эффективной технологии и алгоритмов позиционирования, и развертывания соответствующей инфраструктуры для позиционирования. Современные системы определения местоположения внутри помещений могут предложить не только позиционирование с приемлемой точностью, но и решение задач навигации внутри помещения, обеспечения пользователя контекстной информацией в текстовом, аудио- или видео-формате, сбор статистической информации о плотности и перемещениях людей в зданиях и другое. Таким образом, имеется большой потенциал для создания различных сервисов на базе систем позиционирования внутри помещений.

В статье рассмотрены два основных метода позиционирования внутри помещений с помощью мобильного устройства: метод трилатерации сигналов Wi-Fi и метод радиоотпечатков (fingerprinting). Кроме того, рассмотрены подходы к позиционированию внутри помещений, различающиеся определением расстояния до источника сигнала, которые основываются на вышеуказанных методах и открывают перспективу создания комбинированного метода позиционирования. Данные подходы основаны на модели распространения сигнала в свободном пространстве и методе радиоотпечатков. Было установлено, что модель распространения сигнала в свободном пространстве является не эффективной для использования даже в пределах одной комнаты.

В заключении были сделаны выводы о точности предлагаемого метода, и сделаны предложения по дальнейшей работе в целях увеличения точности метода.

## SUMMARY

### *Shchekotov M.S.* **Indoor Localization Method Based on Wi-Fi Trilateration Technique.**

Creating systems and solutions for solving the problem of indoor positioning and navigation is a complex task. This problem requires of creating maps based on floor plans of indoors, choosing the effective positioning technology and algorithms, and deploying the appropriate positioning devices inside buildings. Modern systems can offer much more than just positioning with acceptable accuracy. For example, the tasks of navigation, providing context information in text-, audio- or video-format, collection statistical information about the movements of people in buildings. The functions of these systems are depending on location type. Thus, the developing companies have great potential for creating various services.

This paper describes two main indoor positioning methods for mobile devices based on Wi-Fi trilateration and fingerprinting. In addition, two indoor positioning approaches based on these methods and having different schemes of distance between transmitter and receiver calculation are considered. These approaches are based on free-space path loss model and received signal strength measurement collection. It was found the signal propagation problem is not effective to positioning within one room.

Conclusion includes ways of further work and increasing positioning accuracy of this method.

О.И. Смоктий

**ВЛИЯНИЕ АТМОСФЕРНОЙ ТУРБУЛЕНТНОСТИ, МОЛЕКУЛЯРНО-АЭРОЗОЛЬНОГО РАССЕЯНИЯ И ПАРАМЕТРОВ БОРТОВОЙ ОПТИЧЕСКОЙ АППАРАТУРЫ НА КАЧЕСТВО СПУТНИКОВОЙ ВИДЕОИНФОРМАЦИИ ПРИ СЪЕМКЕ ЗЕМЛИ ИЗ КОСМОСА**

*Смоктий О.И.* Влияние атмосферной турбулентности, молекулярно-аэрозольного рассеяния и параметров бортовой оптической аппаратуры на качество спутниковой видеоинформации при съемке Земли из космоса.

**Аннотация.** Рассмотрена актуальная проблема совместного влияния развитой атмосферной турбулентности, рассеивающей молекулярно-аэрозольной атмосферы и параметров бортовой оптической аппаратуры на качество спутниковой многоспектральной видеоинформации о природной среде, получаемой при дистанционном зондировании Земли из космоса в видимой области спектра  $\lambda=400-850\text{нм}$ . Представленные результаты получены в рамках теории динамических систем и Фурье-преобразований линейных оптических сигналов. При анализе указанной проблемы использованы репрезентативные данные о вертикальной зависимости структурной функции атмосферной турбулентности, характеристиках бортовой оптической аппаратуры и параметрах пространственно-частотной фильтрации космической видеоинформации, определяемых многократным некогерентным рассеянием солнечных фотонов в подсистеме «молекулярно-аэрозольная атмосфера Земли – горизонтально-неоднородная подстилающая поверхность». Рассчитаны соответствующие передаточные функции и частотно-контрастные характеристики общего комплексного оптического канала. Показано, что в этом случае качество космических изображений и спектров яркости наземных (надводных) объектов среднего ( $\Delta\sim 10-10^2\text{м}$ ) и низкого ( $\Delta\sim 10^3-10^6\text{м}$ ) пространственного разрешения практически не зависит от атмосферной турбулентности. При этом ее влияние по сравнению с многократным некогерентным молекулярно-аэрозольным рассеянием и бортовым оптическим регистратором наиболее существенно только для наиболее высокочастотных и мелко-масштабных фрагментов космических изображений ( $\Delta\ll 10\text{м}$ ).

**Ключевые слова:** функция рассеяния, атмосферная турбулентность, молекулярно-аэрозольное рассеяние, бортовой оптический прибор, пространственно-частотное разрешение, линейная оптическая система, оптическая передаточная функция, Фурье-преобразование, структурная функция турбулентности, флуктуации показатели преломления, атмосферный пространственно-частотный фильтр, наземная мирра, пороговая пространственная частота, прямая и диффузная прозрачности атмосферы.

*Smokty O.I.* Influence of turbulent atmosphere, molecular-aerosol scattering and on-board optical devices' parameters on the satellite video data quality at the Earth's remote sensing from space.

**Abstract:** Topic issue of conjoint influence of developed atmospheric turbulence, molecular-aerosol light scattering and on-board optical devices parameters on the multispectral satellite environmental video data's quality obtained at the Earth's remote sensing from space in the visual spectral field  $\lambda = 400-850\text{nm}$  is considered. Given results are obtained in the framework of both dynamic systems theory and Fourier transformations of the linear optical signals. While analyzing above-considered problem, the representative data concerning vertical dependence of atmospheric turbulence structural function, on-board optical devices characteristics and space video data spatial-frequency filtration parameters determined by multiple non-coherent scatter-

ing of solar photons in the subsystem “the Earth’s molecular-aerosol atmosphere – horizontal and non-uniform underlying surface” have been made used. Both proper transfer functions and frequency-contrast characteristics of general complex optical channel are calculated. For this particular case, is demonstrated that the quality of the both space images and ground-based/water-based objects brightness spectra of medium spatial resolution ( $\Delta \sim 10 - 10^2$ m) and low spatial resolution ( $\Delta \sim 10^3 - 10^4$ m), basically, does not depend on the atmospheric turbulence. At that, the above-mentioned turbulence influence emerges, essentially, for the mostly high-frequency and low-scale space images fragments ( $\Delta \ll 10$ m) only in comparison with the results obtained from both multiple non-coherent molecular-aerosol scattering and on-board optical recording device.

**Keywords:** scattering function, atmospheric turbulence, molecular-aerosol scattering, on-board optical device, spatial-frequency permission, linear optical system, optical transfer function, Fourier transformation, structural turbulence function, refraction index fluctuation, atmospheric spatial-frequency filter, landed myrra, threshold spatial frequency, direct and diffuse atmospheric transparency.

---

**1. Введение.** Проблема пространственно-частотной атмосферной коррекции спутниковой многоспектральной видеоинформации, получаемой при дистанционном зондировании Земли из космоса, имеет принципиально важное значение для решения таких прикладных задач, как селекция, распознавание и количественная оценка параметров состояния наземных (надводных) объектов природного и техногенного происхождения [1-3]. Трудности исследования этой проблемы обусловлены необходимостью предварительного решения ряда фундаментальных физических и математических проблем, связанных, в первую очередь, с проведением численного имитационного моделирования совместных, но независимых процессов атмосферной турбулентности и многократного некогерентного рассеяния солнечных фотонов в системе «атмосфера-объект-фон» при наличии на борту космического носителя регистрирующей оптической аппаратуры [4-6]. Дополнительная сложность решения указанной проблемы состоит в том, что учет влияния турбулентной рассеивающей молекулярно-аэрозольной атмосферы, расположенной между спектрофотометрируемым наземным (надводным) объектом и бортовым регистрирующим прибором, определяется сложными многопараметрическими процессами, имеющими различную физическую природу и пространственно-временные масштабы, несопоставимые характеристики и базовые параметры состояния [7]. Естественно, что для адекватного описания этих независимых процессов даже в рамках единого формализма теории линейных динамических систем и преобразований оптических сигналов необходим достаточно сложный математический аппарат, основанный на использовании различных приближенных методов и алгоритмов [8-10].

В условиях развитой атмосферной турбулентности распространение электромагнитных волн в случайно-неоднородной турбулентной среде определяется когерентным многократным рассеянием солнечных фотонов на пространственно-временных флуктуациях ее показателя преломления, который зависит от температуры, давления и упругости водяных паров атмосферного воздуха [11]. С другой стороны, влияние молекулярно-аэрозольной компоненты земной атмосферы на качество спектральной спутниковой видеоинформации о природной среде обусловлено, в основном, некогерентным многократным рассеянием на флуктуациях плотности молекулярной составляющей атмосферного воздуха (рэлеевское рассеяние) и независимо на взвешенных полидисперсных неоднородных частицах естественного и антропогенного происхождения (аэрозольное рассеяние) [12, 13].

С математической точки зрения учет искажающего радиационного влияния атмосферной турбулентности на качество многоспектральной космической съемки Земли сводится к численному решению уравнений классической электродинамики (уравнений Максвелла) в том или ином приближении для сред со случайным распределением флуктуаций ее показателя преломления [14]. При этом статистика и вероятностные распределения указанных флуктуаций факторизуются в таком базовом имитационном параметре, каким является структурная функция турбулентности [15].

В свою очередь, некогерентное многократное рассеяние солнечных фотонов на молекулах и аэрозольных полидисперсных частицах при определенных и разумных физических предположениях, а также при отсутствии фазовых искажений оптических сигналов может рассматриваться как диффузный перенос излучения (вероятностная цепь Маркова) в атмосфере, ограниченной снизу подстилающей поверхностью [16]. При этом пространственно-угловое и спектральное распределения излучения, многократно рассеянного атмосферой и переотраженного горизонтально-неоднородной подстилающей поверхностью, определяются точными или приближенными решениями краевых задач классической теории переноса солнечных фотонов при соответствующих условиях на верхней и нижней границах атмосферного слоя (см., например, [13]).

Что касается чисто прикладного аспекта данной проблемы, то применительно к рассматриваемой задаче он состоит в необходимости учета и численного моделирования искажающего фильтрующего

влияния рассеивающей и поглощающей турбулентной молекулярно-аэрозольной атмосферы на пространственно-частотное, угловое и спектральное разрешения регистрирующей спутниковой оптической аппаратуры, формирующей в конечном итоге спектрофотометрические образы наземных (надводных) объектов при их съемке из космоса [17]. В этой связи важное значение для их последующей автоматизированной обработки, анализа и тематической интерпретации имеет не только совместный учет атмосферной турбулентности и некогерентного многократного рассеяния на молекулах и аэрозольных частицах, но и численное имитационное моделирование селективного влияния, которое оказывают параметры спектрального многоканального бортового регистрирующего прибора. При этом ухудшение качества космических изображений и спектров яркости природных образований даже при совершенной оптике бортовой регистрирующей аппаратуры, как правило, не связано с дифракционным пределом ее пространственно-углового разрешения [18]. Например, при малой апертуре бортового оптического прибора оно определяется, в основном, случайными перемещениями изображения объекта в фокальной плоскости линзы, где формируется само изображение [19]. Однако в случае большой апертуры и высоком пространственно-угловом разрешении оптической аппаратуры именно развитая атмосферная турбулентность обуславливает появление в видимой области спектра ( $\lambda=400-850$  нм) при коротких экспозициях съемки эффектов расплывание «изображений и характерной пятнистой» структуры в виде «спеклов» их яркостей [20].

Как следствие, проблема построения количественных имитационных моделей, учитывающих влияние всех указанных выше разноплановых факторов на конечные результаты первичной и межотраслевой обработки спутниковой видеoinформации, требует для своего решения комплексного и системного подхода. Ниже используется именно такой подход [21], основанный на едином формализованном описании отдельных составляющих спектрального оптического канала дистанционного зондирования Земли. Такой подход позволяет с помощью понятий оптических передаточных функций и частотно-контрастных характеристик (ЧКХ) комплексной динамической системы построить репрезентативные численные модели пространственно-частотного фильтра турбулентной молекулярно-аэрозольной атмосферы при учете параметров спутниковой оптической аппаратуры.

Схематическое представление комплексной оптической системы, формирующей спутниковые изображения и спектры яркостей

природных и техногенных образований при дистанционном зондировании Земли из космоса, дано на рисунке 1.

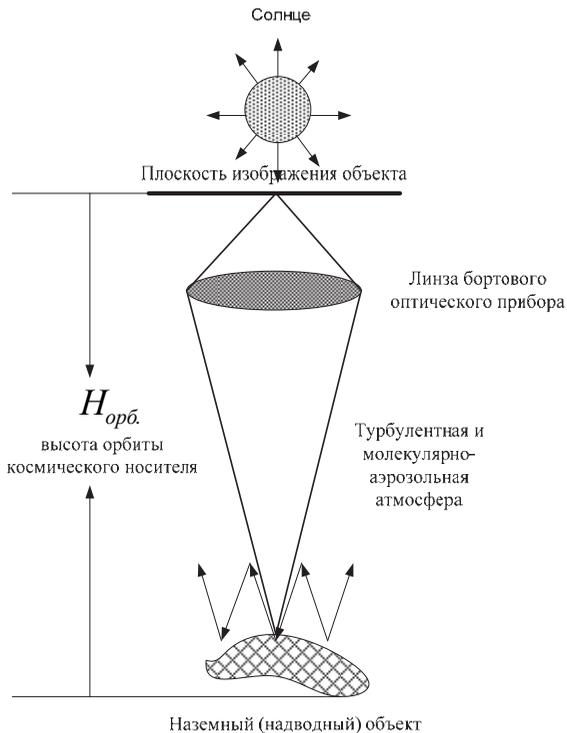


Рис. 1. Схематическое представление комплексной оптической системы, формирующей изображения и спектры яркостей природных объектов при съемке Земли из космоса

Таким образом, прохождение солнечного излучения, отраженного от зондируемого природного или техногенного объекта в направлении бортовой регистрирующей аппаратуры, а затем распространяющегося в обратном направлении к наземному центру приема и обработки спутниковой информации, рассматривается как единый процесс пространственно-частотной фильтрации оптических сигналов не только самим бортовым прибором и турбулентной атмосферой, но и многократно рассеивающей и поглощающей молекулярно-аэрозольной атмосферой, ограниченной снизу произвольным горизонтально-неоднородным отражающим дном.

## 2. Структурные теоретические модели комплексной оптической системы дистанционного зондирования Земли из космоса.

Будем считать, что турбулентная компонента комплексной оптической системы атмосферы, расположенной между спектрофотометрируемым из космоса наземным (надводным) объектом и бортовым регистрирующим прибором, формируется протяженной случайно-неоднородной средой с плавными и достаточно слабыми пространственно-временными флюктуациями ее показателя преломления.

Другая важная компонента этой системы определяется некогерентным многократным рассеянием солнечных фотонов на флюктуациях плотности молекулярной атмосферы (рэлеевское рассеяние) и присутствующих в ней взвешенных полидисперсных частицах в условиях их неплотной упаковки и малого размера по сравнению с длиной волны  $\lambda$  падающего излучения (аэрозольное рассеяние). Кроме того, в обоих случаях необходимо принимать во внимание наличие отражающей подстилающей поверхности, расположенной на уровне нижней границы молекулярно-аэрозольной атмосферы.

Параметры бортовой оптической аппаратуры, формирующей в конечном итоге многоспектральные спутниковые изображения зондируемых объектов, определяют третью базовую компоненту указанной комплексной системы.

Используем далее для построения структурных теоретических моделей линейной динамической системы, определяемой когерентным рассеянием солнечных фотонов турбулентной атмосферой и независимо некогерентным многократным рассеянием молекулярно-аэрозольной атмосферой при учете влияния параметров бортовой оптической аппаратуры, основные положения теории линейных систем и преобразований в оптике [22]. Опуская для простоты дальнейшего рассмотрения угловые и спектральные зависимости моделируемых полей излучения, рассмотрим наземный самосветящийся 2-D объект в виде мирры с пространственным распределением монохроматической яркости  $I(x', y')$  в плоскости, перпендикулярной направлению ее визирования из космоса. Тогда в плоскости изображения  $D(x, y)$  спутникового оптического прибора распределение яркости  $I(x, y)$  данного объекта, наблюдаемого сквозь атмосферу Земли, можно представить в следующем виде:

$$I(x, y) = \iint I_0(x', y') S(x - x', y - y') dx' dy', \quad (1)$$

где  $x = qx'$ ,  $y = qy'$ , а  $q$ -линейное увеличение изображения, обусловленное комплексной оптической системой. Функция  $S(x, y)$  описыва-

ет физические процессы, формирующие 2-D изображение наземного объекта на выходе этой системы при наличии совокупности точечных источников излучения на ее входе. При этом количественное описание процесса передачи пространственных частот  $v_x$  и  $v_y$  исходного изображения объекта можно получить, осуществляя двухмерное преобразование Фурье его спектральных яркостей  $I_0(x, y)$  на входе и  $I(x, y)$  выходе рассматриваемой комплексной линейной оптической системы, а именно:

$$\hat{I}(v_x, v_y) = \iint_{-\infty}^{+\infty} I(x, y) e^{-2\pi i(v_x x + v_y y)} dx dy, \quad (2)$$

$$\hat{I}_0(v_x, v_y) = \iint_{-\infty}^{+\infty} I_0(x, y) e^{-2\pi i(v_x x + v_y y)} dx dy. \quad (3)$$

Выполняя затем 2-D преобразование Фурье функции рассеяния  $S(x, y)$ , получаем основное соотношение, связывающее эту функцию с передаточной функцией комплексной оптической системы  $\hat{P}(v_x, v_y)$ :

$$\hat{P}(v_x, v_y) = \hat{S}(v_x, v_y) = \iint_{-\infty}^{+\infty} S(x, y) e^{-2\pi i(v_x x + v_y y)} dx dy. \quad (4)$$

Учитывая (2)-(4) и используя классическую теорему о свертке функций [22], вместо (1) имеем известное базовое соотношение, связывающее Фурье-преобразования изображений на входе  $\hat{I}_0(x, y)$  и выходе  $\hat{I}(x, y)$  линейной оптической системы:

$$\hat{I}(v_x, v_y) = \hat{P}(v_x, v_y) \hat{I}_0(v_x, v_y). \quad (5)$$

Соотношение (5) в общем виде определяет пространственно-частотную трансформацию космического изображения наземного (надводного) 2-D объекта, искажаемого комплексной спутниковой оптической системой «турбулентная атмосфера - молекулярно-аэрозольная атмосфера – бортовой оптический регистратор». В общем случае передаточная функция  $\hat{P}(v_x, v_y)$  является комплексно-значной и может быть представлена в следующем виде:

$$\hat{P}(v_x, v_y) = |\hat{P}(v_x, v_y)| e^{-i\Phi(v_x, v_y)}. \quad (6)$$

Функция  $|\hat{P}(v_x, v_y)|$ , равная модулю оптической передаточной функции  $\hat{P}(v_x, v_y)$ , является оптической передаточной функцией или частотно-контрастной характеристикой указанной линейной системы, фильтрующей пространственные частоты исходного изображения объекта при съемке Земли из космоса. Фазочастотной характеристикой

такой системы  $\Phi(v_x, v_y)$  можно пренебрегать в случае симметричной функции рассеяния  $\hat{S}(v_x, v_y)$  или при космической съемке объекта в надир  $\{\Phi(v_x, v_y) = 0\}$ , что упрощает рассмотрение проблемы атмосферной фильтрации спутниковой видеоинформации.

Далее, если комплексная линейная оптическая система состоит из  $N$  независимых динамических компонент, каждая из которых характеризуется своей оптической передаточной функцией  $|\Pi_i(v_x, v_y)|$ ,  $i=1, 2, \dots, N$ , то результирующая передаточная функция системы будет равна их произведению:

$$|\Pi_{\Sigma}(v_x, v_y)| = \prod_{i=1}^N |\Pi_i(v_x, v_y)|, \quad i=1, 2, \dots, N. \quad (7)$$

Отметим, что в этом случае разрешающую пространственно-частотную способность  $\Delta$  всей линейной оптической системы оценивают, как правило, либо частотной шириной функции рассеяния  $\Delta[S(v_x, v_y)]$ , либо аналогичной величиной  $\Delta[|\Pi(v_x, v_y)|]$  оптической передаточной функции  $|\Pi(v_x, v_y)|$ . При этом между собой они связаны известным соотношением:

$$\Delta[|\Pi(v_x, v_y)|] \approx \frac{1}{\Delta[S(v_x, v_y)]}, \quad (8)$$

где

$$\Delta^2 [|\Pi(v_x, v_y)|] \approx \iint_{-\infty}^{+\infty} |\Pi(v_x, v_y)| dv_x dv_y. \quad (9)$$

**3. «Теоретические модели линейной динамической системы «рассеивающая турбулентная атмосфера – бортовой оптический прибор».** При рассмотрении репрезентативных теоретических моделей развитой атмосферной турбулентности будем принимать во внимание гипотезу «замороженной турбулентности» (гипотеза Тейлора) (см., например, [23]). Согласно этой гипотезе временные флуктуации метеорологических параметров атмосферы в точке  $(x, y, z)$  обусловлены не изменениями самих параметров, а переносом пространственной картины флуктуаций показателя преломления и его средним потоком. Это базовое предположение позволяет не рассматривать временные флуктуации показателя преломления  $n$  отдельно от его пространственных изменений, поскольку они оказываются связанными между собой. В этом случае временные спектры динамических полей метеорологиче-

ских параметров могут быть выражены через пространственные величины, и наоборот. Таким образом, в этом приближении распространение электромагнитных волн в случайно-неоднородной среде в условиях ее турбулентности факторизуется благодаря флуктуациям показателя преломления среды  $n$ . При этом их статистика определяется статистикой самой турбулентности. Вне полос поглощения величина  $n$  является вещественной и описывается следующей формулой [24]:

$$n = 1 + 77,6(1 + 7,52 \cdot 10^{-3} \lambda^{-2}) \frac{P}{T} 10^{-6}, \quad (10)$$

где длина волны  $\lambda$  задается в мкм, давление атмосферы  $P$  в микробаргах и температура  $T$  в градусах Кельвина. Отметим, что в условиях земной атмосферы флуктуации показателя преломления  $\delta n$  обусловлены, главным образом, флуктуациями температуры и скорости ветра  $V$  в земной атмосфере. Полагая статистику флуктуаций  $\delta n$  пространственно стационарной и изотропной, имеем в инерционном интервале турбулентности ( $l_0, L_0$ ) для структурной функции  $\mathcal{D}_n(r)$  флуктуаций показателя преломления  $n$ , определяющей влияние турбулентности на разрешающую способность спутникового оптического прибора, следующее основное соотношение (закон Колмогорова 2/3) [25]:

$$\mathcal{D}_n(r) = C_n^2(h)r^{2/3}, \quad (11)$$

где  $r = \sqrt{(\Delta x)^2 + (\Delta y)^2}$  - расстояние между точками измерения турбулентного потока,  $C_n^2(h)$  - структурный параметр турбулентности, зависящий от атмосферных метеорологических условий и высоты наблюдения  $h$  над поверхностью Земли. Инерционный интервал развитой турбулентности ( $l_0, L_0$ ) в земной атмосфере соответствует диапазонам ( $10\text{м} \leq L_0 \leq 100\text{м}$ ) и ( $10^{-3}\text{м} \leq l_0 \leq 10^{-2}\text{м}$ ), удовлетворяя условию  $l_0 \ll D \ll L_0$ , где  $D$  - диаметр приемной линзы бортового оптического прибора. Отметим, что согласно теоретическим оценкам вблизи земной поверхности ( $h \sim 2,5\text{м}$ ) величина параметра  $C_n^2(h_0)$  меняется от ( $10^{-13} \text{ м}^{-2/3}$ ) для сильной турбулентности и до ( $10^{-17} \text{ м}^{-2/3}$ ) для слабой турбулентности. Типичным средним значением считается  $C_n^2(h_0) = 10^{-15} \text{ м}^{-2/3}$ . Зависимость структурного параметра турбулентности  $C_n^2(h)$  от высоты  $h$  по данным [26] представлена на рисунке 2.

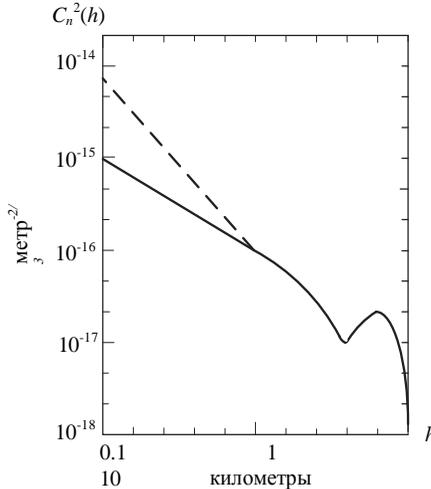


Рис. 2. Структурный параметр турбулентности  $C_n^2(h)$  в зависимости от высоты наблюдения над поверхностью Земли

В дальнейшем для учета влияния атмосферной турбулентности на качество изображений, формируемых бортовым оптическим прибором, и в частности, на его разрешающую  $\Delta$  способность, необходимо выбрать метод решения исходных уравнений распространения электромагнитного излучения в случайно-неоднородной среде. В настоящее время для решения указанной задачи используется метод малых возмущений в его первых приближениях, при которых пренебрегают влиянием многократного рассеяния излучения в среде (приближение Борна) [27]. В методе Борна учитывается возмущение электромагнитного поля за счет рассеянного излучения, обусловленного большим числом независимых рассеивателей в среде. Поэтому в задачах многократного рассеяния электромагнитного излучения в случайно-неоднородных средах при наблюдении наземных (надводных) объектов сквозь турбулентную атмосферу (вне прямой видимости и на длинных трассах) использование метода Борна предпочтительнее по сравнению с другими методами, например, с методом плавных возмущений Рытова [28]. В методе Рытова, который так же, как и метод Борна, широко применяется в теории распространения электромагнитных волн в турбулентной атмосфере, используется не само электромагнитное поле в виде суммы случайных вкладов от независимых рассеивателей, а его логарифмически преобразованные амплитудные флуктуации. Отметим, что согласно экспериментальным данным в задачах многократного рассеяния электромагнитных волн в турбу-

лентной атмосфере, в пределах прямой видимости (однократное рассеяние) приближение Рытова более точно отражает существующие реальности [29].

С учетом вышеизложенного рассмотрим теперь имитационные теоретические модели, позволяющие оценить влияние атмосферной турбулентности на качество изображения, формируемое линейной оптической системой, состоящей из собственно турбулентной атмосферы и бортовой регистрирующей аппаратуры. Согласно [30] оптическая передаточная функция  $|\Pi|$  тонкой сферической линзы в составе бортовой регистрирующей аппаратуры определяется по формуле:

$$\left\{ \begin{array}{l} \left| \Pi \left( \frac{\rho}{D} \right) \right|_{\text{прибор}} = \frac{2}{\pi} \left\{ \arccos \left( \frac{\rho}{D} \right) - \left( \frac{\rho}{D} \right) \left[ 1 - \left( \frac{\rho}{D} \right)^2 \right]^{1/2} \right\}, \left( \frac{\rho}{D} \right) \leq 1, \\ \left| \Pi \left( \frac{\rho}{D} \right) \right|_{\text{прибор}} = 0, \left( \frac{\rho}{D} \right) > 1, \end{array} \right. \quad (12)$$

где параметр  $r = \lambda \sqrt{(\Delta x)^2 + (\Delta y)^2} f$ ,  $f$  - фокусное расстояние линзы,  $D$  - диаметр линзы изображающей оптики.

Если считать время экспозиции космической съемки бесконечно большим ( $t_{\text{экс}} \gg 1$ сек.), то в этом случае основной теоретический результат состоит в заключении [31], что оптическая передаточная функция  $|\Pi_{\Sigma}(v_x, v_y)|$  системы «турбулентная атмосфера-бортовой оптический прибор», нормированная на единицу в начале координат ( $v_x = 0, v_y = 0$ ), равна произведению оптических передаточных функций турбулентной атмосферы  $|\Pi(v_x, v_y)|_{\text{турб.}}$  и оптического прибора  $|\Pi(v_x, v_y)|_{\text{приб.}}$ , именно:

$$|\Pi_{\Sigma}(v_x, v_y)| = |\Pi(v_x, v_y)|_{\text{турб.+приб.}} = |\Pi(v_x, v_y)|_{\text{турб.}} \times |\Pi(v_x, v_y)|_{\text{приб.}}. \quad (13)$$

Оптическая передаточная функция собственно турбулентной атмосферы  $|\Pi(v_x, v_y)|_{\text{турб.}}$  в приближении Рытова равна:

$$|\Pi(v_x, v_y)|_{\text{турб.}} = \exp \left[ -1,46 \frac{k^2}{\cos \theta} \rho^{5/3} \int C_n^2(h) \left( 1 - \frac{h}{H_{\text{орб}}} \right)^{5/3} dh \right], \quad (14)$$

где волновое число  $k = \frac{2\pi}{\lambda}$ ,  $\theta$  - угол наклона трассы распространения

излучения,  $\rho = \frac{\lambda H_{\text{орб}}}{M \cos \theta}$ ,  $M$  - период наземной гармонической само-

светящейся мирры,  $H_{обр}$  – высота расположения бортового оптического прибора.

В случае больших, но конечных времен экспозиции съемки ( $t_{экс} < \infty$ ) не удается разделить искажение изображения за счет турбулентной атмосферой [32] и оптического прибора [33]. Однако, как будет показано ниже на основе проведенного численного моделирования, величины результирующих передаточных функций в обоих случаях практически совпадают в широком диапазоне метеорологических условий земной атмосферы.

Соотношение (14) показывает, что для расчета оптической передаточной функции турбулентной атмосферы  $\left| \Pi(\nu_x, \nu_y) \right|_{\text{турб.}}$  необходима информация о высотной зависимости структурной характеристики флюктуации показателя преломления  $C_n^2(h)$ . Учитывая существенную экспериментальную зависимость этой базовой величины от характера и интенсивности атмосферной турбулентности, в [34] предложено использовать для нее полуэмпирические аппроксимирующие функции для «очень хорошей» и «очень плохой» атмосферной видимости, а также для промежуточных условий в виде среднего геометрического для них. Именно для «очень хороших» условий видимости имеем:

$$C_n^2(h) = 5.19 \cdot 10^{-16} \cdot 10^{-0.86h} + 10^s, \quad (15)$$

где  $s = -18.34 + 0.29h - 2.84 \cdot 10^{-2} h^2 + 7.43 \cdot 10^{-4} h^3$

Для «очень плохих» условий видимости в [31] предложено следующее представление величины  $C_n^2(h)$ :

$$C_n^2(h) = 9.52 \cdot 10^{-14} \cdot 10^{-2.09h} + 10^s, \quad (16)$$

где  $s = -14.39 + 0.17h - 3.48 \cdot 10^{-2} h^2 + 9.59 \cdot 10^{-4} h^3$ .

Соответственно для промежуточных условий атмосферной видимости имеем:

$$C_n^2(h) = \sqrt{C_{n,хор.}^2 \times C_{n,плох.}^2}. \quad (17)$$

В таблице 1 представлены значения оптической передаточной функции рассеивающей турбулентной атмосферы  $\left| \Pi(\nu_x, \nu_y) \right|_{\text{турб.}}$  для трех указанных выше приближенных моделей высотного хода  $C_n^2(h)$  в случае бесконечного ( $t_{экс.} = \infty$ ) и конечного ( $t_{экс.} < \infty$ ) времен экспозиции.

Таблица 1. Значения оптической передаточной функции турбулентной атмосферы  $|\Pi(\nu_x, \nu_y)|_{\text{турб.}}$  в зависимости от метеорологических условий видимости и времени экспозиции космической съемки

хорошие условия видимости				
F \ t <sub>экс.</sub>	0,1047·10 <sup>4</sup>	0,1047·10 <sup>5</sup>	0,1047·10 <sup>6</sup>	0,1047·10 <sup>7</sup>
t <sub>экс.</sub> =∞	1,0000	0,9996	0,9811	0,4566
t <sub>экс.</sub> <∞	1,0000	0,9996	0,9811	0,4566
средние условия видимости				
t <sub>экс.</sub> =∞	0,9990	0,9949	0,7926	0,7194·10 <sup>-4</sup>
t <sub>экс.</sub> <∞	0,9990	0,9949	0,7926	0,7194·10 <sup>-4</sup>
плохие условия видимости				
t <sub>экс.</sub> =∞	0,9982	0,9100	0,1969·10 <sup>-1</sup>	0,9978·10 <sup>-7</sup>
t <sub>экс.</sub> <∞	0,9982	0,9100	0,1969·10 <sup>-1</sup>	0,5228·10 <sup>-8</sup>

Данные таблице 1 показывают, что оптические передаточные функции рассеивающей турбулентной атмосферы  $|\Pi(\nu_x, \nu_y)|_{\text{турб.}}$  практически совпадают для бесконечного (t<sub>экс.</sub>=∞) и конечного времени (t<sub>экс.</sub><∞) экспозиции космической съемки.

**4. Влияние атмосферной турбулентности и параметров бортового оптического прибора на разрешающую способность спутникового канала изображения.** Рассмотрим теперь влияние атмосферной турбулентности и параметров бортового оптического прибора на разрешающую способность рассматриваемой линейной оптической системы. Параметр, который характеризует влияние рассеивающей турбулентной атмосферы на разрешающую способность системы, включающей бортовой оптический прибор, можно определить радиусом атмосферной когерентности  $r_0$  (см., например, [17]):

$$r_0(\Delta) = 0,185\lambda^{6/5} \left[ \int_0^{\mathcal{L}} C_n^2 dh \right]^{-3/5}, \quad (18)$$

где  $\mathcal{L}$  - длина трассы наблюдения.

При коротких временах экспозиции (t<sub>экс.</sub><<0,01сек) и, следовательно, для коротких трасс визирования, когда можно пренебрегать зависимостью величины  $C_n^2(h)$  от высоты  $h$  и не учитывать эффекты многократного молекулярно-аэрозольного рассеяния, из (18) получаем соотношение:

$$r_0(\Delta) = 0,185\lambda^{6/5} \mathcal{L}^{-3/5} (C_n^2)^{-3/5}. \quad (19)$$

При больших временах экспозиции ( $t_{\text{экс.}} \gg 0,01 \text{ сек}$ ) и соответственно в случае длинных трасс наблюдения расчетным для передаточной функции бортового оптического прибора и турбулентной атмосферой является выражение:

$$|\Pi(\Omega)|_{\text{турб+прибор}} = |\Pi(\Omega)|_{\text{прибор}} \exp \left\{ -3,44 \left( \frac{\lambda \Omega}{r_0} \right) \right\}, \quad (20)$$

где  $\Omega = \nu f$ .

В общем случае мерой разрешающей способности  $\Delta$ , рассматриваемой линейной оптической системы со сферически-симметричной оптической передаточной функцией  $|\Pi(\Omega)|_{\text{турб+прибор}}$ , нормированной на единицу в начале координат ( $\nu_x = \nu_y = 0$ ), может служить объем  $\mathcal{R}(\Delta)$  под ее поверхностью:

$$\mathcal{R}(\Delta) = 2\pi \int_0^\infty \Omega |\Pi(\Omega)|_{\text{турб+прибор}} d\Omega. \quad (21)$$

Величину  $\mathcal{R}(\Delta)$  можно оценить численно (см., например, [17]). Соответствующая зависимость приведена на рисунке 3. При  $\left(\frac{d_0}{r_0}\right) = 1$  пространственно-частотное разрешение оптической системы растет с увеличением размера апертуры  $d_0$  как квадрат отношения  $\left(\frac{d_0}{r_0}\right)$ . При  $\left(\frac{d_0}{r_0}\right) \approx 1$  рост разрешающей способности замедляется настолько, что дальнейшее увеличение параметра  $d_0$  практически не влияет на величину  $\mathcal{R}(\Delta)$ . При  $\left(\frac{d_0}{r_0}\right) \ll 1$  величина  $\mathcal{R}(\Delta)$  стремится к постоянному максимальному значению  $\mathcal{R}_{\text{max}}(\Delta)$ .

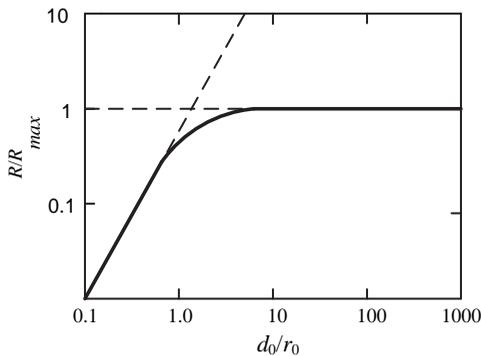


Рис.3. Нормализованная разрешающая способность  $\left(\mathcal{R}/\mathcal{R}_{\text{max}}\right)$  оптической системы в зависимости от приведенного диаметра изображающей оптики  $\left(\frac{d_0}{r_0}\right)$  при длинных экспозициях ( $t_{\text{экс}} \gg 0,01 \text{ сек.}$ )

Пересечение двух асимптот на рисунке 3 происходит при выполнении равенства  $d_0=r_0$ , чем и объясняется введение параметра  $r_0$  согласно формуле (19). Значения  $r_0$  для условий наблюдения наземного (надводного) объекта через незамутненную турбулентную атмосферу в надир ( $\theta=0$ ) располагаются в пределах  $r_0 \sim 5-20$  см в зависимости от степени ее загрязнения и метеорологических условий. Для наклонных трасс и при углах наблюдения  $\theta$ , отличных от 0, величина параметра  $r_0$  может быть много меньше. Отметим, что максимальную меру разрешающей способности  $\mathcal{R}_{max}(\Delta)$  можно асимптотически оценить, заменяя при  $\left(\frac{d_0}{r_0}\right) \rightarrow \infty$  входящей в (21) интеграл его известным табличным аналогом (см. [17]):

$$\frac{\pi}{2} \int_0^{\infty} x \exp[-(ax)^{5/3}] dx, \quad (22)$$

где параметр  $a = (3,44)^{3/5} \left(\frac{d_0}{r_0}\right)$ .

В результате такой замены вместо (21) получаем для  $\mathcal{R}_{max}(\Delta)$  приближенное асимптотическое соотношение:

$$\mathcal{R}_{max}(\Delta) \approx \left(\frac{r_0}{\lambda}\right)^2. \quad (23)$$

Выражение (23) для  $\mathcal{R}_{max}(\Delta)$  соответствует плоскому углу разрешения  $\left(\frac{\lambda}{r_0}\right)$  оптического прибора или адекватно пространственному линейному размеру  $\left(\frac{\lambda}{r_0}\right) \mathcal{L}$ .

Отметим теперь принципиальные отличия в работе бортовой оптической системы, формирующей изображение в условиях турбулентной атмосферы, при переходе от длинных ( $t_{экс} \gg 10^{-2}$  сек.) к коротким экспозициям съемки ( $t_{экс} \ll 10^{-2}$  сек.). В последнем случае появляется возможность пренебречь влиянием на качество изображения тех искажений волнового фронта, которые сводится к изменению среднего угла наклона трассы наблюдения. При этом наклон волнового фронта падающей волны при коротких экспозициях приводит лишь к смещению центра изображения, но не меняет его пространственной структуры. Что касается длинных экспозиций, то флюктуация угла наклона трассы наблюдения приводит к расширению усредненной функции рассеяния  $|S(\nu_x, \nu_y)|$  и, следовательно, к сужению частотной ширины оптической передаточной функции рассматриваемой системы  $|\Pi(\nu_x, \nu_y)|$ , (см. рисунок 4).

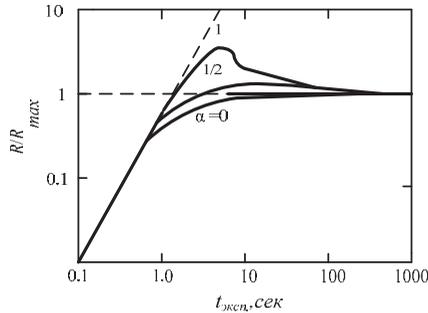


Рис. 4. Разрешающая способность оптической системы при длинных и коротких экспозициях

На рисунке 4 представлены численные значения нормализованной разрешающей способности ( $R / R_{max}$ ) оптической системы в зависимости от времени экспозиции  $t_{эксп}$ . съемки:  $\lambda=500\text{мм}$ , диаметр атмосферной когерентности  $r_0=10\text{см}$ ,  $\alpha=0$  – длинные экспозиции съемки,  $\alpha=1$  – короткие экспозиции съемки,  $\alpha=1/2$  – промежуточный случай.

**5. Влияние многократного молекулярно-аэрозольного рассеяния на пространственно-частотную фильтрацию космических изображений природных образований.** Рассмотрим теперь в рамках теории линейных оптических систем и преобразований влияние многократного молекулярно-аэрозольного рассеяния на пространственно-частотную фильтрацию спутниковых изображений природных образований. На рисунке 5 представлена приближенная аналитическая модель пространственно-частотного фильтра земной атмосферы при учете многократного рассеяния фотонов в молекулярно-аэрозольной земной атмосфере [35].

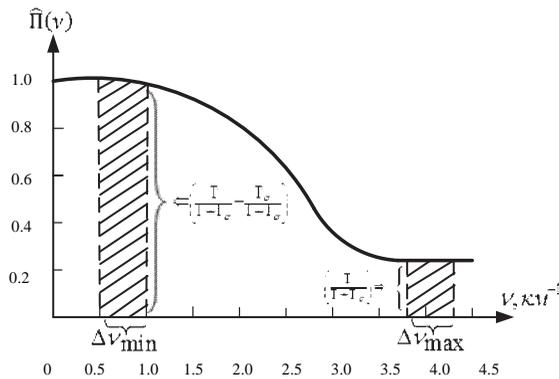


Рис. 5. Приближенная аналитическая модель пространственно-частотного фильтра аэрозольно-молекулярной земной атмосферы в видимой области спектра  $\lambda= (400-800)\text{нм}$

Из данной модели следует, что при многоспектральной съемке Земли из космоса молекулярно-аэрозольной рассеивающая и поглощающая компонента атмосферы является полосовым пространственно-частотным фильтром, пропускающим в основном низкочастотные и подавляющим высокочастотные (мелкомасштабные) фрагменты спутниковых изображений и спектров яркости наземных (надводных) природных образований. В рамках использованной концепции комплексных линейных оптических систем расчетные данные для оптической передаточной функции  $|\Pi(v_x, v_y)|_{\text{аэр}}$  молекулярно-аэрозольной атмосферы, представленные на рисунке 5, получены на основе линейности и инвариантности решений соответствующих краевых задач теории переноса излучения [2]. При этом для получения приближенных аналитических выражений передаточной функции  $|\Pi(v_x, v_y)|_{\text{аэр}}$  использовалась полиномиальная (квадратичная) аппроксимация атмосферного частотного фильтра и точные численно-аналитические решения соответствующих краевых задач D-1 теории переноса излучения.

Согласно [35] в отсутствие фазочастотной зависимости  $[\Phi(v) = 0]$  модуль передаточной функции  $\hat{\Pi}_{\text{аэр}}(v)$  рассеивающей и поглощающей молекулярно-аэрозольной земной атмосферы, нормированной при  $v = 0$ , определяется согласно выражению:

$$\hat{\Pi}(v) = \frac{|\Pi_{\text{аэр}}(v)|}{|\Pi_{\text{аэр}}(0)|} = \delta_*(v) + [1 - \delta_*(v)] \frac{T(\theta, r_0)}{T_\sigma(\theta, r_0) + T(\theta, r_0)}, \quad (24)$$

где модельный полуэмпирический аппроксимирующий параметр  $\delta_*(v)$  рассчитывается следующим образом:

$$\delta_*(v) = - \frac{(v - v_{*,\min})}{(v_{*,\max} - v_{*,\min})}. \quad (25)$$

Параметр  $\tau_0$  - полная оптическая толщина молекулярно-аэрозольной земной атмосферы в вертикальном направлении  $\theta=0^\circ$ . Экстремальные пороговые частоты в (25), за пределами которых атмосферный полосовой фильтр превращается в полосовой модулятор, равны  $v_{*,\min} = (0,5 - 1)\text{км}^{-1}$  и  $v_{*,\max} = 3,5\text{км}^{-1}$ . Функция  $T(\theta, \tau_0) = \exp\{-\tau_0 \cos \theta\}$  является прямой прозрачностью земной атмосферы, обусловленной однократным рассеянием излучения, а  $T_\sigma(\theta, \tau_0)$  равна диффузной прозрачности атмосферы, определяемой многократным рассеянием излучения в направлении текущего угла визирования  $\theta$ .

Таким образом, пространственно-частотный фильтр многократно рассеивающей и поглощающей молекулярно-аэрозольной земной

атмосферы определяется согласно (24) двумя пороговыми пространственными частотами  $\nu_{*,min} = (0,5 - 1)\text{км}^{-1}$  и  $\nu_{*,max} = 3,5\text{км}^{-1}$ , а также соотношением между ее прямой (Т) и диффузной (Т<sub>д</sub>) прозрачностью. При этом за пределами указанных пороговых частот  $\nu \leq \nu_{*,min}$  и  $\nu \geq \nu_{*,max}$  атмосферный пространственно-частотный фильтр ведет себя как полосовой модулятор, равномерно искажающий по всем частотам  $\nu$  фрагменты спутниковых изображений и спектров яркости спектрофотометрируемых объектов (см. таблицу 2).

Таблица 2. Зависимость оптической передаточной функции молекулярно-аэрозольной атмосферы от пространственной частоты  $\nu$  (замкнутая аэрозольная модель Элтермана [36], съемка в надир  $\theta=0^\circ$ ,  $\nu_{*,min} = (0,5 - 1)\text{км}^{-1}$  и  $\nu_{*,max} = 3,5\text{км}^{-1}$

$\nu, \text{км}^{-1}$	$ H(\nu) _{aэp.}$		$\nu, \text{км}^{-1}$	$ H(\nu) _{aэp.}$	
	$a_0=0,2;$ $a_1=0,2$	$a_0=0,6;$ $a_1=0,3$		$a_0=0,2;$ $a_1=0,2$	$a_0=0,6;$ $a_1=0,3$
$10^{-1}$	2,110	2,739	5	1,070	2,050
1	1,615	2,410	10	1,070	2,050
3	1,080	2,053	$10^2$	1,070	2,050
4	1,075	2,050	$10^4$	1,070	2,050

В таблице 2 представлены значения нормированной оптической передаточной функции  $|H(\nu)|_{aэp.}$  рассеивающей и поглощающей молекулярно-аэрозольной земной атмосферы (модель Элтермана) [36] в зависимости от пространственной частоты  $\nu$  наземной миры (26), наблюдаемой из космоса.

**6. Численные модели пространственно-частотной фильтрации комплексным оптическим каналом дистанционного зондирования Земли из космоса.** Проведенный выше анализ позволяет провести численное имитационное моделирование всей комплексной оптической системы, состоящей из следующих независимых компонент:

- рассеивающая и поглощающая турбулентная земная атмосфера;
- рассеивающая и поглощающая молекулярно-аэрозольная земная атмосфера в условиях ее турбулентного состояния;
- регистрирующий бортовой прибор, формирующий много-спектральные изображения наземных (надводных) объектов при съемке Земли из космоса.

В этой связи отметим, что если известна оптическая передаточная функция  $|P_{\Sigma}(\nu_x, \nu_y)|$  комплексной системы «атмосферная турбулентность – молекулярно-аэрозольное рассеяние – бортовой оптический прибор», то общее трансформирующее влияние этой системы на

качество спутниковой видеoinформации при съемке Земли из космоса можно также адекватно оценить с помощью ее частотно-контрастной характеристики (ЧКХ). Учитывая, что яркость любого наземного (надводного) протяженного ортотропного объекта, наблюдаемого из космоса, представима в виде отрезка ряда Фурье, рассмотрим наземную плоскую мирру, альбеда  $A(x, y)$  которой меняется вдоль одной из декартовых координат  $(x, y)$  по гармоническому закону, а вдоль другой постоянно, именно:

$$A(x, y) = a_0 + a_1 \cos\left(\frac{2\pi}{M}, x\right), \quad (26)$$

где  $M \in [0, \infty)$  - период мирры.

Освещенность такой мирры в плоскости изображения бортового оптического прибора  $I(x, y)$  связана с исходной яркостью объекта, определяемой величиной альбеда  $A(x', y')$ , следующим соотношением:

$$I(x, y) = \iint S(x - x', y - y') A(x', y') dx' dy', \quad (27)$$

где

$$A(x', y') = a_0 + a_1 \cos\left(\frac{2\pi}{m}, x'\right), \quad m = qM. \quad (28)$$

Параметр  $q$  равен коэффициенту линейного увеличения линзы бортового оптического прибора:

$$q = \frac{f}{f + H_{\text{орб}} / \cos\theta}, \quad \text{при } H_{\text{орб}} \gg f, \quad (29)$$

где  $f$  - фокусное расстояние приемной линзы,  $H_{\text{орб}}$  - высота орбиты космической станции,  $\theta$  - текущий угол визирования наземного (надводного) объекта.

Проведя пространственное статическое усреднение величины  $I(x, y)$ , получаем:

$$\begin{aligned} \langle I(x, y) \rangle &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} A(x', y') S(x - x', y - y') dx' dy' = \\ &= a_0 + a_1 \cos\left(\frac{2\pi}{m}, x\right) \hat{S}\left(\frac{2\pi}{m}, \theta\right). \end{aligned} \quad (30)$$

Учитывая соотношение (4), имеем:

$$\hat{\Pi}\left(\frac{2\pi}{m}, 0\right) = \hat{S}\left(\frac{2\pi}{m}, 0\right). \quad (31)$$

С другой стороны, размытие изображения наземной мирры (26) в плоскости изображения бортового оптического прибора можно харак-

теризовать контрастом этого изображения на уровне орбиты космического носителя:

$$K_{\infty} = \frac{\langle I \rangle_{\max} - \langle I \rangle_{\min}}{\langle I \rangle_{\max} + \langle I \rangle_{\min}}, \quad (32)$$

где  $\langle I \rangle_{\max}$  и  $\langle I \rangle_{\min}$  соответственно максимальные и минимальные значения средней яркости наземного (надводного) объекта в плоскости изображения бортового оптического прибора.

В рассматриваемом случае видимый контраст  $K_{\infty}$  изображения мирры (26) с учетом соотношения (30) определяется следующим выражением:

$$K_{\infty} \left( \frac{2\pi}{m}, 0 \right) = \frac{a_1}{a_2} \hat{S} \left( \frac{2\pi}{m}, 0 \right) = \frac{a_1}{a_2} \hat{\Pi} \left( \frac{2\pi}{m}, 0 \right). \quad (33)$$

Таким образом, контраст космического изображения  $K_{\infty}$  наземной гармонической мирры (26), с точностью до постоянного множителя  $\left( \frac{a_0}{a_1} \right)$  совпадает с нормированной оптической передаточной функцией оптической системы  $\hat{\Pi} \left( \frac{2\pi}{m}, 0 \right)$ . Принимая во внимание, что величина  $\left( \frac{a_0}{a_1} \right)$  равна контрасту  $K_0$  наземной мирры (26) на уровне подстилающей поверхности, для частотно-контрастной характеристики (ЧКХ) линейной оптической системы соответствующее равенство эквивалентности имеет следующий вид:

$$P(v_x, v_y) = \left| \frac{K_{\infty}}{K_0} \right| = \left| \frac{\Pi_{\Sigma}(v_x, v_y)}{\Pi_{\Sigma}(0,0)} \right|. \quad (34)$$

Учитывая формулу (34), для расчета ЧКХ рассматриваемой комплексной линейной оптической системы при отсутствии фазочастотной зависимости [ $\Phi(v_x, v_y) = 0$ ] необходимо использовать представленные выше нормированные выражения для  $\left| \Pi(v_x, v_y) \right|$  в случае турбулентной атмосферы  $\hat{\Pi}_{\text{турб}}(v_x, v_y) = \left| \frac{\Pi_{\text{турб}}(v_x, v_y)}{\Pi_{\text{турб}}(0,0)} \right|$ , для бортового оптического прибора  $\hat{\Pi}_{\text{приб}}(v_x, v_y) = \left| \frac{\Pi_{\text{приб}}(v_x, v_y)}{\Pi_{\text{приб}}(0,0)} \right|$  и молекулярно-аэрозольной атмосферы  $\hat{\Pi}_{\text{атм}}(v_x, v_y) = \left| \frac{\Pi_{\text{атм}}(v_x, v_y)}{\Pi_{\text{атм}}(0,0)} \right|$ :

$$\hat{\Pi}_{\Sigma}(v_x, v_y) = \hat{\Pi}_{\text{турб}}(v_x, v_y) \times \hat{\Pi}_{\text{приб}}(v_x, v_y) \times \hat{\Pi}_{\text{атм}}(v_x, v_y). \quad (35)$$

**7. Результаты численного моделирования оптического пространственно-частотного атмосферного фильтра при съемке Земли из космоса.** Численное моделирование комплексного пространственно-частотного атмосферного фильтра (35) было проведено для соответствующих оптико-физических параметров каждой из его компонент и использовании соотношений (15)-(17) применительно к условиям космической съемки Земли в видимой области спектра  $\lambda = (400-800)\text{нм}$ . Для численного моделирования передаточной функции турбулентной атмосферы  $\hat{\Pi}_{\text{турб}}(v_x, v_y)$  прежде всего необходима информация о высотном профиле ее структурной функции  $C_n^2(h)$ . Как отмечено выше, обработка экспериментальных данных высотной зависимости структурной характеристики флуктуаций показателя преломлений  $C_n^2(h)$ , проведенная в [34], позволила построить полуэмпирические модельные структурные функции  $C_n^2(h)$  в виде соотношений (15)-(17) для четырех метеорологических состояний земной атмосферы, соответствующих «хорошим», «плохим» и «средним» условиям ее видимости. Выше на указанной основе были рассчитаны численные модели нормированных оптических передаточных функций турбулентной атмосферы  $\left| \frac{\Pi_{\text{турб}}(v_x, v_y)}{\Pi_{\text{турб}}(0, 0)} \right| = \hat{\Pi}_{\text{турб}}(v_x, v_y)$  для бесконечного ( $t_{\text{экс.}} = \infty$ ) и конечного ( $t_{\text{экс.}} < \infty$ ) времен экспозиции в зависимости от волнового числа  $\sqrt{v_x^2 + v_y^2} F$ , где  $F = f/D$  (таблица 1).

Входными параметрами проведенного численного моделирования являются: высота орбиты космической станции  $H_{\text{орб}}(m)$ , период наземной (надводной) гармонической мирры  $M(m)$ , фокусное расстояние линзы бортового оптического прибора  $f$  (м), диаметр  $D$  линзы (м), длина волны излучения  $\lambda$  (нм), угол наклона трассы визирования  $\theta$  (рад), время экспозиции  $t$  (сек), скорость движения космической станции  $V$ (м/сек),  $\beta$ - угол обзора съемки (град).

Как отмечено выше, величины  $|\Pi_{\text{турб}}(v)|$  в случае бесконечного ( $t_{\text{экс.}} = \infty$ ) и конечного ( $t_{\text{экс.}} < \infty$ ) времен экспозиции практически совпадают и, следовательно, одинаково влияют на пространственное разрешение бортовой оптической системы  $\left( \frac{\lambda H_{\text{орб}}}{M \cos \theta} \right)$ . Из анализа проведенного численного имитационного моделирования следует, что разрешающая способность  $\Delta$  бортовой оптической аппаратуры, опреде-

ляемая предельным периодом  $M$  наземной мирры (26), при параметрах наблюдений с космической станции ( $H_{орб} \sim 100$  км,  $\lambda = 0,55 \cdot 10^{-6}$  м,  $v = 10^4$  км/сек,  $\theta = \theta^0$ ,  $D = 1$  м,  $f = 1$  м,  $T = 2 \cdot 10^{-6}$  сек,  $\beta = 2^0$ ), составляет при очень хороших условиях видимости  $\Delta \sim 1$  м и для плохой видимости  $\Delta \sim 10$  м.

Далее, как следует из табл.1, для заданных параметрах расчетов значения  $\left| \frac{I_{мирр, \infty}(v)}{I_{мирр, \infty}(0)} \right|$  и  $\left| \frac{I_{мирр, t}(v)}{I_{мирр, t}(0)} \right|$  практически совпадают. Это означает, что скорость космической станции  $V$  и время экспозиции  $t_{эксн.}$  съемки наземной мирры (26) практически не влияют на пространственное разрешение спутникового изображения системой «турбулентная атмосфера - бортовой оптический прибор». Кроме того, проведенное численное моделирование показало, что значения оптической передаточной функции  $|P(v)|_{турб.}$  в пределах погрешности ( $10^{-5}$ ) практически не зависят от высоты орбиты  $H_{орб}$  космической станции при ее изменении в пределах от 50 км до 250 км.

Далее, из данных таблицы 2 следует, что для указанных выше параметров  $|P(v)|_{аэр.}$  предельный период пространственного разрешения наземной (надводной) гармонической мирры (26), наблюдаемой из космоса сквозь молекулярно-аэрозольную атмосферу, составляет  $\approx 200-300$  м. Таким образом, полученные количественные оценки позволяют сделать вывод, что основной вклад в атмосферное искажение космических снимков и спектров яркости природных образований большого пространственного масштаба обусловлен не флуктуациями показателя преломления атмосферы, и, следовательно, не ее турбулентностью, а наличием в ней рассеивающих аэрозольных частиц и молекул. Следовательно, при проведении процедуры атмосферной коррекции спутниковых изображений и спектров яркостей природных и техногенных объектов среднего и низкого пространственного масштаба эффектом турбулентности можно пренебречь. Подобный эффект следует учитывать только для атмосферной коррекции космических снимков очень высокого пространственного разрешения ( $\Delta \ll 1$  м).

Для рассмотренных выше теоретических моделей оптических передаточных функций чисто турбулентной  $|P(v, \theta)|_{турб.}$  и молекулярно-аэрозольной атмосферы  $|P(v, \theta)|_{аэр.}$ , включая бортовой оптический прибор  $|P(v, \theta)|_{прибор}$ , было проведено комплексное численное моделирование общей нормированной оптической передаточной функции

$\left| \frac{\Pi_{\Sigma}(\nu, \theta)}{\Pi_{\Sigma}(0, \theta)} \right| = \hat{\Pi}_{\Sigma}(\nu, \theta)$  в зависимости от указанных выше параметров, а

также от пространственной частоты  $\nu = 1/M$  и длины волны  $\lambda$ .

Результаты проведенного численного имитационного моделирования показали, что в отдельности оптические нормированные передаточные функции турбулентной атмосферы  $|\Pi(\nu, \theta)|_{\text{турб}}$ , молекулярно-аэрозольного рассеяния  $|\Pi(\nu, \theta)|_{\text{аэр}}$  и бортового оптического прибора  $|\Pi(\nu, \theta)|_{\text{приб}}$  для различных метеорологических условий видимости существенно отличаются друг от друга. Особенно это заметно на больших пространственных частотах  $\nu$ . Спектральная зависимость передаточных функций турбулентной атмосферы и космического прибора показывает уменьшение величины  $|\Pi(\nu, \theta)|_{\text{приб}}$  и относительно слабое увеличение величины  $|\Pi(\nu, \theta)|_{\text{турб}}$  при увеличении длины волны  $\lambda$ . Однако общая оптическая передаточная функция системы  $|\Pi(\nu, \theta)|_{\text{турб}} \times |\Pi(\nu, \theta)|_{\text{аэр}}$  уменьшается с увеличением длины волны  $\lambda$ . Учет многократного рассеяния и поглощения излучения в молекулярно-аэрозольной атмосфере приводит к увеличению общей оптической передаточной функции  $|\Pi_{\Sigma}|$  при  $\Delta > 10\text{ м}$  и при увеличении длины волны  $\lambda$ . Далее при выбранных параметрах численного моделирования большее влияние на качество спутникового изображения наземной гармонической мишени (26) оказывают искажения, обусловленные бортовым оптическим прибором, а не турбулентной атмосферой. Аналогичный вывод справедлив и для оптических передаточных функций многократно рассеивающей молекулярно-аэрозольной атмосферы  $|\Pi(\nu, \theta)|_{\text{аэр}}$  по сравнению с турбулентной атмосферой  $|\Pi(\nu, \theta)|_{\text{турб}}$ . Именно величины  $|\Pi(\nu, \theta)|_{\text{аэр}} \times |\Pi(\nu, \theta)|_{\text{турб}}$  в видимой ( $\lambda=400\text{-}800\text{ нм}$ ) и ближней ИК ( $\lambda=800\text{-}1500\text{ нм}$ ) областях спектра показывают, что чисто рассеивающая и поглощающая молекулярно-аэрозольная атмосфера в отличие от турбулентной атмосферы подавляет все высокие частоты, пропуская только низкочастотные детали спутниковых изображений и спектров яркостей природных наземных (надводных) образований.

В качестве примеров проведенного численного моделирования на рисунке 6 представлены зависимости оптических передаточных функций молекулярно-аэрозольной атмосферы в условиях ее турбу-

лентного состояния от пространственной частоты  $\nu$  наземной гармонической мурры (26), наблюдаемой из космоса:  $\lambda=500$  нм,  $F = 0,3M$ ,  $\beta = 2^0$ ,  $\theta = 0^0$ .

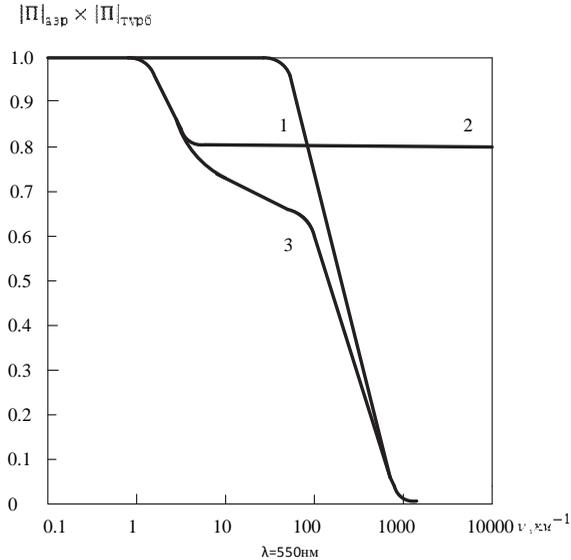


Рис.6. Частотная зависимость нормированных передаточных функций молекулярно-аэрозольной атмосферы в условиях турбулентности при «плохих» условиях видимости

Кривая 1 представляет частотную зависимость оптической передаточной функции  $|\Pi(\nu, \theta)|_{\text{турб}}$  чисто турбулентной атмосферы в случае «плохих условий видимости». Кривая 2 демонстрирует частотную зависимость оптической передаточной функции многократно рассеивающей молекулярно-аэрозольной атмосферы  $|\Pi(\nu, \theta)|_{\text{аэр}}$ , рассчитанную по формуле (24) для замкнутой модели Элтермана [2] в случае аппроксимирующей трехчленной индикатрисы рассеяния. Кривая 3 показывает зависимость оптической передаточной функции системы «атмосферная турбулентность – молекулярно-аэрозольное рассеяние» от частоты  $\nu$  для «плохих» условий видимости и многократного рассеяния в случае аппроксимирующей трехчленной атмосферной индикатрисы рассеяния [2].

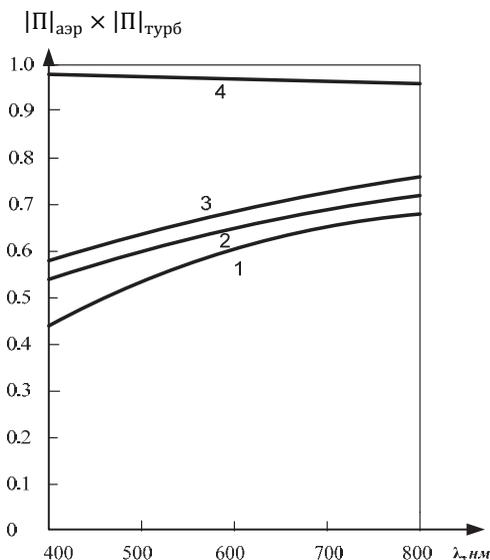


Рис. 7. Спектральная зависимость нормированных передаточных функций турбулентной атмосферы при наличии молекулярно-аэрозольной компоненты

На рисунке 7 представлены зависимости оптических передаточных функций системы «турбулентная атмосфера – молекулярно-аэрозольная атмосфера» от длины волны  $\lambda$  в видимой области спектра (400-800) нм: высота орбиты космического носителя  $H_{орб.} = 100 \text{ км}$ , фокусное расстояние бортового оптического прибора  $F = 0,3 \text{ м}$ ,  $M = 100 \text{ м}$ , угол обзора космической съемки  $\beta = 2^\circ$ , угол наблюдения  $\theta = 0^\circ$ . Кривая 1 представляет спектральную зависимость оптической передаточной функции турбулентной атмосферы при учете однократного рассеяния излучения в молекулярно-аэрозольной атмосфере. Кривая 2 показывает спектральную зависимость оптической передаточной функции турбулентной атмосферы при учете многократного молекулярно-аэрозольного рассеяния в случае трехчленной аппроксимирующей индикатрисы рассеяния [2]. Кривая 3 демонстрирует спектральную зависимость оптической передаточной функции  $|\Pi(\nu, \theta)|_{аэ}$  молекулярно-аэрозольной атмосферы в случае трехчленной аппроксимирующей индикатрисы рассеяния без учета атмосферной турбулентности. Кривая 4 представляет зависимость от длины волны  $\lambda$  оптической передаточной функции  $|\Pi(\nu, \theta)|_{турб}$  чисто турбулентной атмосферы без учета многократного аэрозольно-молекулярного рассеяния в ней.

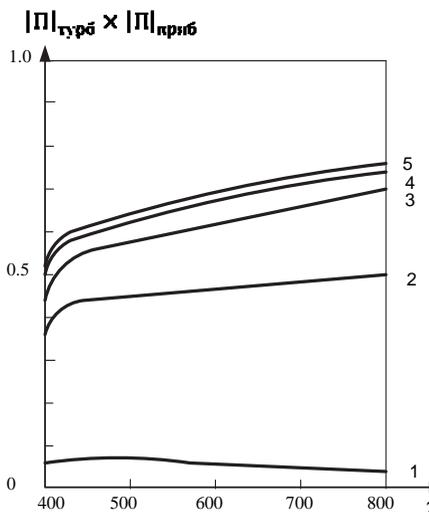


Рис. 8 Спектральная зависимость оптической передаточной функции системы «турбулентная атмосфера - бортовой оптический прибор»: высота орбиты космического носителя

На рисунке 8 представлены спектральные зависимости передаточных функций системы «турбулентная атмосфера - бортовой оптический прибор» для разных пространственных частот  $\nu$ : высота орбиты космического носителя  $H_{\text{орб}}=100\text{км}$ , фокусное расстояние прибора  $F=0,3\text{м}$ , угол обзора съемки  $\beta = 2^\circ$ , угол наблюдения  $\theta = 0^\circ$ ,  $1 - \nu = 100\text{км}^{-1}$ ,  $2 - \nu = 25\text{км}^{-1}$ ,  $3 - \nu = 10\text{км}^{-1}$ ,  $4 - \nu = 4\text{км}^{-1}$ ,  $5 - \nu = 1\text{км}^{-1}$ .

**8. Заключение.** Главный вывод проведенного выше анализа состоит в следующем. В видимой области спектра ( $\lambda=400\text{-}800\text{ нм}$ ) для рассмотренных условий космической съемки наземных (надводных) природных образований в надир ( $\theta=0^\circ$ ) эффектом турбулентности, обусловленной флуктуациями показателя преломления земной атмосферы, можно пренебречь по сравнению с искажениями, вносимыми некогерентным многократным аэрозольным и молекулярным рассеянием, а также влиянием регистрирующей бортовой спутниковой аппаратуры.

Таким образом, при проведении атмосферной коррекции космических снимков и спектров яркости наземных (надводных) природных объектов невысокого и среднего пространственного разрешения достаточно учитывать только искажения, обусловленные некогерентным многократным рассеянием и поглощением излучения в молекулярно-аэрозольной земной атмосфере и трансформирующим влиянием регистрирующей бортовой оптической аппаратуры. Что касается искажений, вносимых

чисто турбулентной земной атмосферой, то их влияние на качество указанной спутниковой видеоинформации следует учитывать только при коррекции фрагментов многоспектральных мелкомасштабных космических изображений очень высокого пространственного разрешения.

### Литература

1. *Кондратьев К.Я., Смоктий О.И., Козодеров В.В.* Влияние атмосферы на исследование природных ресурсов из космоса // М: Машиностроение. 1985. 272 с.
2. *Смоктий О.И.* Моделирование полей излучения в задачах космической спектrophотометрии // Л.: Наука. 1986. 352 с.
3. *Kondratyev K.Ya., Kozoderov V.V., Smokty O.I.* Remote Sensing of the Earth from Space: Atmospheric Correction // Springer-Verlag. 1992. 478p.
4. *Smokty O. I., Guzzi R.* The Influence of the Earth's Atmospheric Turbulence on the Optical System Resolution // Proc. Progress in Electromagnetics Research Symp. (PIERS'1994). Noordwijk. 1994. pp. 2665–2669.
5. *Smokty O.I.* The Influence of the Earth's Atmospheric Turbulence on the Space Optical System Resolution // Proc. Current Problems in Atmospheric Radiation (IRS'2012). New-York: American Inst. Phys. 2013. pp. 372–376.
6. *Kawata Y., Kusaka T., Uenno S., Smokty O.I.* Atmospheric Correction Algorithm for Remote Sensing Data with Multiviewing Angles // In: Proc. IGARSS'93. Tokyo. Japan. 1993. vol. 4. pp. 1113–1115.
7. *Исмару А.* Распространение и рассеяние волн в случайно-неоднородных средах // М.: Мир. 1981. Т.1. 280 с., Т.2. 317 с.
8. *Кляцкин В.И.* Статистическое описание динамических систем с флуктуирующими параметрами // М.: Наука. 1975. 239 с.
9. *Smokty O.I., Kondratyev K.Ya.* Analytical Approximation of the Spatial Frequency Filtration of Radiation Fields for the Space Spectrophotometry Problems // Appl. Math. and Computation. 1995. vol. 80. pp. 1–17.
10. *Smokty O. I.* Applied Mathematical Problems of Satellite Data Filtration and Atmospheric Correction // In: Proc. Current Problems in Atmospheric Radiation (IRS'2000). A.Deepak Publ.Co. Hampton (USA). 2001. pp. 433–437.
11. *Татарский В.И.* Распространение волн в турбулентной атмосфере // М: Наука. 1967. 548 с.
12. *Чандрасекар С.* Перенос лучистой энергии // М.: ИЛ. 1953. 432 с.
13. *Соболев В.В.* Рассеяние света в атмосферах планет // М.: Наука, 1972, 335 с.
14. *Чернов Л.А.* Волны в случайно-неоднородных средах // М.: Наука. 1975. 174 с.
15. *Татарский В.И.* Теория флуктуационных явлений при распространении волн в турбулентной атмосфере // М.: Изд-во АН СССР. 1959. 127 с.
16. *Апресян Х. А., Кравцов Ю.А.* Теория переноса излучения // М: Наука. 1983. 216 с.
17. *Алдошина О.А., Бачериков В.В., Смоктий О.И., Фабриков В.А.* Влияние турбулентной атмосферы на формирование оптических изображений объектов // Л.: ЛИИАН. Препринт. 1991. №186. 61с.
18. *Labeyrie A.* Attainment of Diffraction Limited Resolution in Large Telescopes by Fourier's Analysis Speckle Patterns in Star Images // Astron. and Astrophys. 1970. vol. 6. pp. 85–87.
19. *Roddier F.* The Effect of Atmospheric Turbulence in Optical Astronomy // Progress in Optics. 1981. vol. XIX. pp. 283–376.
20. *Франсон М.* Оптика спеклов // М: Мир. 1980. 171с.
21. *Зуев В.У., Кабанов М.В.* Перенос оптических сигналов в земной атмосфере (в условиях помех) // М.: Советское радио. 1977. 367 с.

22. *Смокты О.И., Фабриков В.А.* Методы теории систем и преобразований в оптике // Л.: Наука. 1989. 312 с.
23. *Browne L.W., Antonia R.A., Rajagopalam S.* The Spectral Derivative of Temperature in a Turbulent Flow and Taylor's Hypothesis // *Phys. Fl.* 1983. vol. 26. no. 5. pp. 1222–1227.
24. *Клиффорд С.Ф.* Классическая теория распространения волн в турбулентной среде // В кн.: Проблемы прикладной физики. М.: Мир. 1981. Т.25. С. 18–60.
25. *Колмогоров А.Н.* Рассеяние энергии при локально-изотропной турбулентности // Докл. АН СССР. 1941. Т. 32. С. 19–21.
26. *Ламли Д., Пановский Г.* Структура атмосферной турбулентности // М.: Мир. 1966. 264 с.
27. *Борн М., Вольф Э.* Основы оптики // М.: Наука. 1970. 856 с.
28. *Рытов С.М.* Введение в статистическую радиофизику. Часть I. Случайные процессы // М.: Наука. 1976. 280 с.
29. *Рытов С.М., Крайцов Ю.Ф., Татарский В.И.* Введение в статистическую радиофизику. Часть II. Случайные поля // М.: Наука. 1978. 463 с.
30. *Марешаль А., Франсон М.* Структура оптического изображения // М.: Мир. 1964. 295 с.
31. *Hufnagel R.E., Stanley N.R.* Modulation Transfer Function Associated with Image Transmission through Turbulent Media // *Journ. Opt. Soc. Amer. (JOSA)*. 1964. vol. 54. no. 1. pp. 52–61.
32. *Lutomirsky R.F., Yura H.T.* Imaging of Extended Objects through a Turbulent Atmosphere // *Appl. Optics*. 1974. vol. 13. pp. 431–437.
33. *Fried D.L.* Optical Resolution through a Randomly Inhomogeneous Medium for Very Long and Very Short Exposures // *J. Opt. Soc. Amtr. (JOSA)*. 1966. vol. 56. no. 10. pp. 1372–1379.
34. *Беленький М.С., Боронаев В.В., Гамбоев Н.Ц., Миронов В.Л.* Оптическое зондирование атмосферы // Новосибирск: Наука. 1986. 91с.
35. *Smokty O. I.* Applied Mathematical Problems of Atmospheric Correction and Filtration of Multispectral Satellite Data // In: Proc. CEReS Intern. Symp. «The Atmospheric Correction of Satellite Data and Its Application to Global Environment». Chiba Univ. Japan. 1998. pp. 143–150.
36. *Elterman L.* Vertical Attenuation Model with Eight Meteorological Ranges 2 to 13 kilometers // Rep. ACCRL-70-0200. Environ. Res. Papers. Bedford. USA. 1970. vol. 318. 57p.

## References

1. Kondratyev K.Ya., Smokty O.I., Kozoderov V.V. *Vliyanie atmosfery na issledovanie prirodnih resursov iz kosmosa* [Atmospheric Influence on the Natural Resources Investigation from Space]. M: Mashinostroenie. 1985. 272 p. (In Russ.).
2. Smokty O.I. *Modelirovanie poley izlucheniya v zadachakh kosmicheskoy spectrophotometrii* [Radiation Fields Modeling in the Problems of Space Spectrophotometry]. L. Nauka. 1986. 352 p. (In Russ.).
3. Kondratyev K.Ya., Kozoderov V.V., Smokty O.I. Remote Sensing of the Earth from Space: Atmospheric Correction. Springer-Verlag. 1992. 478 p.
4. Smokty O. I., Guzzi R. The Influence of the Earth's Atmospheric Turbulence on the Optical System Resolution. Proc. Progress in Electromagnetics Research Symp. (PIERS'1994). Noordwijk. 1994. pp. 2665–2669.
5. Smokty O.I. The Influence of the Earth's Atmospheric Turbulence on the Space Optical System Resolution. Proc. Current Problems in Atmospheric Radiation (IRS'2012). New-York: American Inst. Phys. 2013. pp. 372–376.

6. Kawata Y., Kusaka T., Ueno S., Smokty O.I. Atmospheric Correction Algorithm for Remote Sensing Data with Multiviewing Angles. In: Proc. IGARSS'93. Tokyo. Japan. 1993. vol. 4. pp. 1113–1115.
7. Ishimaru A. *Rasprostranenie i rasseyanie voln v sluchaino-neodnorodnykh sredakh* [Wave Propagation and Scattering in Random media]. M.: Mir. 1981. vol. 1. 280 p., vol. 2. 317 p. (In Russ.).
8. Kljatzkin V.I. *Statisticheskoe opisanie dinamicheskikh sistem s flyktuirushimi parametrami* [Statistical description of dynamical systems with the fluctuating parameters]. M.: Nauka. 1975. 239 p. (In Russ.).
9. Smokty O.I., Kondratyev K.Ya. Analytical Approximation of the Spatial Frequency Filtration of Radiation Fields for the Space Spectrophotometry Problems. Appl. Math. and Computation. 1995. vol. 80. pp. 1–17.
10. Smokty O. I. Applied Mathematical Problems of Satellite Data Filtration and Atmospheric Correction. In: Proc. Current Problems in Atmospheric Radiation (IRS'2000). A. Deepak Publ.Co. Hampton (USA). 2001. pp. 433–437.
11. Tatarskiy V.I. *Rasprostranenie voln v turbulentnoy atmosfere* [Wave Propagation in a Turbulent Atmosphere]. M.: Nauka. 1967. 548 p. (In Russ.).
12. Chandrasekhar S. *Perenos luchistoy energii* [Radiative transfer]. M.: IL. 1953. 432 p. (In Russ.).
13. Sobolev V.V. *Rasseyanie sveta v atmosferah planet* [Light Scattering in Planetary Atmospheres]. M.: Nauka. 1972. 335 p. (In Russ.).
14. Chernov L.A. *Volny v sluchaino-neodnorodnykh sredakh* [Waves in the Random-Non-uniform Media]. M.: Nauka. 1975. 174 p. (In Russ.).
15. Tatarskiy V.I. *Teoriya fluktuazionnykh javleniy pri rasprostranении voln v turbulentnoy atmosfere* [Fluctuating Phenomena Theory for Waves Propagation in the Turbulent Atmosphere]. M.: AN SSR. 1959. 127 p. (In Russ.).
16. Apresjan X.A., Kravzov I.A. *Teoriya perenosa izlucheniya* [Radiative Transfer Theory]. M.: Nauka. 1983. 216 p. (In Russ.).
17. Aldoshina O.A., Bacherikov V.V., Smokty O. I. Fabrikov V.A. *Vliyanie turbulentnoy atmosfery na formirovaniye opticheskikh izobrascheniy obyektov* [Turbulent Atmosphere Influence on the Objects Optical Images Forming]. L. LIAN. Preprint. 1991. no. 186. 61 p. (In Russ.).
18. Labeyrie A. Attainment of Diffraction Limited Resolution in Large Telescopes by Fourier's Analysis Speckle Patterns in Star Images. Astron. and Astrophys. 1970. vol. 6. pp. 85–87.
19. Roddier F. The Effect of Atmospheric Turbulence in Optical Astronomy // Progress in Optics. 1981. vol. XIX. pp. 283–376.
20. Franon M. *Optika spektrov* [Spectral optics]. M.: Mir. 1980. 17 p. (In Russ.).
21. Zuev V.E., Kabanov M.V. *Perenos opticheskikh signalov v zemnoy atmosfere (v uslovijah pomeh)* [Optical Signals Transfer in the Earth's Atmosphere (in encumbrances conditions)]. M.: Sovetskoe radio. 1977. 367 p. (In Russ.).
22. Smokty O. I., Fabrikov V.A. *Metody teorii sistem i preobrazovaniya v optike* [Systems Theory Methods and Optics Transformations]. L.: Nauka. 1989. 312 p. (In Russ.).
23. Browne L.W., Antonia R.A., Rajagopalam S. The Spectral Derivative of Temperature in a Turbulent Flow and Taylor's Hypothesis. Phys. Fl. 1983. vol. 26. no. 5. pp. 1222–1227.
24. Klifford S.F. [Classical Theory of Wave Propagation in a Turbulent Media]. In.: *Problemy prikladnoy fiziki – Applied Physics Problems*. M.: Mir. 1981. vol. 25. pp. 1–60. (In Russ.).
25. Kolmogorov A. N. *Rassejanie energii pri lokalno-izotropnoy turbulentnosti* [Energy scattering at local-isotropic turbulence]. Dokl. AN SSR. 1941. vol. 32. pp. 19–21. (In Russ.).
26. Lamli D., Panovsky G. *Structura atmosfernoy turbulentnosti* [The structure of

- atmospheric turbulence]. M.: Mir. 1966. 264 p. (In Russ.).
27. Born M., Volf E. *Osnovy optiki* [Optics Basics]. M.: Nauka. 1970. 856 p. (In Russ.).
  28. Rytov S.M. *Vvedenie v statisticheskuyu radiofiziku. Ch.I. Sluchaynye protsessy* [Introduction to Statistical Radiophysics. Part 1. Random processes]. M.: Nauka. 1976. 280 p. (In Russ.).
  29. Rytov S.M., Kravzov J.A. Tatarskiy V.I. *Vvedenie v statisticheskuyu radiofiziku. Ch.II. Sluchaynye polja* [Introduction to Statistical Radiophysics. Part 2. Random fields]. M.: Nauka. 1978. 463 p. (In Russ.).
  30. Marechal A., Franson M. *Struktura opticheskogo izobrascheniya* [Optical Images Structure]. M.: Mir. 1964. 295p. (In Russ.).
  31. Hufnagel R.E., Stanley N.R. Modulation Transfer Function Associated with Image Transmission through Turbulent Media. *Journ. Opt. Soc. Amer. (JOSA)*. 1964. vol. 54. no. 1. pp. 52–61.
  32. Lutomirsky R.F., Yura H.T. Imaging of Extended Objects through a Turbulent Atmosphere. *Appl.Optics*. 1974. vol. 13. pp. 431–437.
  33. Fried D.L. Optical Resolution through a Randomly Inhomogeneous Medium for Very Long and Very Short Exposures. *J.Opt.Soc.Amtr. (JOSA)*. 1966. vol. 56. no. 10, pp. 1372–1379.
  34. Belenkiy M.S., Boronae V.V., Gamboev N.Z., Mironov V.I. *Opticheskoe zondirovanie atmosfery* [Optical Remote Sensing of the Atmosphere]. Novosibirsk: Nauka. 1986. 91p.
  35. Smokty O. I. Applied Mathematical Problems of Atmospheric Correction and Filtration of Multispectral Satellite Data. In: *Proc. CERES Intern. Symp. «The Atmospheric Correction of Satellite Data and Its Application to Global Environment»*. Chiba Univ. Japan. 1998. pp. 143–150.
  36. Elterman L. Vertical Attenuation Model with Eight Meteorological Ranges 2 to 13 kilometers. *Rep. ACCRL-70-0200. Environ. Res. Papers. Bedford. USA*. 1970. vol. 318. 57p.

**Смоктый Олег Иванович** — д-р физ.-мат. наук, профессор, Заслуженный Деятель Науки Российской Федерации, академик Международной Академии Астронавтики (Франция), главный научный сотрудник лаборатории прикладной информатики и проблем информатизации общества Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: теория переноса излучения, моделирование полей излучения в задачах дистанционного зондирования Земли из космоса, космическая геоинформатика. Число научных публикаций — 270 научных трудов, в том числе 10 монографий, одного открытия и 6 изобретений, e-mail: soi@ias.spb.su СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, Россия. р.т. +7-(812)-328-19-19

**Smokty Oleg Ivanovich**— Ph.D., Dr. Sci., professor, Honored Scientist of Russian Federation, Full Member of the International Academy of Astronautics (France), Chief Researcher at the Laboratory of Applied Informatics and Problems of Sociert Informatization of St. Petersburg Institute for Informatics and Automation of Russian Academy of Sciences (SPIIRAS). Research interests: Radiative Transfer Theory; Radiation Field Modeling in Problems of the Earth Remote Sensing from Space; Geoinformatics. The number of publications — 270 scientific works, including 10 monographs, 1 discovery and 6 inventions, e-mail: soi@ias.spb.su, SPIIRAS, 39,14<sup>th</sup> Line VO, 199178 St. Petersburg, Russia; office phone +7-(812)-328-19-19.

**Поддержка исследований.** Работа выполнена по Программе Фундаментальных исследований ОНИТ РАН и в соответствии с планом НИР СПИИРАН.

**Acknowledgements.** This research is supported by ONIT RAS project and as consistent with schedule of SPIIRAS research.

## РЕФЕРАТ

**Смоктый О.И. Влияние атмосферной турбулентности, молекулярно-аэрозольного рассеяния и параметров бортовой оптической аппаратуры на качество спутниковой видеoinформации при съемке Земли из космоса.**

В настоящей работе рассмотрена важная проблема совместного влияния атмосферной турбулентности, молекулярно-аэрозольного многократного рассеяния и параметров спутникового бортового оптического прибора на качество космической видеoinформации, получаемой при дистанционном спектрофотометрировании Земли. Аналитические и численные оценки комплексной пространственно-частотной передаточной функции линейной оптической системы «атмосфера – подстилающая поверхность – бортовой прибор» выполнены в рамках концепции линейных динамических систем и Фурье-преобразований оптических сигналов. Главное заключение проведенного анализа состоит в следующем. Для дистанционного зондирования природной среды из космоса в видимой ( $\lambda = 400-800$  нм) и ближней инфракрасной областях ( $\lambda = 850-1500$  нм) спектра при условиях съемки в надир ( $\theta = 0^\circ$ ), высоте орбиты космического носителя ( $H_{\text{орб}} \geq 10^5$  м) и его скорости ( $V \geq 10^3$  м/сек) влиянием атмосферной турбулентности, вызываемой флуктуациями показателя преломления земной атмосферы, можно пренебрегать в масштабах  $\Delta > 10\mu$  по сравнению с искажениями, вызываемыми молекулярно-аэрозольным многократным рассеянием и регистрирующим бортовым прибором. Таким образом, при атмосферной коррекции спутниковой видеoinформации невысокого пространственного разрешения достаточно ограничиться рассмотрением процессов многократного некогерентного рассеяния и поглощения солнечного излучения на пути от бортового оптического регистратора до наземного (надводного) объекта.

## SUMMARY

### ***Smokty O.I.* Influence of turbulent atmosphere, molecular-aerosol scattering and on-board optical devices on the satellite video data quality at the Earth's remote sensing from space.**

Important problem of conjoint influence of atmospheric turbulence, molecular-aerosol multiple light scattering and on-board satellite optical devices' parameters on the quality of video data obtained at the Earth's spectrophotometric remote sensing from the space is considered. Analytical and numeric estimations of the complex spatial-frequency transfer function for the linear optical system "atmosphere – underlying surface – on-board device" considered in the framework of both linear dynamic systems concept and Fourier transformations of the optical signals are given. The main conclusion of given analysis is as follows. In order to carry out remote sensing of environment from space in both visual ( $\lambda=400-800$  nm) and nearest infra-red ( $\lambda=850-1500$  nm) spectral fields under condition of nadir survey ( $\Theta = 0^\circ$ ), space carrier's orbit height ( $H_{\text{orb}} \geq 10^5$  m) and its velocity ( $v \geq 10^3$  m/sec), the atmospheric turbulence influence induced by the Earth's atmospheric refraction index fluctuations may be neglected in scale  $\Delta > 10$  m in comparison with distortions induced by both molecular-aerosol multiple light scattering and registering on-board device. Therefore, in case of the atmospheric correction applied for the low-quality satellite video data spatial resolution, the analysis of solar radiance multiple non-coherent scattering and absorption processes taken on its route from on-board device to the ground-based or overwater-based object should be sufficiently used.

## РУКОВОДСТВО ДЛЯ АВТОРОВ



Вып. 5(36)

ТРУДЫ СПИИРАН  
proceedings.spiiras.nw.ru

Взаимодействие автора с редакцией осуществляется через личный кабинет на сайте журнала «Труды СПИИРАН» <http://www.proceedings.spiiras.nw.ru>. При регистрации авторам рекомендуется заполнить все предложенные поля данных, так как это значительно ускорит процесс оформления метаданных к новым статьям.

Подготовка статьи ведется с помощью текстовых редакторов MS Word 2007 и выше. При подаче материала в редакцию сначала отправляется только статья в формате \*.docx. Для обеспечения требований слепого рецензирования при представлении статьи в журнал авторам необходимо удалить персональные данные, содержащиеся в тексте файла и его свойствах.

Объем основного текста – от 5 до 20 страниц включительно. Формат страницы документа – А5 (148 мм ширина, 210 мм высота); ориентация – портретная; все поля – 20 мм. Верхний и нижний колонтитулы страницы – пустые. Основной шрифт документа – Times New Roman, основной кегль (размер) шрифта – 10 pt. Переносы разрешены. Абзацный отступ устанавливается размером в 10 мм. Межстрочный интервал – одинарный. Номера страниц не проставляются.

Не допускается использования цветных шрифтов, цветовых выделений и цветных рисунков. Статьи должны быть полностью готовы к черно-белой печати.

Основная часть текста статьи разбивается на разделы, среди которых являются обязательными: введение, хотя бы один «содержательный» раздел и заключение. Допускается также мотивированное содержанием и структурой материала выделение подразделов.

В основную часть допускается помещать рисунки, таблицы, листинги и формулы. Правила их оформления подробно рассмотрены на нашем сайте в разделе «Руководство для авторов».

ISSN 2078-9181



9 772078 918785 >

С-П6  
2014