

# В Е С Т Н И К

## НОВОСИБИРСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

Научный журнал  
Основан в ноябре 1999 года

Серия: Информационные технологии

2018. Том 16, № 4

### СОДЕРЖАНИЕ

<i>Абрамский М. М., Тимерханов Т. И.</i> Сравнительный анализ использования реляционных и графовых баз данных в разработке цифровых образовательных систем	5
<i>Барамия Д. А., Горбенко Н. И., Лаврентьев М. М.</i> Описание долговременной эволюции берегового профиля на основе диффузионной модели	13
<i>Баталин К. В., Мамеев Н. С., Попова К. Ю., Рыжаков И. Д., Яхьяева Г. Э.</i> Программная система управления образовательным процессом ИТОС	20
<i>Боронина М. А., Вшивков В. А.</i> Параллельный алгоритм для численного моделирования встречных пучков ультрарелятивистских частиц с учетом угла встречи	31
<i>Глинских В. Н., Никитенко М. Н.</i> Теоретико-алгоритмическая база и компьютерное моделирование данных диэлектрического каротажа для изучения частотного спектра электрофизических параметров геологической среды	41
<i>Журавлев С. С., Рудометов С. В., Окольников В. В., Шакиров С. Р.</i> Применение модельно-ориентированного проектирования к созданию АСУ ТП опасных промышленных объектов	56
<i>Карин Ю. Г., Балков Е. В., Фадеев Д. И., Манитейн А. К., Панин Г. Л., Алымов А. О., Романов Д. Б.</i> Электромагнитное профилирование компактной аппаратурой: новый подход и результаты применения	68
<i>Лысаков К. Ф., Облаухов К. К., Шадрин М. Ю.</i> Применение FPGA для ускорения расчета волнового фронта цунами	79
<i>Лях Т. В., Зюбин В. Е., Гаранина Н. О.</i> Автоматизированная верификация алгоритмов управления сложными технологическими объектами на программных имитаторах	85
<i>Махина Е. Д., Пальчунов Д. Е.</i> Программная система для определения речевых действий в текстах естественного языка	95
<i>Молородов Ю. И., Ходорченко Д. А.</i> Разработка информационной системы по клещевой опасности на основе онтологии предметной области	107
<i>Мызников П. В.</i> Разработка прецедентно-ориентированного подхода обратного реинжиниринга web-интерфейсов	115

<i>Нечаев О. В., Глинских В. Н.</i> Трехмерное моделирование и инверсия данных комплекса методов электрокаротажа в моделях сред с наклоном главных осей тензора электрической анизотропии	127
<i>Нугуманова А. Б., Байбурин Е. М., Мансурова М. Е., Барахнин В. Б.</i> Автоматическое извлечение решеток понятий из медицинских текстов на основе комбинации анализа формальных понятий и технологий бутстраппинга	140
<i>Титов П. А.</i> Моделирование упругих волн в средах со сложной топографией свободной поверхности	153
Сведения об авторах	167
Информация для авторов	170

# VESTNIK

## NOVOSIBIRSK STATE UNIVERSITY

Scientific Journal  
Since 1999, November  
In Russian

Series: Information Technologies

2018. Volume 16, № 4

---

---

### CONTENTS

<i>Abramskiy M. M., Timerkhanov T. I.</i> Comparative Analysis of Relational and Graph Databases Use in Digital Education Systems Development	5
<i>Baramiya D. A., Gorbenko N. I., Lavrentiev M. M.</i> Description of Coastal Profile Long Term Evolution by Diffusion Model	13
<i>Batalin K. V., Mameev N. S., Popova K. Yu., Ryzhakov I. D., Yakhyaeva G. E.</i> Software System for Management of Educational Process ITOS	20
<i>Boronina M. A., Vshivkov V. A.</i> Parallel Algorithm for Numerical Simulations of Ultra Relativistic Particle Beams Colliding with Crossing Angle	31
<i>Glinskikh V. N., Nikitenko M. N.</i> Theoretical-Algorithmical Base and Computer Simulation of Dielectric Logging Data to Study Frequency Spectrum of Electrophysical Parameters of Geological Environment	41
<i>Zhuravlev S. S., Rudometov S. V., Okolnishnikov V. V., Shakirov S. R.</i> Model-Based Design Approach for Development Process Control Systems of Hazardous Industrial Facilities	56
<i>Karin Yu. G., Balkov E. V., Fadeev D. I., Manshtein A. K., Panin G. L., Alymov A. O., Romanov D. B.</i> Electromagnetic Profiling by Compact Apparatus: New Approach and Results of Application	68
<i>Lysakov K. F., Oblaukhov K. K., Shadrin M. Yu.</i> Using FPGA to Accelerate Computation of Tsunami Wavefront	79
<i>Liakh T. V., Zyubin V. E., Garanina N. O.</i> Automatic Verification of Control Algorithms for Complex Technological Objects on Software Simulators	85
<i>Makhina E. D., Palchunov D. E.</i> Software System for the Determination of Speech Acts in Natural Language Texts	95
<i>Molorodov Yu. I., Khodorchenko D. A.</i> Development of the Information System for Tick-Hazard on the Basis of the Objective Ontology	107
<i>Myznikov P. V.</i> Development of the Case-Based Approach of Web Interfaces Reverse Reengineering	115
<i>Nechaev O. V., Glinskikh V. N.</i> Three-Dimensional Simulation and Inversion of Lateral Logging Sounding and Lateral Logging Data in Media with Tilt of the Main Axes of the Dielectric Anisotropy Tensor	127

<i>Nugumanova A. B., Bayburin E. M., Mansurova M. E., Barakhnin V. B.</i> Automatic Extraction of Formal Lattices from Medical Texts Based on The Combination of the Formal Concept Analysis and Bootstrapping Technologies	140
<i>Titov P. A.</i> Elastic Waves Modeling in Media with Complex Free Surface Topography	153
Our Contributors	167
Instructions to Contributors	170

*Editor in Chief* Anatolij M. Fedotov

*Vice-Editor* A. V. Avdeev

*Executive Secretary* N. N. Pestereva

*Editorial Board of the Series*

- I. V. Bychkov*, professor, academician (Irkutsk), *B. M. Glinsky*, professor (Novosibirsk)  
*A. N. Gorban'*, professor (Lester, GB), *E. P. Gordov*, professor (Tomsk)  
*B. S. Dobronets*, professor (Krasnoyarsk), *A. M. Elizarov*, professor (Kazan)  
*G. N. Erokhin*, professor (Kaliningrad), *A. I. Kamyshnikov*, professor (Khanty-Mansijsk)  
*G. P. Karev*, professor (Maryland, USA), *N. A. Kolchanov*, professor, academician (Novosibirsk)  
*M. M. Lavrentjev*, professor (Novosibirsk), *V. E. Malyshkin*, professor (Novosibirsk)  
*N. N. Mirenkov*, professor (Aizu, Japan), *N. M. Oskorbin*, professor (Barnaul)  
*D. E. Palchunov*, professor (Novosibirsk), *T. Pizansky*, professor (Ljubljana, Slovenia)  
*V. P. Potapov*, professor (Kemerovo), *O. I. Potaturkin*, professor (Novosibirsk)  
*V. A. Serebryakov*, professor (Moscow), *A. V. Starchenko*, professor (Tomsk)  
*S. I. Smagin*, professor, corresponding member of RAS (Khabarovsk)  
*D. A. Tusupov*, professor (Astana, Kazakhstan)  
*V. V. Shajdurov*, professor, corresponding member of RAS (Krasnoyarsk)  
*Yu. I. Shokin*, professor, academician (Novosibirsk)

*The journal is published quarterly in Russian since 1999  
by Novosibirsk State University Press*

*The address for correspondence*

Faculty of Information Technologies, Novosibirsk State University

1 Pirogov Street, Novosibirsk, 630090, Russia

*Tel.* +7 (383) 363 42 46

*E-mail address:* [inftech@vestnik.nsu.ru](mailto:inftech@vestnik.nsu.ru)

*On-line version:* <http://elibrary.ru>

**М. М. Абрамский, Т. И. Тимерханов**

*Казанский (Приволжский) федеральный университет  
ул. Кремлевская, 35, Казань, 420008, Россия*

*ma@it.kfu.ru, timurgrunge@gmail.com*

## **СРАВНИТЕЛЬНЫЙ АНАЛИЗ ИСПОЛЬЗОВАНИЯ РЕЛЯЦИОННЫХ И ГРАФОВЫХ БАЗ ДАННЫХ В РАЗРАБОТКЕ ЦИФРОВЫХ ОБРАЗОВАТЕЛЬНЫХ СИСТЕМ**

Рассмотрены вопросы выбора варианта реализации базы данных при разработке цифровых образовательных систем. Проведен сравнительный анализ реляционного и графового подходов к хранению базовых сущностей. Продемонстрирован ряд преимуществ графовой модели в контексте удобства разработки и эффективности реализации запросов.

*Ключевые слова:* графовые базы данных, реляционные базы данных, Neo4j, цифровые образовательные системы.

### **Введение**

В настоящее время лидирующее положение среди средств хранения данных занимают системы управления базами данных (СУБД), основанные на реляционном подходе<sup>1</sup> (рис. 1): на трех лидирующих позициях – реляционные СУБД Oracle, MySQL и MS SQL, на 4-м месте – также реляционная СУБД PostgreSQL; самая распространенная нереляционная СУБД MongoDB занимает 5-е место, на 6-м месте вновь находится реляционная БД DB2.

Реляционный подход был сформулирован в 1969–1970 гг. Э. Ф. Коддом [1]. Он заключается в хранении данных в таблицах (*реляциях*) с определенными между ними связями и ограничениями для поддержки целостности, корректности и неизбыточности данных.

Однако за последние несколько лет<sup>2</sup> начала расти доля рынка СУБД, реализующих альтернативные модели хранения данных, которые объединяют в класс нереляционных (NoSQL, “not only SQL”). Такие модели имеют преимущество перед реляционными БД при работе со специфическими структурами данных, в частности они могут превосходить реляционные СУБД

в работе со слабоструктурированными данными или с данными, явным образом связанными с визуальным представлением информации (графы, диаграммы и др.). На рис. 2 представлена динамика роста популярности использования всех нереляционных СУБД независимо от модели, например, MongoDB – документоориентированная СУБД, Redis – СУБД, реализующая модель «ключ – значение», Neo4j – графовая БД.

---

<sup>1</sup> DB-Engines Ranking 2018. URL: <https://db-engines.com/en/ranking>

<sup>2</sup> Ranking of the most popular database management systems worldwide, as of February 2018. URL: <https://www.statista.com/statistics/809750/worldwide-popularity-ranking-database-management-systems>

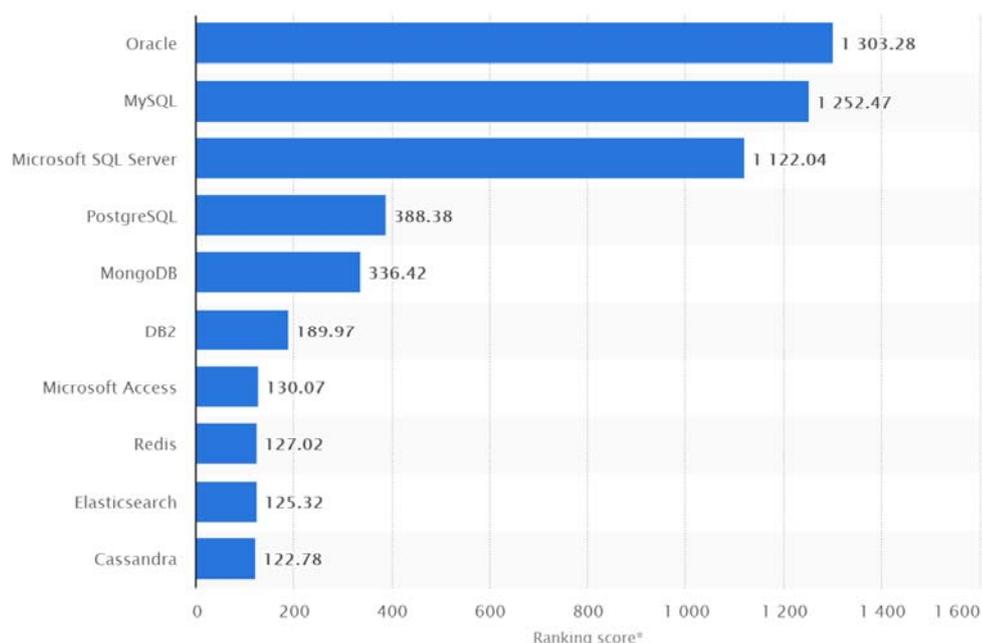


Рис. 1. Доли рынка, занимаемые системами управления базами данных  
(URL: <https://www.statista.com/statistics/809750/worldwide-popularity-ranking-database-management-systems>)

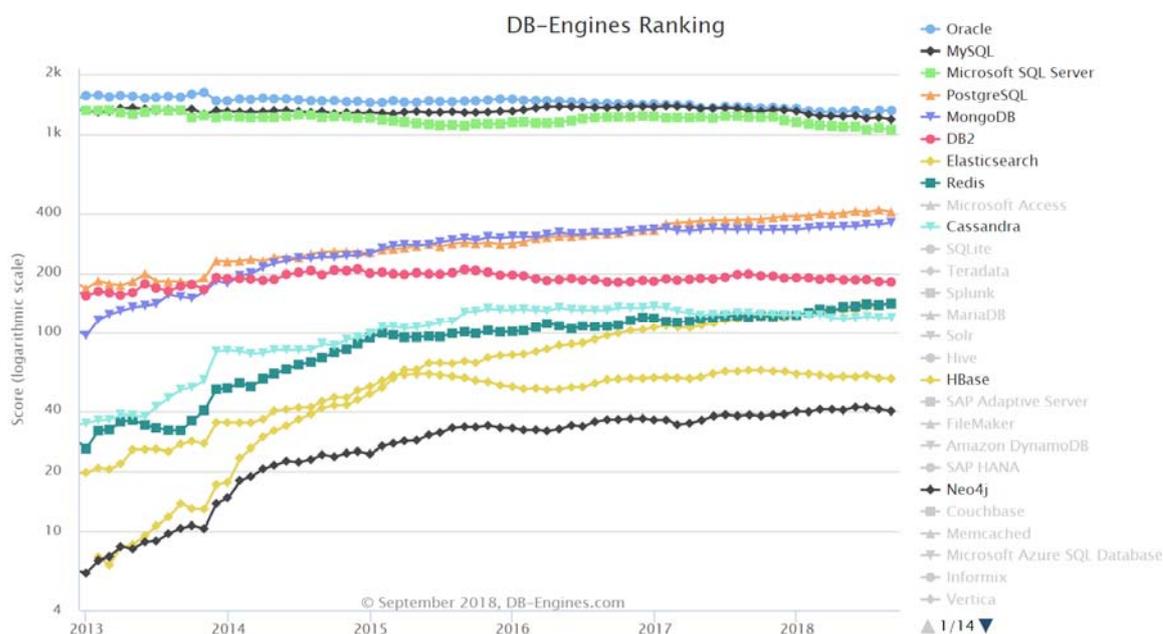


Рис. 2. Динамика использования СУБД с 2013 по 2018 г.  
(URL: <https://www.statista.com/statistics/809750/worldwide-popularity-ranking-database-management-systems>)

Вопрос выбора подхода к организации базы данных актуален при разработке цифровых образовательных систем и сред. В последние годы в сфере образования наблюдается тенденция к индивидуализации [2], что означает организацию учебного процесса с учетом индивидуальных особенностей каждого обучающегося. В [3] показана возможность реализации такой индивидуализации с помощью специальных цифровых инструментов, которые анализируют перечень компетенций, достижений и характеристик обучающегося и генерируют его индиви-

дуальную траекторию, представляющую собой граф. Предложенный способ генерации основан на обработке образовательных программ вуза. Сущность «образовательная программа» с точки зрения данных может быть представлена в виде связанных ориентированных графов, описывающих зависимости между учебными дисциплинами, компетенциями, модулями и др.

Авторами данной статьи было сделано предположение, что при работе с сущностями предметной области графовые базы данных могут иметь преимущество в эффективности работы и удобстве разработки в сравнении с реляционными. В данной работе рассмотрены реализации хранения данных, характерных для цифровых образовательных систем, в реляционной (PostgreSQL) и графовой (Neo4J) системах управления базами данных. Сравнительный анализ реляционных и графовых баз данных проводился в некоторых работах (например, [4]), нашей же целью является сравнительный анализ именно удобства использования реляционной и графовой моделей при разработке цифровых образовательных систем в контексте модели данных.

### Общие сущности образовательных систем

Нами были проанализированы образовательные программы высшего и дополнительного образования и выявлены сущности баз данных, характерные для каждой образовательной системы. Любая информационная система, которую можно назвать образовательной, хранит данные, характерные для таких систем.

Отметим, что работа посвящена разработке образовательных систем с технической точки зрения, поэтому, несмотря на различия в образовательной терминологии, некоторые разные понятия были рассмотрены с точки зрения данных как одна и та же сущность. Именно поэтому ниже приведены все подходящие термины, а в скобках – вариант, выбранный в рамках данной статьи.

- **Модуль / предмет / дисциплина (модуль).** Содержит информацию об учебном предмете: название, цель, методы обучения. Термин *модуль* может также подразумевать некий обособленный раздел дисциплины, который можно изучать отдельно, например, модуль «Язык SQL» в курсе «Базы данных».

- **Образовательный результат / компетенция / знание, умение, навык (компетенция).** То, что может / чем владеет обучающийся по окончании образования. Если рассматривать высшее образование, то образовательными результатами являются компетенции ФГОС и связанные с ними знания, умения и навыки. Программы дополнительного образования, корпоративного образования, переподготовки могут также определять конкретные требования к сотрудникам, прошедшим обучение.

- **Направление подготовки / позиция / специальность (направление).** То, на что ориентировано образование – на получение диплома по конкретному направлению или, в случае дополнительного или корпоративного образования, на некоторую специализацию или соответствие некоторой вакансии на рынке труда.

Между выделенными сущностями можно определить следующие типы отношений:

- **Модуль – Модуль, связь «пререквизит».** Описывает зависимость между двумя модулями, когда изучение одного должно предшествовать изучению другого.

- **Модуль – Модуль, связь «иерархия».** Описывает зависимость между двумя предметами, когда один модуль является частью другого модуля (как показано выше с модулем «Язык SQL» в модуле «Базы данных»).

- **Модуль – Компетенция.** Компетенции / знания / образовательные результаты, получаемые при изучении предмета.

- **Направление – Компетенции.** Компетенции / знания / умения / навыки, которые необходимо освоить в рамках обучения по данному направлению.

### Реализация модели данных в реляционных БД

Структурным аспектом реляционной модели данных являются отношения (таблицы) [1]. Для описания структуры реляционной базы данных, а также управления (создание, модифи-

кация, удаление) данными в уже существующем хранилище используется декларативный язык программирования Structured Query Language (SQL) <sup>3</sup>.

Таблицы реляционных баз данных можно рассматривать в двух аспектах. С одной стороны, они могут пониматься как структуры для хранения выделенных сущностей. На этом понимании основан, в том числе, подход Object-Relational Mapping (объектно-реляционное отображение) <sup>4</sup>, когда структура таблицы понимается как некоторый класс, столбец – как атрибут класса, а строка – как объект класса.

С другой стороны, реляционный подход к БД предполагает, что таблица хранит сами отношения между сущностями. Так, например, можно рассмотреть отношение «Подписка» между пользователями в социальных сетях. Подписка является не сущностью в привычном понимании, а отношением между двумя экземплярами сущности «Пользователь» (Пользователь1 подписан на Пользователь2).

Схема реляционной базы данных для модели данных представлена на рис. 3. Таблицы Module (модуль), Major (направление), Competence (компетенции) содержат сущности, описанные выше. Все связи являются отношением типа «многое-ко-многом» (Many-To-Many), что традиционно реализуется введением так называемых «промежуточных таблиц» для хранения этих связей. Кроме внешних ключей, эти таблицы могут содержать дополнительные данные. Например, таблица modules-major может содержать информацию о семестрах, в которых изучается модуль для конкретной образовательной программы. Отметим, что в рамках этой статьи отдельная спецификация таких данных не требуется.

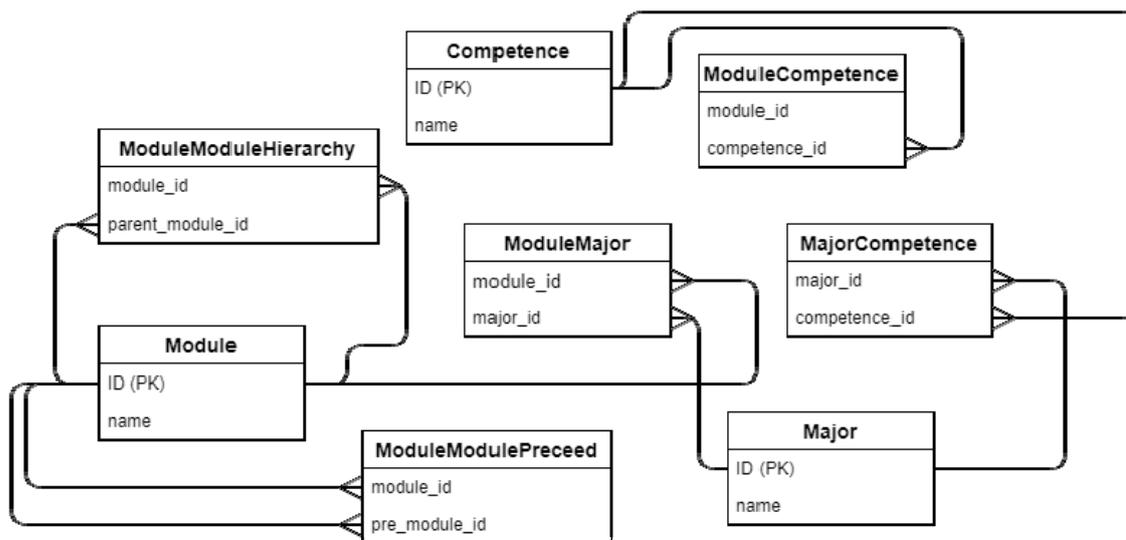


Рис. 3. Схема реляционной БД

Использование реляционного подхода для баз данных образовательных систем имеет несколько недостатков. Во-первых, реляционная схема не различает таблицы, хранящие сущности, и таблицы, хранящие связи между объектами. Во-вторых, визуализация результатов запросов к реляционным БД опять же выглядит как таблица, которая не является естественным способом визуализации данных, носящих графовый характер.

Не уменьшая достоинств реляционных БД, для повышения качества и удобства разработки цифровых образовательных систем необходимо обратиться к другим моделям хранения дан-

<sup>3</sup> Information Technology – Database Language SQL. URL: <http://www.contrib.andrew.cmu.edu/~shadow/sql/sql1992.txt>

<sup>4</sup> Mapping Objects to Relational Databases: O/R Mapping In Detail. URL: <http://www.agiledata.org/essays/mappingObjects.html>

ных, имеющим собирательное название NoSQL (“not only SQL”). Они призваны не целиком заменить существующие реляционные решения, а дополнить их там, где они недостаточно гибки и удобны. Создатели нереляционных решений отмечают более высокую производительность при использовании специфических моделей данных и легкость работы с ними [5].

Одна из классификаций NoSQL базы данных приведена в [6]. Основываясь на модели данных, можно разделить нереляционные хранилища на 4 группы:

- 1) Базы данных NoSQL на основе модели «ключ – значение» (Redis, MemcacheDB и т. п.);
- 2) хранилища колонок (Cassandra, HBase);
- 3) документоориентированные базы данных (MongoDB, Couchbase);
- 4) графовые СУБД (OrientDB, Neo4J).

Графовые БД являются одним из наиболее популярных и актуальных подвидов нереляционных хранилищ [5]. Такие БД оперируют представлениями данных в виде графов и позволяют эффективно совершать привычные для них операции.

### Использование графовых баз данных в цифровых образовательных системах

Для реализации хранения модели данных, описанной выше, была выбрана Neo4j – самая популярная на сегодняшний день графовая СУБД [6]. В основе ее работы лежит аппарат ориентированных графов. Узлы (nodes) хранят основную информацию об объектах и сгруппированы по видам с помощью специальных меток (labels). Также узлы связаны друг с другом отношениями (ребрами), в метках которых можно хранить данные<sup>5</sup>. Для манипуляций с данными используется язык Cypher, синтаксис которого достаточно близок к современным языкам разработки.

Как было отмечено ранее, представленная выше модель данных может быть описана ориентированным графом, что означает естественную возможность реализации в Neo4j. Для основных сущностей определены три метки: Модуль, Направление и Компетенция. Вершины разных меток могут хранить разный набор атрибутов.

Определены следующие связи между сущностями:

Связь	Описание
:ПРЕДШЕСТВУЕТ	Описывает зависимость между двумя модулями, когда один модуль необходим для освоения другого
:СОДЕРЖИТСЯ	Описывает зависимость между двумя модулями, когда один модуль является составной частью другого
:ИМЕЕТ_НАПРАВЛЕНИЕ	Связывает предмет и направление подготовки. Ребро также содержит свойства – первый и последний семестр преподавания предмета
:РАЗВИВАЕТ_КОМПЕТЕНЦИИ	Связывает предмет и компетенции. Описывает компетенции, которые являются результатом освоения предмета. Также связывает направление подготовки и компетенции. Описывает компетенции, которыми должен обладать выпускник данного направления

Для сравнительного анализа удобства использования реляционного и графового подходов в контексте разработки цифровых образовательных систем обратимся к конкретным примерам типовых запросов, которые могут быть использованы при работе с такими системами.

*Запрос 1. Получить все модули, которые непосредственно (длина связей = 1) предшествуют (связь “пререквизит”) модулю “Управление программными проектами”.*

<sup>5</sup> Neo4J Basics. URL: <https://neo4j.com/product/>.

Реализация на языке SQL:

```
SELECT m.name FROM modules m
  INNER JOIN modules_modules_preceed m2 ON m.id = m2.module_id
  INNER JOIN modules m3 ON m2.pre_module_id = m3.id WHERE m3.name='Управление про-
граммными проектами'
```

Реализация на языке Cypher:

```
MATCH (steps)-[rels:ПРЕДШЕСТВУЕТ]->(m:Module) WHERE m.name='Управление программными
проектами' RETURN m, steps;
```

*Запрос 2. Получить все модули, которые предшествуют (связь “пререквизит”) данному (за все время обучения).*

Для сокращения запроса считаем, что id модуля заранее известен и равен 0.

Реализация на языке SQL:

```
WITH RECURSIVE traverse AS (
  SELECT pre_module_id FROM ModuleModulePreceed
  WHERE module_id = 0
  UNION ALL
  SELECT ModuleModulePreceed.from_subject_id
  FROM ModuleModulePreceed
  INNER JOIN traverse
  ON ModuleModulePreceed.module_id =
    traverse.pre_module_id
)
SELECT distinct pre_module_id, s2.name FROM traverse INNER JOIN Module s2
  ON traverse.pre_module_id = s2.id;
```

Реализация на языке Cypher:

```
MATCH (steps)-[rels:ПРЕДШЕСТВУЕТ*]->(m:Module) WHERE m.id=0 RETURN DISTINCT steps;
```

Приведенные примеры показывают, что обе модели данных эмулируют обход графа. Однако в силу своей структуры реляционная модель для прохода по графу требует явного рекурсивного прохода с операцией соединения на каждом уровне, при этом осуществляется многократный проход по одному и тому же набору записей. Графовые БД, опираясь на представление графа, естественным образом осуществляют переход от вершины к ее смежным вершинам за один шаг. Обратим внимание, что запросы 1 и 2 практически не отличаются. Это связано с тем, что Cypher позволяет указать максимальную и минимальную глубины связей, которые надо пройти. В данном случае эти значения не указаны – возвращаются все предки.

Помимо выигрыша в удобстве проектирования запросов к базе данных и в количестве операций, Neo4J предоставляет удобный интерфейс визуализации данных в запросах, встраиваемый автоматически в каждую БД. Так, выполняя запрос на получение предмета с названием «Системное программное обеспечение», можно также получить информацию о связанных узлах (рис. 4).

Представленная визуализация при работе с графовой БД позволяет видеть данные в естественном для них представлении, что облегчает восприятие данных и позволяет быстрее проверять корректность запросов.

## Заключение

Проведенный анализ позволяет говорить о преимуществе в цифровых образовательных системах в ряде случаев графовых баз данных перед реляционными в контексте проектирования

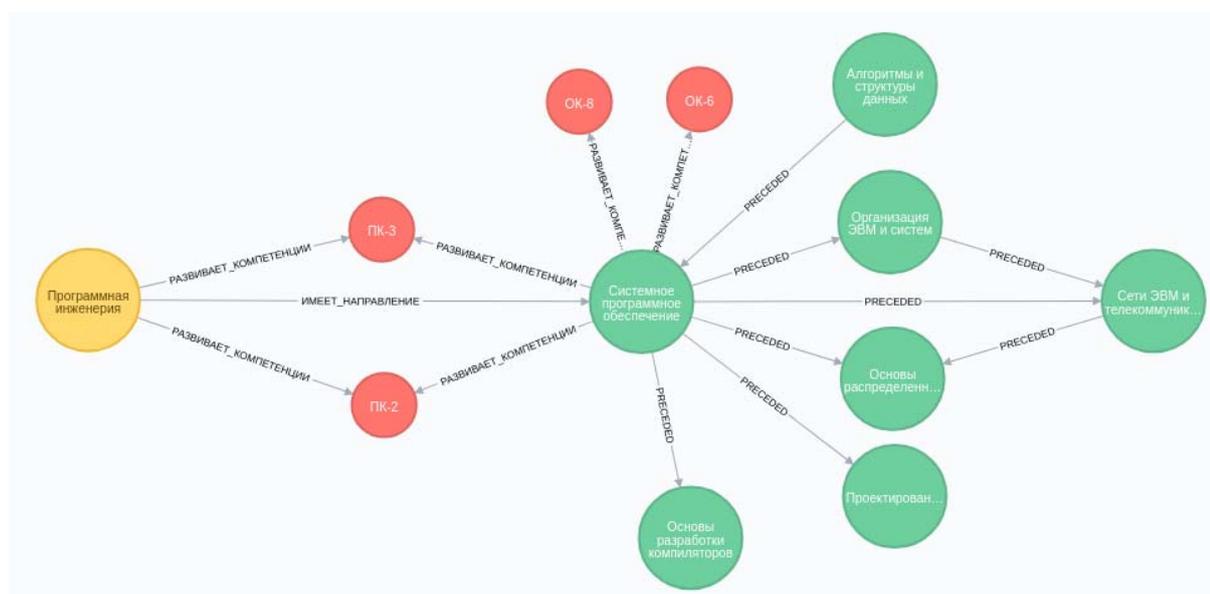


Рис. 4. Фрагмент структуры графовой базы данных

запросов, скорости их выполнения и удобства работы с самой базой данных. Возможно применение графовых БД с опорой на большое количество различных связей между объектами в образовательных системах, что позволяет рассматривать такие объекты как графы.

Еще раз подчеркнем, что графовая структура воспроизводима в реляционных базах данных, но для эффективной работы с «графовыми» данными в таких БД потребуются больше времени, более сложные запросы, а также, возможно, использование сторонних библиотек. Чем больше запрос использует графовую структуру, тем он становится сложнее и медленнее [7]. В дальнейших исследованиях предполагается продолжить анализ NoSQL-средств для работы с данными в цифровых образовательных средах, а также выполнить разработку специализированных расширений для графовых баз данных, позволяющих более эффективно разрабатывать соответствующие приложения в области образования.

### Список литературы

1. Дейт К. Дж. Введение в системы баз данных. М.: Вильямс, 2006. 92 с.
2. Cheng Yin Cheong. New Paradigm for Re-engineering Education: Globalization, Localization and Individualization. Springer, 2005. 26 с.
3. Абрамский М. М., Батырова Э. Ф., Марданова А. Р., Ахметзянова Т. А. Генерация индивидуальных образовательных траекторий и расписания обучения в парадигме индивидуализации образования // Электронные библиотеки. 2018. Т. 21 (3–4). С. 129–145.
4. Мизюков Г. С., Сергеева Е. А., Чернов А. В. Технологии графовых баз данных для информационной поддержки электронного портфолио студентов // Успехи современной науки и образования. 2016. Т. 3, № 8. С. 57–63.
5. Бартенев М. В., Вишняков И. Э. Использование графовых баз данных в целях оптимизации анализа биллинговой информации // Инженерный журнал: наука и инновации. 2013. Вып. 11. URL: <http://engjournal.ru/catalog/it/hidden/1058.html>
6. Antonios Makris. A classification of NoSQL data stores based on key design characteristics // Procedia Computer Science. 2016. No. 97. P. 94–103.
7. Vicknair C., Macias M., Zhao Zhendong, Nan Xiaofei, Chen Yixin, Wilkins D. A Comparison of a Graph Database and a Relational Database // Proc. of the 48<sup>th</sup> Annual Southeast Regional Conference. 2010. Article No. 42.

**М. М. Abramskiy, T. I. Timerkhanov**

*Kazan Federal University  
35 Kremlevskaya Str., Kazan, 420008, Russian Federation*

*ma@it.kfu.ru, timurgrunge@gmail.com*

## **COMPARATIVE ANALYSIS OF RELATIONAL AND GRAPH DATABASES USE IN DIGITAL EDUCATION SYSTEMS DEVELOPMENT**

Questions on database implementation model choice for development of digital education systems are discussed. The comparative analysis of relational and graph approaches for main entities data storing. Several advantages of graph model in case of development usability and query efficiency are shown.

*Keywords:* Graph Databases, Relational Databases, Neo4j, Digital Education Systems.

### **References**

1. Deit K. J. Vvedenie v systemy baz dannykh. Moscow, Viliams Publ., 2006, 92 p. (in Russ.)
2. Cheng Yin Cheong. New Paradigm for Re-engineering Education: Globalization, Localization and Individualization. Springer, 2005, 26 p.
3. Abramskiy M. M., Batyrova E. F., Mardanova A. R., Ahmetzyanova T. A. Generatsiya individual'nykh obrazovatel'nykh traektorij i raspisaniya obucheniya v paradigme individualizatsii obrazovaniya. *Elektronnye biblioteki*, 2018, Tvol. 21 (3–4), p. 129–145. (in Russ.)
4. Mizyukov G. S., Sergeeva E. A., Chernov A. V. Tekhnologii grafovnykh baz dannykh dlya informatsionnoy podderzhki elektronnoho portfolio studentov. *Uspekhi sovremennoy nauki i obrazovaniya*, 2016, vol. 3, no. 8, p. 57–63. (in Russ.)
5. Bartenev M. V., Vishnyakov I. E. Ispol'zovanie grafovnykh baz dannykh v tselyakh optimizatsii analiza billingovoy informatsii. *Inzhenernyj zhurnal: nauka i innovatsii*, 2013, no. 11. URL: <http://engjournal.ru/catalog/it/hidden/1058.html/> (in Russ.)
6. Antonios Makris. A classification of NoSQL data stores based on key design characteristics. *Procedia Computer Science*, 2016, no. 97, p. 94–103.
7. Vicknair C., Macias M., Zhao Zhendong, Nan Xiaofei, Chen Yixin, Wilkins D. A Comparison of a Graph Database and a Relational Database. *Proc. of the 48<sup>th</sup> Annual Southeast Regional Conference*, 2010, article no. 42.

*Received 11.09.2018*

*For citation:*

Abramskiy M. M., Timerkhanov T. I. Comparative Analysis of Relational and Graph Databases Use in Digital Education Systems Development. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 4, p. 5–12. (in Russ.) DOI 10.25205/1818-7900-2018-16-4-5-12

Д. А. Барамия<sup>1</sup>, Н. И. Горбенко<sup>2</sup>, М. М. Лаврентьев<sup>1,3</sup>

<sup>1</sup>Новосибирский государственный университет  
ул. Пирогова, 2, Новосибирск, 630090, Россия

<sup>2</sup>Институт вычислительной математики  
и математической геофизики СО РАН  
пр. Академика Лаврентьева, 6, Новосибирск, 630090, Россия

<sup>3</sup>Институт автоматики и электрометрии СО РАН  
пр. Академика Коптюга, 1, Новосибирск, 630090, Россия

*baramiyadenis@gmail.com, nikolay.gorbenko@gmail.com, mmlavrentiev@gmail.com*

## ОПИСАНИЕ ДОЛГОВРЕМЕННОЙ ЭВОЛЮЦИИ БЕРЕГОВОГО ПРОФИЛЯ НА ОСНОВЕ ДИФФУЗИОННОЙ МОДЕЛИ \*

Рассматривается диффузионная модель для описания долговременной эволюции профиля глубин в прибрежной зоне. Для калибровки модели решалась обратная задача восстановления коэффициентов по дополнительным граничным условиям путем минимизации функционала невязки. Многомерная оптимизация функционала проводилась методом дифференциальной эволюции. Для вычисления профиля глубин применялся метод конечных элементов. Тестирование проводилось на базе данных JARKUS при различных пространственных и временных интервалах. Достигнута относительная погрешность описания эволюции берегового профиля на уровне не более 5 %.

*Ключевые слова:* эволюция берегового профиля, обратная задача, функционал невязки, дифференциальная эволюция, метод конечных элементов.

### Введение

Возрастающее влияние деятельности человека на окружающую среду начинает приводить к заметным экологическим последствиям. Это относится, в частности, к инженерным сооружениям в прибрежной зоне, таким как дамбы, искусственные острова и каналы. Естественные процессы массопереноса приводят к постоянному изменению профиля глубин, что влечет за собой ощутимые последствия как для экологии прибрежной зоны, так и к необходимости дополнительных инженерных работ. Понимание механизмов, влияющих на изменение профиля глубин, а также возможность предсказания такого изменения позволяют минимизировать отрицательное влияние на экологию прибрежной зоны при проектировании и осуществлении инженерных работ.

---

\* Работа выполнена при финансовой поддержке проекта 0319-2018-0010 IV.36.1.4. «Исследование и развитие методов и технологий построения интегрированных программно-аппаратных комплексов для задач моделирования и управления динамическими системами обработки и отображения данных». Регистрационный номер: АААА-А17-117062110016-4.

В настоящей работе рассматривается программный инструментарий решения задачи калибровки диффузионной модели для описания долговременной эволюции берегового профиля.

Для моделирования процесса долговременной эволюции профиля глубин в прибрежной зоне, в работе [1] была предложена следующая диффузионная модель:

$$\frac{\partial(\delta X)}{\partial t} = D(z) \frac{\partial^2(\delta X)}{\partial z^2} + \Psi(t, z, \delta X, \frac{\partial(\delta X)}{\partial z}), \quad (1)$$

где  $\delta X(z, t)$  – изменение глубины на расстоянии  $z$  от береговой линии,  $D(z)$  – коэффициент диффузии,  $\Psi$  – некоторая функция, описывающая внешние источники. Данная модель применима только для песчаного дна. Согласно [1] наличие компоненты  $\Psi$  дает возможность ввести эффекты случайных воздействий, различных процессов переноса и вмешательства человека (таких как добыча песка и полезных ископаемых). Что касается общих подходов по изучению долговременной эволюции берегового профиля, см. работы [2–6].

Модель (1) была выбрана в качестве основы для дальнейших исследований, со следующими изменениями:

- функция изменения глубины  $\delta X(z, t)$  представлена как функция глубины  $u(z, t)$ ;
- функция, описывающая внешние источники  $\Psi$ , по аналогии с другими физическими процессами, была выбрана как функция переноса вещества  $C(z, t)u_z$ ;
- добавлена временная зависимость в коэффициент диффузии  $D(z, t)$ .

### Постановка задачи

В данной работе мы предполагаем, что долговременная эволюция профиля глубин в прибрежной зоне описывается линейным диффузионным уравнением:

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial z} (D(z, t) \frac{\partial u}{\partial z}) + C(z, t) \frac{\partial u}{\partial z}, \quad 0 \leq z \leq H, \quad t > 0. \quad (2)$$

Параметр  $H$  представляет собой «расстояние замыкания», т. е. расстояние от береговой линии, где моделируемые процессы диффузии и переноса вещества практически не имеют влияния на изменение профиля глубин. Для решения уравнения (2), в качестве начальных и граничных условий использовались данные измерений (например, с помощью сонаров [7]) профиля глубин в прибрежной зоне:

$$u(z, 0) = u^{\text{meas}}(z, 0), \quad u(0, t) = u^{\text{meas}}(0, t), \quad u(H, t) = u^{\text{meas}}(H, t). \quad (3)$$

Поскольку основная информация, необходимая для описания эволюции профиля глубин, содержится в коэффициентах  $D(z, t)$  и  $C(z, t)$ , задача состоит в подборе данных коэффициентов таким образом, чтобы уравнение (2) с начальными и граничными условиями (3) «оптимально» описывало измеренные данные, т. е. эволюцию профиля глубин. Термин «оптимально» означает, что разница между измеренными данными  $u^{\text{meas}}(z, t)$  и вычисленным профилем  $u^{\text{comp}}(z, t)$  должна быть возможно меньше.

### Калибровка модели

Калибровка модели (вычисление коэффициентов  $D(z, t)$  и  $C(z, t)$ ) проводилась в пространственно-временной области  $\Omega = \{0 \leq z \leq H, 0 \leq t \leq T\}$ . Для получения лучшего приближения коэффициентов решалась обратная задача восстановления коэффициентов уравнения

(2) по дополнительным граничным условиям (данным измерений). Задача решалась путем минимизации функционала невязки на каждом шаге по времени [8; 9]:

$$\Phi(D(z, t_i), C(z, t_i)) = \int_z \left| u^{\text{meas}}(z, t_i) - u^{\text{comp}}(z, t_i, D(z, t_i), C(z, t_i)) \right| dz. \quad (4)$$

Для минимизации функционала невязки (4) использовался метод дифференциальной эволюции, реализованный в библиотеке «Pallas Solver»<sup>1</sup>. Данный метод представляет собой стохастический алгоритм многомерной оптимизации, основанный на генетическом алгоритме. Общий принцип работы метода дифференциальной эволюции можно представить следующим образом (рис. 1).

1. Алгоритм генерирует начальную *популяцию*, состоящую из  $M$  случайных векторов  $n$ -мерного пространства, в котором определена целевая функция  $f(x)$ . Размерность популяции определяется эмпирически и задается в настройках алгоритма.

2. Выполняется воспроизводство *потомков* (набора векторов) текущего поколения с использованием стратегии *мутации* и *скрещивания*. Стратегия мутации генерирует новые мутантные вектора путем стохастического комбинирования векторов предыдущего поколения (*базовых векторов*). Стратегия скрещивания замещает координаты мутантных векторов на соответствующие координаты базовых векторов с некоторой вероятностью  $p$ . Значение  $p$  задается в настройках алгоритма.

3. Выполняется этап *селекции*. Если после скрещивания полученный *пробный вектор* оказывается лучше базового вектора (значение целевой функции улучшилось), то в новом поколении базовый вектор заменяется пробным.

4. На каждой итерации алгоритма (или с заданной периодичностью) определяется лучший вектор поколения для контроля скорости поиска оптимального решения. Помимо этого, на данном этапе проверяются условия остановки алгоритма, такие как достижение предельного количества итераций, нахождение глобального или субоптимального решения, неизменность лучшего вектора поколения на протяжении заданного предельного количества итераций.

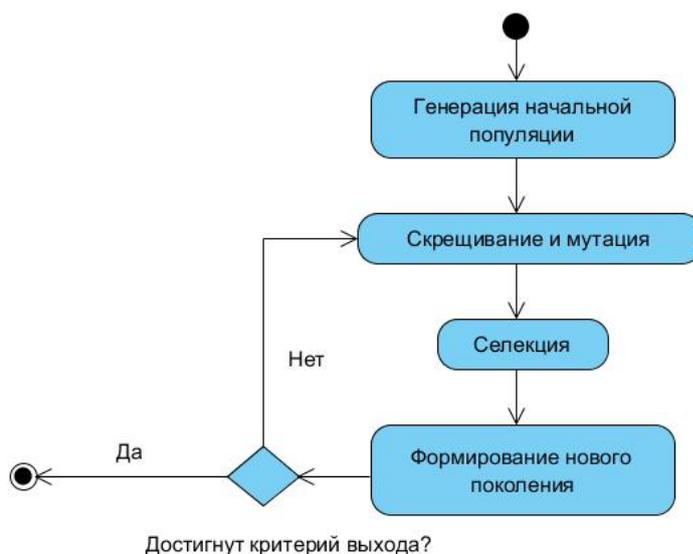


Рис. 1. Общая схема работы алгоритма дифференциальной эволюции

<sup>1</sup> <http://lattice.github.io/pallas-solver/> – библиотека «Pallas Solver» с открытым исходным кодом, содержащая набор алгоритмов глобальной оптимизации.

Преимущества метода дифференциальной эволюции заключаются в следующем: возможность применения к недифференцируемым, нелинейным, мультимодальным функциям от многих переменных; отсутствие требования начального приближения для искомого решения, что позволяет рассматривать решения разных порядков на области определения целевой функции; метод не пытается поддерживать одно текущее решение, а развивает популяцию кандидатов решений.

Для решения прямой задачи, а именно, вычисления профиля  $u^{\text{comp}}(z, t)$  в области  $z$  при заданных значениях  $t_i$ ,  $D(z, t_i)$ ,  $C(z, t_i)$  был реализован метод конечных элементов [10]. В качестве базисных функций для  $u(z, t)$  и коэффициентов  $D(z, t_i)$ ,  $C(z, t_i)$  использовались полиномы первой степени по  $z$  на каждом интервале по времени. Другими словами, функции рассматривались как кусочно-линейные. Отладка и тестирование разработанного метода конечных элементов проводилась с использованием аналитически заданных функций по следующей схеме.

1. Аналитически задаем функцию  $v(z, t)$  и коэффициенты  $D(z, t_i)$ ,  $C(z, t_i)$ .
2. Определяем функцию  $\varphi(z, t)$  вида

$$\varphi(z, t) = \frac{\partial v}{\partial t} - \frac{\partial}{\partial z} \left( D(z, t) \frac{\partial v}{\partial z} \right) - C(z, t) \frac{\partial v}{\partial z}. \quad (5)$$

3. Подставляем функцию (5) в исходное уравнение (2) в качестве правой части:

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial z} \left( D(z, t) \frac{\partial u}{\partial z} \right) + C(z, t) \frac{\partial u}{\partial z} + \varphi(z, t). \quad (6)$$

4. Используя разработанный метод конечных элементов, вычисляем численное решение уравнения (6) и сравниваем с аналитическим решением  $v(z, t)$ .

Результаты тестирования показали, что относительная погрешность меньше 0,01 % при различных аналитически заданных функциях  $v(z, t)$ .

### Тестирование диффузионной модели

Тестирование проводилось на многолетних данных измерений профиля глубин у побережья Голландии, которые собираются в рамках проекта JARKUS<sup>2</sup>. Поскольку предложенная модель зависит только от одной пространственной переменной, задача рассматривается как одномерная относительно пространственных переменных. Поэтому для тестирования была выбрана одна точка на побережье (рис. 2), относительно которой и решалась задача калибровки модели.

Набор тестовых данных состоит из массива  $201 \times 34$  точек (рис. 3): 201 точка наблюдений по расстоянию от берега с шагом 3 метра и 34 точки по времени с шагом 1 год. Калибровка модели выполнялась на временных интервалах длиной в 10 лет.

Точность описания измерений берегового профиля предложенной моделью оценивалась путем вычисления относительной ошибки (7) между измеренными данными  $u^{\text{meas}}(z, t)$  и вычисленными данными  $u^{\text{comp}}(z, t)$ :

$$\delta(z, t) = 100\% \frac{\int_{z_1}^{z_2} \int_{t_1}^{t_2} |u^{\text{meas}}(z, t) - u^{\text{comp}}(z, t)| dz dt}{\int_{z_1}^{z_2} \int_{t_1}^{t_2} |u^{\text{meas}}(z, t)| dz dt}. \quad (7)$$

<sup>2</sup> <https://publicwiki.deltares.nl/display/OET/Dataset+documentation+JarKus>



Рис. 2. Остров Амеланд в Голландии

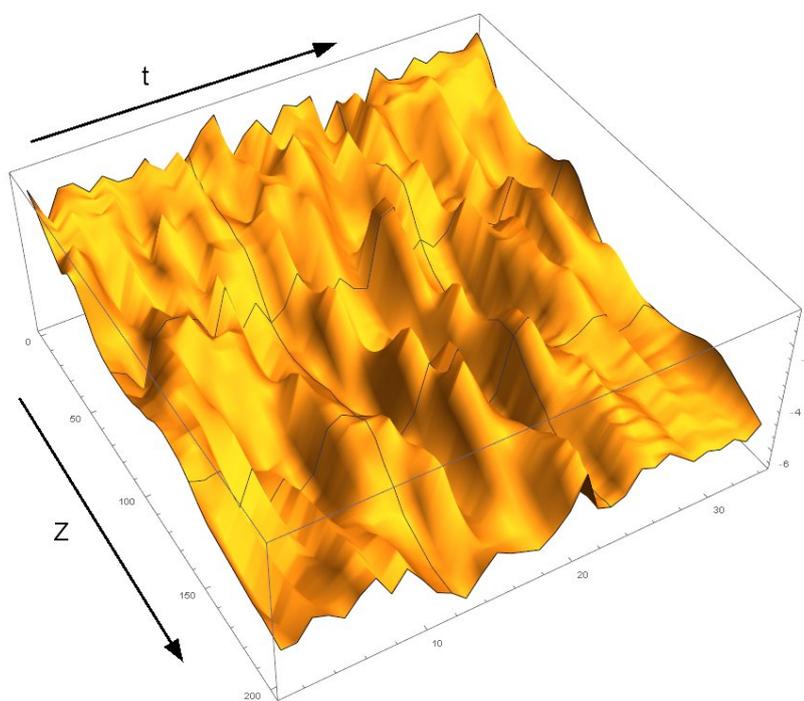


Рис. 3. Визуальное представление голландских данных эволюции берегового профиля

Относительная погрешность описания эволюции берегового профиля глубин на разных временных и пространственных интервалах, %

	$0 \leq t \leq 10$	$5 \leq t \leq 15$	$10 \leq t \leq 20$	$15 \leq t \leq 25$
$0 \leq z \leq H$	3,17	2,9	3,62	3,85
$H/10 \leq z \leq H$	3,01	2,23	2,53	2,71
$H/4 \leq z \leq H$	2,56	1,44	1,82	2,33

При тестировании использовались также разные пространственные интервалы. Это проводилось из соображений, что вблизи берега, помимо диффузии и переноса, существуют другие процессы, которые влияют на данные измерений. Из результатов, представленных в таблице,

видно, что при смещении пространственного интервала от берега относительная погрешность уменьшается. Также следует отметить, что при смещении окна по времени относительная погрешность вычисленных профилей сохраняется в пределах 5 %.

### Заключение

Для моделирования процесса долговременной эволюции профиля глубин в прибрежной зоне использовалась диффузионная модель (2). Решалась задача подбора коэффициентов этой модели, которые обеспечивают соответствие решения данным измерений. Для получения лучшего приближения коэффициентов решалась обратная задача путем минимизации функционала невязки (4). Модель тестировалась на наборе данных JARKUS для различных пространственных и временных интервалов измерений. Из приведенных результатов видно, что предложенная диффузионная модель способна описать долговременную эволюцию профиля глубин с относительной погрешностью на уровне не более 5 %. В дальнейшем будет исследоваться возможность предсказания долговременной эволюции профиля глубин путем экстраполяции найденных коэффициентов по времени.

### Список литературы

1. *De Vriend H. I., Capobianco M., Chesher T., Swart Th. E. de, Latteux B., Stive M. J. F.* Approaches to long-term modelling of coastal morphology: a Review // *Coastal Engineering*. Amsterdam: Elsevier Science Publ., 1993. Vol. 21. P. 225–269.
2. *Bakker W. T.* The dynamics of a coast with a groyne system // *Proc. of the 11<sup>th</sup> Coastal Engineering Conference*. American Society of Civil Engineers (ASCE). 1968. P. 492–517.
3. *Capobianco M.* A procedure for parameter identification of partial differential equations of parabolic type // *G6-M Workshop “System Dynamics”, Delft Hydraulics*. 1992.
4. *Dean R. G.* Equilibrium beach profiles: characteristics and applications // *J. Coastal Res.* 1991. Vol. 7. No. 1. P. 53–84.
5. *Hanson H.* GENESIS, a generalized shoreline change numerical model for engineering use // *Univ. of Lund, Dept. of Water Res. Eng.* 1987. Report Vol. 1007. 206 p.
6. *Larson M., Hanson H., Kraus N. C.* Analytical solutions of the one-line model of shoreline change // *Report to U.S. Army Corps of Engineers, Coastal Engineering Research Center*. 1987. 6 p.
7. *Beavers R., Howd P., Birkemeier W., Hathaway K.* Evaluating profile data and depth of closure with sonar altimetry // *Proc. Hauppauge Coastal Sediments*. 1999. Vol. 1. P. 479–490.
8. *Avdeev A. V., Lavrentiev M. M., Priimenko V. I.* Inverse Problems and Some Applications. Novosibirsk, 1999. 342 p.
9. *Lavrentiev M. M., Romanov V. G., Shishatskii Y.* Ill-Posed Problems of Mathematical Physics and Analysis // *AMS Translations of Math. Monographs*. 2017. 290 p.
10. *Larson M. G., Bengzon F.* The Finite Element Method: Theory, Implementation, and Applications. Springer, 2013. 395 p.

Материал поступил в редколлегию 10.10.2018

**Д. А. Барамия<sup>1</sup>, Н. И. Горбенко<sup>2</sup>, М. М. Лаврентьев<sup>1,3</sup>**

<sup>1</sup> *Novosibirsk State University  
1 Pirogov Str., Novosibirsk, 630090, Russian Federation*

<sup>2</sup> *Institute of Computational Mathematics and Mathematical Geophysics SB RAS  
6 Academician Lavrentiev Ave., Novosibirsk, 630090, Russian Federation*

<sup>3</sup> *Institute of Automation and Electrometry SB RAS  
1 Academician Koptuyug Ave., Novosibirsk, 630090, Russian Federation*

*baramiyadenis@gmail.com, nikolay.gorbenko@gmail.com, mmlavrentiev@gmail.com*

### DESCRIPTION OF COASTAL PROFILE LONG TERM EVOLUTION BY DIFFUSION MODEL

Diffusion model is considered to describe coastal depth profile. For model calibration an inverse problem for equation coefficient reconstruction was solved by minimization of a certain misfit func-

tional using an extra boundary conditions. Multy-dimensional optimization was arranged with the help of differential evolution method. Finite elements method was used to calculate model depth profile. The JARKUS dataset was used for testing at different temporal and spatial intervals. Relative error at the level of 5 % was achieved.

*Keywords:* coastal profile evolution, inverse problem, cost functional, differential evolution, finite element method.

## References

1. De Vriend H. I., Capobianco M., Chesher T., Swart Th. E. de, Latteux B., Stive M. J. F. Approaches to long-term modelling of coastal morphology: a Review. *Coastal Engineering*. Amsterdam: Elsevier Science Publ., 1993, vol. 21, p. 225–269.
2. Bakker W. T. The dynamics of a coast with a groyne system. *Proc. of the 11<sup>th</sup> Coastal Engineering Conference. American Society of Civil Engineers (ASCE)*, 1968, p. 492–517.
3. Capobianco M. A procedure for parameter identification of partial differential equations of parabolic type. *G6-M Workshop "System Dynamics", Delft Hydraulics*. 1992.
4. Dean R. G. Equilibrium beach profiles: characteristics and applications. *J. Coastal Res.*, 1991, vol. 7, no. 1, p. 53–84.
5. Hanson H. GENESIS, a generalized shoreline change numerical model for engineering use. *Univ. of Lund, Dept. of Water Res. Eng.*, 1987, vol. 1007, 206 p.
6. Larson M., Hanson H., Kraus N. C. Analytical solutions of the one-line model of shoreline change. *Report to U.S. Army Corps of Engineers, Coastal Engineering Research Center*, 1987, 6 p.
7. Beavers R., Howd P., Birkemeier W., Hathaway K. Evaluating profile data and depth of closure with sonar altimetry. *Proc. Hauppauge Coastal Sediments*, 1999, vol. 1, p. 479–490.
8. Avdeev A. V., Lavrentiev M. M., Priimenko V. I. Inverse Problems and Some Applications. Novosibirsk, 1999, 342 p.
9. Lavrentiev M. M., Romanov V. G., Shishatskii Y. Ill-Posed Problems of Mathematical Physics and Analysis. *AMS Translations of Math. Monographs*, 2017, 290 p.
10. Larson M. G., Bengzon F. The Finite Element Method: Theory, Implementation, and Applications. Springer, 2013. 395 p.

Received 10.10.2018

*For citation:*

Baramiya D. A., Gorbenko N. I., Lavrentiev M. M. Description of Coastal Profile Long Term Evolution by Diffusion Model. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 4, p. 13–19. (in Russ.) DOI 10.25205/1818-7900-2018-16-4-13-19

**К. В. Баталин, Н. С. Мамеев, К. Ю. Попова, И. Д. Рыжаков, Г. Э. Яхьяева**

*Новосибирский государственный университет  
ул. Пирогова, 2, Новосибирск, 630090, Россия*

*kir55rus@yandex.ru, asm\_edf@mail.ru, kr111kr@yandex.ru  
igorusha4@yandex.ru, gul\_nara@mail.ru*

## **ПРОГРАММНАЯ СИСТЕМА УПРАВЛЕНИЯ ОБРАЗОВАТЕЛЬНЫМ ПРОЦЕССОМ ИТОС**

Статья посвящена описанию программной системы ИТОС (индивидуальных траекторий обучения студентов), предназначенной для автоматизации работы с основной образовательной программой (ООП) и соответствующим ей учебным планом. В системе реализована возможность отслеживать логику следования учебных курсов, что позволяет на этапе составления ООП выявлять необходимость перестановки или слияния дисциплин, а также достижимость всех заданных компетенций. Имеется возможность устанавливать взаимосвязь между ООП и различными профессиональными стандартами. В системе реализован алгоритм построения и оценки всевозможных индивидуальных траекторий обучения студента в рамках данной ООП, а также модуль рекомендательной системы, помогающей пользователям (студентам) в выборе индивидуальных траекторий обучения.

*Ключевые слова:* основная образовательная программа, профессиональный стандарт, индивидуальная траектория обучения, рекомендательная система.

### **Введение**

Научно-технический прогресс, развитие производств и технологий, а также изменяющийся рынок труда требуют постоянного развития профессиональных навыков и компетенций работника. Именно поэтому в настоящее время в Российской Федерации производится пересмотр квалификационных справочников и вводятся новые профессиональные стандарты [1].

Профессиональный стандарт – это характеристика квалификации, необходимой работнику для осуществления определенного вида профессиональной деятельности (ст. 195.1 ТК РФ)<sup>1</sup>. Для работодателей профессиональный стандарт будет являться основой для установления более конкретных требований при выполнении трудовой функции работника с учетом специфики деятельности организации.

В соответствии с Федеральным законом «Об образовании в Российской Федерации»<sup>2</sup> (ст. 11, п. 7) с целью наиболее полного и качественного удовлетворения потребности рынка труда федеральные государственные образовательные стандарты должны быть сопряжены с соответствующими профессиональными стандартами. Это позволит обеспечить подготовку высококвалифицированных кадров, востребованных на рынке труда и способных качественно выполнять свои функциональные обязанности в соответствии с действующими профессиональными стандартами [2].

---

<sup>1</sup> <http://www.tkodeksrf.ru/>

<sup>2</sup> <https://fzakon.ru/laws/federalnyy-zakon-ot-29.12.2012-n-273-fz/>

В связи с этим положения соответствующих профессиональных стандартов должны учитываться при формировании федеральных государственных образовательных стандартов высшего образования (ФГОС ВО). Практическая реализация ФГОС ВО ставит основной целью высших учебных заведений формирование личной профессиональной компетентности выпускников [3]. Принципиальная задача ФГОС новых поколений – при массовости образования сделать его индивидуальным, отвечающим требованиям различных профессиональных стандартов.

В содержании образовательных стандартов подчеркиваются права обучающихся на формирование своей индивидуальной образовательной программы, на получение консультаций в вузе по выбору дисциплин (модулей) предусмотренных ООП. В связи с этим обязанностью вуза становится обеспечение обучающимся реальной возможности участвовать в формировании своей программы обучения.

Однако на сегодняшний день составление индивидуальных образовательных программ в вузовской практике сталкивается с определенными трудностями. Во-первых, принятие студентом ответственности за собственное образование, осознание его цели и выбора профессионального стандарта до начала обучения не возможно ввиду постоянных изменений потребностей и индивидуальных возможностей личности. Во-вторых, проектирование такой программы не может учесть индивидуальный темп обучения студента. Таким образом, индивидуальная образовательная программа студента должна рассматриваться как динамическая сущность, которая может меняться и корректироваться в ходе образовательного процесса. Но каждая индивидуальная образовательная программа должна гарантировать, что студент, успешно окончивший вуз, на выходе будет обладать всеми компетенциями, предусмотренными тем или иным профессиональным стандартом.

Таким образом, на сегодняшний день становится актуальной проблема разработки модели, которая будет учитывать вариативность и индивидуализацию высшего образования. В данной модели ФГОС и ООП будут определять лишь исходные и конечные точки для образования студентов, на основе которых можно реализовать значительное количество различных индивидуальных траекторий обучения студента.

В данной статье описывается программная система ИТОС (индивидуальных траекторий обучения студентов), учитывающая индивидуализацию и вариативность образовательного процесса в вузе.

### **Онтологическая модель и база знаний**

В связи с постоянным развитием науки и техники структура и содержание учебных дисциплин, а также требования работодателей постоянно меняются. Следовательно, разрабатываемая программная система должна уметь работать с постоянно пополняемыми базами учебных дисциплин и профессиональных стандартов. Применение методологии онтологического моделирования [4–7] позволяет быстро адаптировать ООП факультета под актуальные профессиональные стандарты.

Основными документами, рассматриваемыми в модели процесса формирования ООП, являются:

- Федеральный государственный образовательный стандарт (ФГОС) – документ, выдвигающий требования к определённому уровню образования или к направлению подготовки;
- Основная образовательная программа (ООП) – набор документов, описывающих структуру образовательного процесса: документ, в котором определены общие сведения об образовательной программе, такие как цель (миссия) программы, объем и сроки освоения, требования к абитуриенту и возможные виды профессиональной деятельности выпускника; учебный план; рабочие программы дисциплин; календарный учебный график;
- Набор профессиональных стандартов – документы, описывающий требования к знаниям и умениям для выполнения определённого вида профессиональной деятельности, обучение которым планируется в ходе реализации ООП.

Для описания связей между данными документами была построена онтологическая модель (рис. 1).

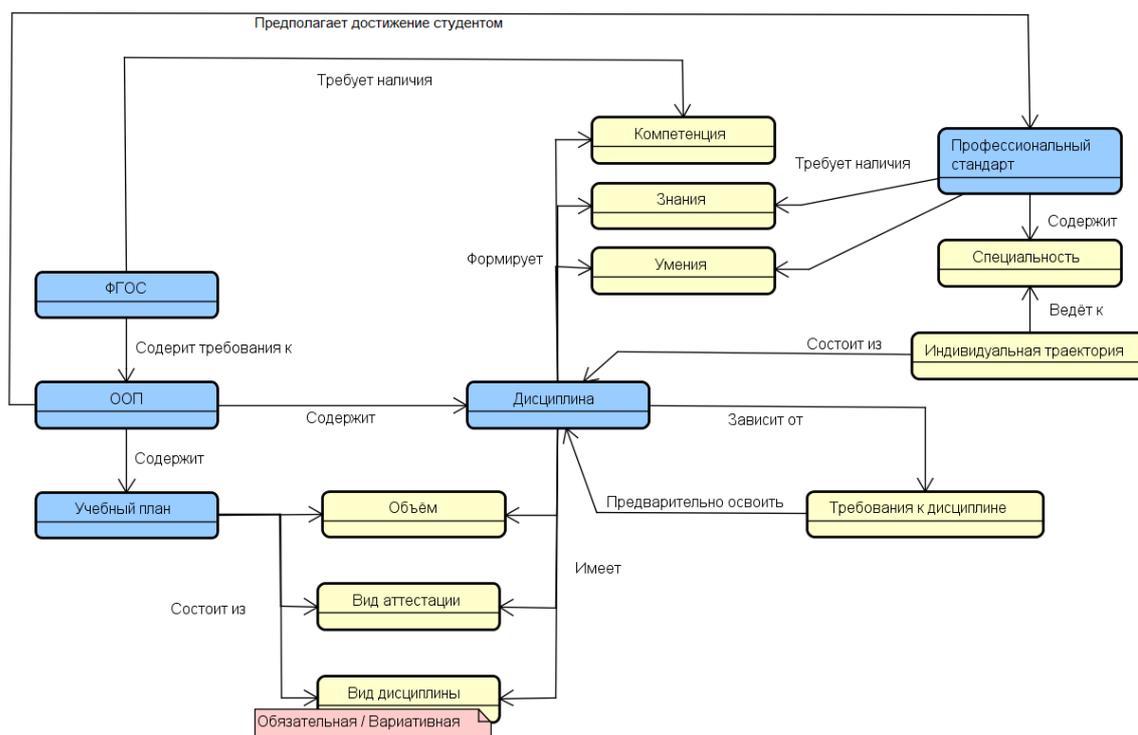


Рис. 1. Онтологическая модель системы ИТОС

В соответствии с данной онтологической моделью разработана база данных, позволяющая хранить все описанные множества документов. Эта база разделена на три функциональных модуля: база нормативных документов, база дисциплин и база профессиональных стандартов.

В базе нормативных документов системы ИТОС хранятся общие требования, взятые из ФГОС и ООП, такие как перечень обязательных дисциплин, максимальный объем образовательной программы, сформированные компетенции и т. п. [8].

База дисциплин системы ИТОС формируется на основе учебного плана и рабочих программ дисциплин. Для каждой дисциплины формируется экземпляр данного документа, содержащий название, аннотацию, формируемые знания и умения, формируемые компетенции, зависимости от других дисциплин, структуру и содержание. В ходе работы с системой имеется возможность редактирования базы дисциплин, т. е. добавление / удаление дисциплин, изменение зависимостей между дисциплинами и т. д.

Одним из ключевых направлений деятельности структурных подразделений вузов, отвечающих за разработку и реализацию образовательных программ, является соотнесение профессиональных и образовательных стандартов, обеспечивающее выполнение требований профессионального стандарта в результате реализации образовательного [9].

Профессиональный стандарт состоит из четырех частей: 1) общие сведения; 2) описание трудовых функций, входящих в профессиональный стандарт (функциональная карта вида профессиональной деятельности); 3) характеристика обобщенных трудовых функций; 4) сведения об организациях – разработчиках профессионального стандарта.

Наибольший интерес для данной работы представляет третья часть, где можно сопоставить трудовые функции и необходимые для них знания и умения. В связи с этим была построена онтология (рис. 2) профессиональных стандартов [10].

База профессиональных стандартов системы ИТОС формируется из реестра<sup>3</sup> профессиональных стандартов с сайта Министерства труда и социальной защиты Российской Федера-

<sup>3</sup> <http://profstandart.rosmintrud.ru/obshchiy-informatsionnyy-blok/natsionalny-reestr-professionalnykh-standartov/reestr-professionalnykh-standartov/>

ции, где они хранятся в виде xml-файла. Программная система ИТОС имеет возможность принимать на вход данный файл, парсить его и переносить необходимую информацию в базу данных в соответствии с составленной онтологической моделью.

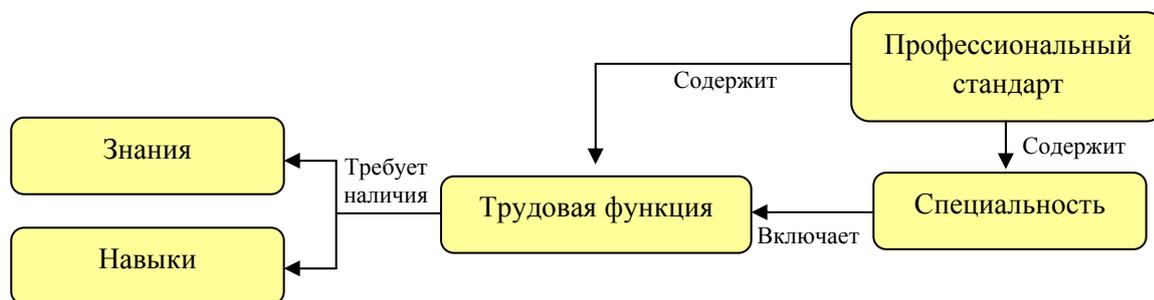


Рис. 2. Онтология профессионального стандарта

В системе ИТОС реализована возможность поиска с подсказками по названию профессионального стандарта, специальности, набору трудовых функций, знаниям, умениям. Также в программе реализован функционал контроля введения профессиональных стандартов, что подразумевает просмотр развиваемых дисциплинами умений и навыков и выяснением, покрывают ли они требования профессионального стандарта.

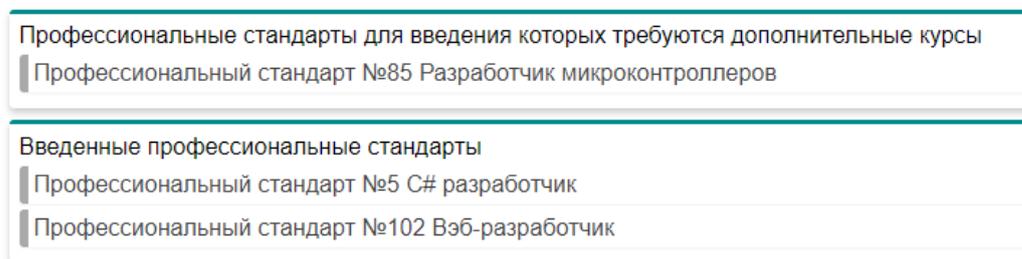


Рис. 3. Статус профессиональных стандартов в модуле

Программа показывает введенные в данный момент профессиональные стандарты и профессиональные стандарты, для введения которых нужны дополнительные курсы (рис. 3).

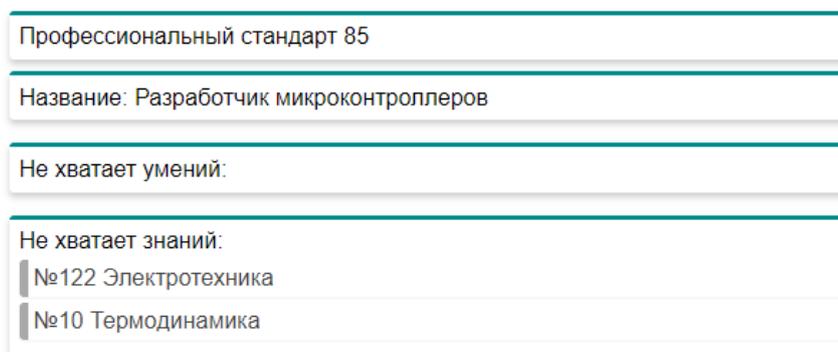


Рис. 4. Необходимые для введения профессиональных стандартов знания и умения

Также при нажатии на не введенный профессиональный стандарт отображаются необходимые для его введения знания и умения (рис. 4).

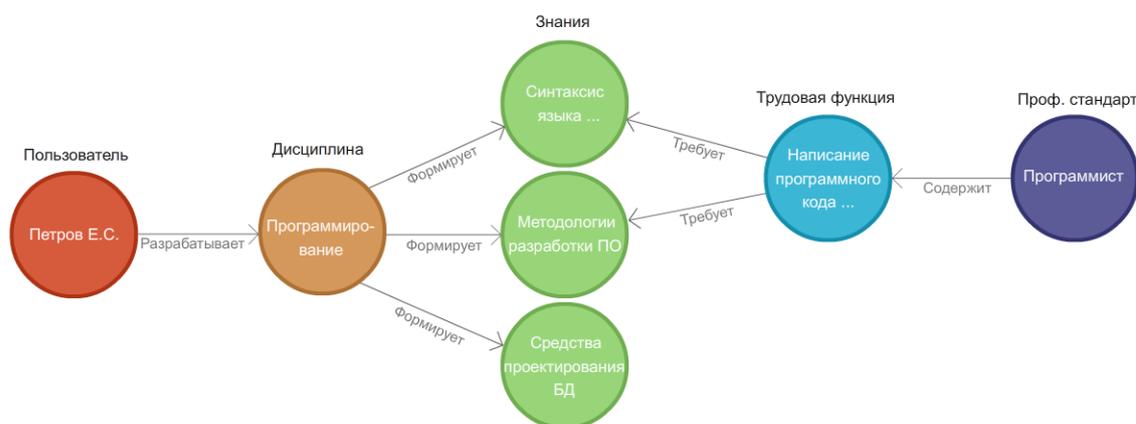


Рис. 5. Пример представления данных в графовой БД

Для хранения данных используется графовая база данных neo4j<sup>4</sup>: вершинами графа являются пользователи, все виды документов и вспомогательные сущности, такие как знания, умения, роли и др. Рёбра описывают отношения принадлежности или зависимости элементов (рис. 5).

### Алгоритм проверки внутренней и внешней согласованности основной образовательной программы

Ввиду большого числа учебных дисциплин, зависимостей в порядке их изучения, а также направленности дисциплин на развитие различных профессиональных навыков появляется потребность контролировать корректность составленной образовательной программы.

Руководство факультета составляет общий вид образовательной программы и задаёт множество профессиональных стандартов, которые должны быть достигнуты. После этого преподаватели составляют учебные планы для своих дисциплин, описывая зависимости между дисциплинами и развиваемые профессиональные навыки и компетенции. В процессе формирования образовательной программы система ИТОС отслеживает необходимость изменения списка дисциплин или порядка их следования и достижимость всех заданных профессиональных стандартов, тем самым проверяет внутреннюю согласованность образовательной программы.

Кроме того, система ИТОС проверяет образовательную программу на предмет соответствия актуальному ФГОСу. Таким образом, в системе ИТОС реализована возможность в автоматическом режиме выявлять ошибки и неточности в структуре образовательной программы ещё на этапе её составления.

### Соответствие ООП требованиям ФГОС

- Пункт 2.1. ФГОС содержит требования к объёму программы бакалавриата. Для проверки данного требования вычисляется суммарный объём образовательной программы, складывающийся из суммы объёмов базовых дисциплин и суммы объёмов блоков вариативных дисциплин. После чего определяется соответствие вычисленного объёма требованиям.
- Пункты 2.2. и 2.3. ФГОС содержат требования к обязательному наличию заданного множества дисциплин. Для проверки данного требования производится поиск всех требуемых дисциплин среди составленных рабочих программ дисциплин.

<sup>4</sup> <https://neo4j.com>

• Пункт 2.9. ФГОС содержит требования к процентному соотношению базовой и вариативной частей образовательной программы. Для проверки данного требования вычисляется отдельно суммарный объём всех базовых дисциплин и отдельно суммарный объём всех блоков, после чего определяется соответствие соотношения данных объёмов требованиям.

### Проверка внутренней согласованности ООП

Представим множество дисциплин в виде ориентированного графа  $G = \langle \mathbb{V}, \mathbb{E} \rangle$ , где  $\mathbb{V}$  – множество дисциплин. Бинарное отношение  $\mathbb{E} \subseteq \mathbb{V}^2$  задается следующим образом:  $(A, B) \in \mathbb{E}$ , если необходимо изучить дисциплину  $B$  перед изучением дисциплины  $A$ .

На каждую вершину  $A$  графа  $G$  накладываем метку

$$\langle m_1(A); m_2(A) \rangle,$$

где  $m_1(A)$  – номер семестра, с которого начинается изучение дисциплины  $A$ , и  $m_2(A)$  либо равна 0, если дисциплина  $A$  базовая, либо равна номеру блока вариативных дисциплин, которому принадлежит дисциплина  $A$ .

Для каждой дисциплины  $A \in \mathbb{V}$  определим множество

$$S(A) = \{B \in \mathbb{V} \mid (A, B) \in \mathbb{E}\}$$

дисциплин, знание которых необходимо для изучения дисциплины  $A$ .

Будем считать, что образовательная программа обладает внутренней согласованностью, если:

- 1) для любой вершины  $A \in \mathbb{V}$  и для любой вершины  $B \in S(A)$  выполняется условие  $m_1(B) < m_1(A)$ ;
- 2) для любой вершины  $A \in \mathbb{V}$  и для любой вершины  $B \in S(A)$  выполняется условие: если  $m_2(A) = 0$ , то  $m_2(B) = 0$ ;
- 3) для любой вершины  $A \in \mathbb{V}$  и для любых вершин  $B_1, B_2 \in S(A)$  выполняется условие: если  $m_2(B_1) > 0$ , то  $m_2(B_1) \neq m_2(B_2)$ .

Алгоритм проверки внутренней согласованности заключается в упорядочивании множества вершин согласно лексическому упорядочиванию меток и последующему обходу графа согласно данному упорядочиванию с проверкой в каждой вершине всех трех условий.

### Алгоритм построения индивидуальных траекторий

В последнее время образовательные стандарты все больше акцентируют внимание на том, что у студентов должна быть возможность корректировки процесса своего обучения. У учебных заведений появляется обязанность предоставить такую возможность. Корректировка отображается на учебной траектории студента. Индивидуальная учебная траектория должна быть корректной: после прохождения всех курсов траектории студент обязан обладать полным набором необходимых умений и знаний, чтобы получить выбранную специальность.

В программной системе ИТОС реализован алгоритм построения всевозможных корректных индивидуальных траекторий в рамках заданной модели основной образовательной программы [11].

Основная образовательная программа определяет набор курсов, которые являются базовыми. Эти курсы обязательно должны присутствовать во всех итоговых траекториях обучения. Также основная образовательная программа определяет блоки вариативных курсов. Студент вправе выбирать какой-либо курс из каждого блока. Следовательно, перед алгоритмом поставлена задача перебрать все варианты курсов из каждого блока, сконструировать

траектории на основе перебора, проверить каждую траекторию на корректность и выдать в качестве результата только те траектории, которые прошли все проверки.

Под корректностью в данном случае подразумевается следующий набор правил.

1. Траектория должна охватывать такой набор курсов, чтобы после завершения обучения по данной траектории студент обладал всеми необходимыми знаниями и умениями для получения какой-либо специальности. Другими словами, не должно получиться так, что студент закончил все курсы из сгенерированной траектории, но при этом не смог получить желаемую специальность.

2. Траектория должна содержать все зависимости своих внутренних курсов. Иначе говоря, если в траектории есть курсы, зависящие от внешних курсов, которых, в свою очередь, нет в сгенерированной учебной траектории, то данную траекторию следует исключить из результата работы алгоритма.

Если во время работы алгоритма были обнаружены траектории, которые не удовлетворяют заданным критериям, алгоритм исключает их из результата работы и оповещает о случившемся пользователя системы.

Полный перебор всех курсов из каждого блока может потребовать больших временных и мощностных ресурсов. Отсюда появляется еще одна деталь реализации алгоритма: программный модуль, содержащий алгоритм генерации всевозможных траекторий обучения, не должен работать отдельно каждый раз для каждого запроса пользователя.

Результат работы алгоритма зависит от основной образовательной программы. Поэтому запускать алгоритм можно только после ее изменения, а результат работы сохранять в базе данных. После этого пользователи должны взаимодействовать не с самим алгоритмом для получения траекторий обучения, а с базой данных, которая хранит результат работы – траектории обучения.

### **Модуль рекомендательной системы**

Множество вариативных дисциплин, реализуемых в рамках данной ООП, может быть достаточно велико (порядка 50–70 дисциплин). Кроме того, оно может меняться из года в год. Следовательно, множество всевозможных траекторий обучения студента также достаточно велико и непостоянно. В такой ситуации студенту сложно самостоятельно анализировать происходящие изменения, потому необходима автоматизация данного процесса. Для того чтобы облегчить студентам выбор траекторий, в системе ИТОС реализован модуль рекомендательной системы, позволяющий выбирать траектории исходя из предпочтений и возможностей каждого конкретного студента и исходя из текущего учебного плана и всех программ курсов [12].

Взаимодействие рекомендательной системы с пользователем происходит в режиме диалога [13]. Входные параметры, по которым система помогает пользователю выбирать индивидуальные траектории следующие.

- **Выбор профессиональных стандартов.** Студент имеет возможность определить список специальностей, одну или несколько из которых он хотел бы получить по завершении обучения, а также список специальностей, в получении которых он не заинтересован.

- **Выбор дисциплин.** Студент имеет возможность определить список дисциплин, которые он предпочитает посещать, а также список дисциплин, в посещении которых он не заинтересован.

- **Освоенные дисциплины.** На момент выбора или смены индивидуальной траектории студент (старших курсов) уже обладает множеством освоенных дисциплин. Система корректирует список рекомендованных траекторий согласно этой информации.

Рекомендательная система предоставляет пользователю список возможных траекторий, подходящих под заданные условия. Кроме того, возможен случай, когда траекторий, удовлетворяющих заданным условиям, не существует. Тогда система рекомендует пользователю траектории, удовлетворяющие части требований. Пользователь, в свою очередь, может пересмотреть критерии поиска.

Также не имеет смысла выдавать пользователю все возможные траектории, в случае, если таких нашлось много. Это вновь приведёт к затруднению выбора из-за большого объема информации. Система упорядочивает траектории в порядке убывания множества требований, которым они соответствуют, и выдает пользователю первые  $n$  траекторий. При этом требования относительно профессиональных стандартов считаются более приоритетными, нежели требования относительно дисциплин.

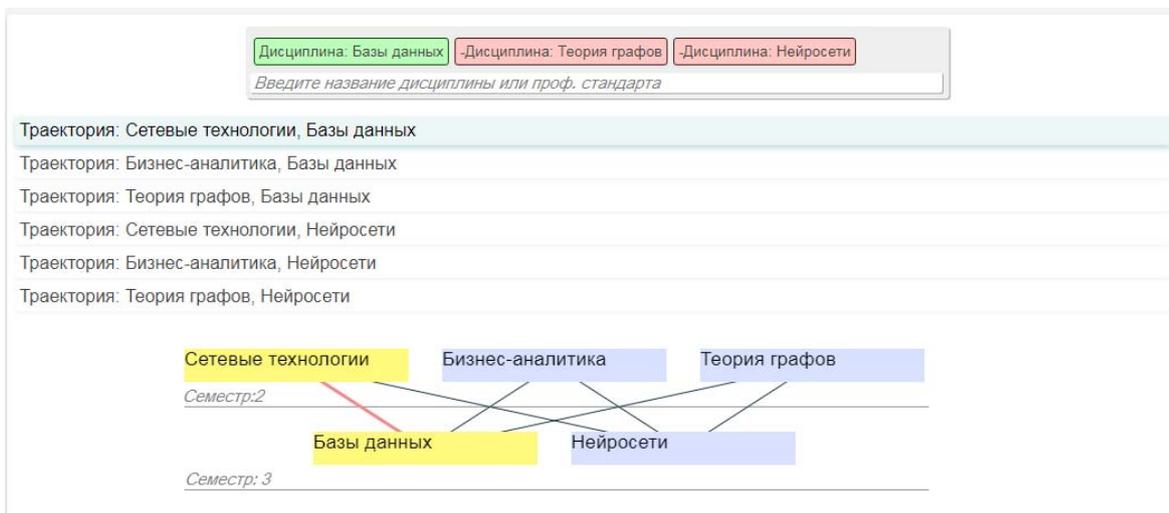


Рис. 6. Пример работы интерфейса рекомендательной системы

На рис. 6 показан интерфейс работы рекомендательной системы. Уже выбранные критерии отображаются над строкой поиска в виде кнопок красного или зелёного цвета, цвет определяет, включена данная дисциплина или профессиональный стандарт как «желаемая» или как «не желаемая». Нажатие на любую из кнопок исключит критерий из запроса.

Результат поиска отображается ниже в виде двух представлений:

- список траекторий, для каждой из которых указаны входящие в неё вариативные дисциплины; траектории отсортированы исходя из соответствия требованиям;
- граф, вершинами которого являются вариативные дисциплины, рёбрами – порядок их прохождения. В одной линии располагаются дисциплины, входящие в состав одного блока. При нажатии на траекторию в списке, на графе подсвечиваются соответствующие ей дисциплины.

## Заключение

В статье описана программная система ИТОС (Индивидуальные траектории обучения студентов). Она представляет собой совокупность четырех модулей, и предназначена для автоматизации работы с основной образовательной программой (ООП) и соответствующим ей учебным планом. Первый модуль предназначен для автоматизации формирования ООП в соответствии с актуальным ФГОСом. Данный модуль помогает отслеживать логику следования учебных курсов и позволяет на этапе составления ООП выявлять необходимость перестановки или слияния дисциплин, а также достижимость всех заданных компетенций. Второй модуль системы реализует алгоритм построения и оценки всевозможных индивидуальных траекторий обучения студента в рамках данной ООП. Данный алгоритм позволяет принимать решения о добавлении новых курсов в ООП с целью увеличения количества различных специальностей, которые можно получить в вузе, или количества дополнительных траекторий для получения уже существующих специальностей. Третий модуль реализует возможность устанавливать взаимосвязь между ООП и различными профессиональными

стандартами. Четвертый модуль реализует рекомендательную систему, помогающую пользователям (студентам) в выборе индивидуальных траекторий обучения.

Программная система реализована в виде web-приложения. Для реализации серверной части программного модуля используется фреймворк Spring, а для реализации клиентской части – фреймворк Angular. Все наборы курсов, все индивидуальные траектории – это граф. В связи с этим с помощью графовой БД работа с данными упрощается по сравнению с использованием реляционной БД. Нами была выбрана графовая база данных Neo4j.

### Список литературы

1. Крылов Д. Б., Леонтьева Ж. Г. Профессиональные стандарты: проблемы применения и перспективы развития // Вестн. Воронеж. гос. ун-та. Серия: Экономика и управление. 2017. № 3. С. 91–94.
2. Морозов М. А., Морозова Н. С. Подходы к оценке соответствия образовательных программ профессиональным стандартам // Высшее образование сегодня. 2017. № 10. С. 13–17.
3. Яхьяевой Г. Э., Абсайдульевой А. Р. Семантический подход к моделированию фонда оценочных средств // Вестн. НГУ. Серия: Информационные технологии. 2018. Т. 16, № 2. С. 113–121.
4. Пальчунов Д. Е. Моделирование мышления и формализация рефлексии. I. Теоретико-модельная формализация онтологии и рефлексии // Философия науки. 2006. № 4 (31). С. 86–114.
5. Пальчунов Д. Е. Моделирование мышления и формализация рефлексии. II. Онтологии и формализация понятий // Философия науки. 2008. № 2 (37). С. 62–99.
6. Пальчунов Д. Е., Яхьяева Г. Э., Ясинская О. В. Применение методологии онтологического моделирования для задач диагностирования заболеваний позвоночника // Вестн. НГУ. Серия: Информационные технологии. 2015. Т. 13, вып. 3, С. 42–51.
7. Мамеев Н. С. Онтологическая модель основной образовательной программы бакалавриата // Материалы Всерос. конф. с международным участием «Знания – Онтологии – Теории» (ЗОНТ-2017). Новосибирск, 2017. 175 с.
8. Михайлов А. С. Разработка онтологического паттерна ФГОС ВО // Развитие и актуальные вопросы современной науки. 2018. № 5 (12). С. 34–42.
9. Каракозов С. Д., Петров Д. А., Худжина М. В. Проектирование образовательных программ бакалавриата в условиях приведения федеральных государственных образовательных стандартов в соответствие с профессиональными стандартами // Преподаватель XXI век. 2015. Т. 1 (2). С. 9–23.
10. Рыжаков И. Д., Попова К. Ю., Мамеев Н. С., Баталин К. В. Разработка сервиса поддержки внедрения образовательных программ основанных на профессиональных стандартах // Актуальные научные исследования в современном мире. 2018. № 4 (36). С. 101–104.
11. Баталин К. В. Алгоритм построения индивидуальных траекторий обучения студента // Материалы Всерос. конф. с международным участием «Знания – Онтологии – Теории» (ЗОНТ-2017). Новосибирск, 2017. 175 с.
12. Попова К. Ю., Мамеев Н. С., Рыжаков И. Д. Разработка рекомендательной системы, помогающей студенту прорабатывать индивидуальную траекторию обучения // Материалы XV Междунар. конф. студентов, аспирантов и молодых ученых «Перспективы развития фундаментальных наук». Томск, 2018.
13. Яхьяева Г. Э., Карманова А. А., Ершов А. А., Савин Н. П. Вопросно-ответная система для управления информационными рисками на основе теоретико-модельной формализации предметных областей // Информационные технологии. 2017. Т. 23, № 2. С. 97–106.

**K. V. Batalin, N. S. Mameev, K. Yu. Popova, I. D. Ryzhakov, G. E. Yakhyaeva**

*Novosibirsk State University  
1 Pirogov Str., Novosibirsk, 630090, Russian Federation*

*kir55rus@yandex.ru, asm\_edf@mail.ru, kr111kr@yandex.ru  
igorusha4@yandex.ru, gul\_nara@mail.ru*

## **SOFTWARE SYSTEM FOR MANAGEMENT OF EDUCATIONAL PROCESS ITOS**

The article is devoted to the description of the program system of individual educational trajectory, intended for automation of work with the basic educational program and the appropriate curriculum. The system has implemented a functional for tracking the order of educational courses. This makes it possible, at the stage of forming the basic educational program, to identify the need for a reorganization or merging of disciplines and the attainability of all the assigned competencies. There is an opportunity to build a relationship between the basic educational program and various professional standards. The system implements an algorithm for constructing and evaluating all possible individual educational trajectories of student from a given basic educational program. And the module of the recommendatory system that helps users (students) in the selection of individual educational trajectories.

*Keywords:* basic educational program, professional standard, individual educational trajectory, recommender System

### **References**

1. Krilov D. B., Leont'eva Zh. G. Professional'nye standarty: problemy primeneniya i perspektivy razvitiya. *Vestnik Voronezhskogo gosudarstvennogo universiteta. Seria: Ekonomika i upravlenie*, 2017, no. 3, p. 91–94. (in Russ.)
2. Morozov M. A., Morozova N. S. Podhody k ocenke sootvetstviya obrazovatel'nyh programm professional'nym standartam. *Visshee obrazovanie segodnia*, 2017, no. 10, p. 13–17. (in Russ.)
3. Yakhyaeva G. E., Absayduleva A. R. Semantic Approach to Modeling of the Fund of Assessment Means. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 2, p. 113–121. (in Russ.)
4. Palchunov D. E. Modeling of intellection and formalization of reflection. I. Model-theoretic formalization of ontology and reflection. *Filosofia nauki*, 2006, no. 4 (31), p. 86–114. (in Russ.)
5. Palchunov D. E. Modeling of intellection and formalization of reflection. II. Ontologies and formalization of concepts. *Filosofia nauki*, 2008, no. 2 (37), p. 62–99. (in Russ.)
6. Palchunov D. E., Yakhyaeva G. E., Yasinskaya O. V. Application of model-theoretic methods and ontological modeling to automate the diagnosis of diseases. *Vestnik NSU. Series: Information Technologies*, 2015, vol. 13, no. 3, p. 42–51. (in Russ.)
7. Mameev N. S. Ontological model of the basic educational Bachelor program. *Materialy Vserossiyskoy konferentsii s mezhdunarodnym uchastiem "Znania – Ontologii – Teorii" (ZONT-2017)*. Novosibirsk, 2017, 175 p. (in Russ.)
8. Mikhailov A. S. Development of the ontological pattern FGOS VO. *Razvitye i aktual'nye voprosy sovremennoy nauki*, 2018, no. 5 (12), p. 34–42. (in Russ.)
9. Karakozov S. D., Petrov D. A., Hudgina M. V. Designing undergraduate educational programs in the context of bringing federal state educational standards in line with professional standards. *Prepodavatel' XXI vek*, 2015, vol. 1 (2), p. 9–23. (in Russ.)
10. Ryzhakov I. D., Popova K. Y., Mameev N. S., Batalin K. V. Development of the service for assistance of implementation of education program based on professional standards. *Aktual'nie nauchnie issledovaniya v sovremennoy mire*, 2018, no. 4 (36), p. 101–104. (in Russ.)

11. Batalin K. V. Algorithm construction of individual paths of the student studing. *Materialy Vserossiyskoy konferentsii s mezhdunarodnym uchastiem "Znania – Ontologii – Teorii" (ZONT-2017)*. Novosibirsk, 2017, 175 p. (in Russ.)

12. Popova K. Y., Mameev N. S., Ryzhakov I. D. Development of an advisory system that helps the student to build an individual student trajectory. *Materialy XV Mezhdunarodnoy konferentsii studentov, aspirantov i molodykh uchenykh "Perspektivi razvitia fundamental'nykh nauk"*. Tomsk, 2018, p. 153–155. (in Russ.)

13. Yakhyaeva G. E., Karmanova A. A., Ershov A. A., Savin N. P. Question-answering system for managing of the information risks based on model-theoretic formalization of the object domains. *Informatsionnye tekhnologii*, 2017, vol. 23, no. 2, p. 97–106. (in Russ.)

*For citation:*

Batalin K. V., Mameev N. S., Popova K. Yu., Ryzhakov I. D., Yakhyaeva G. E. Software System for Management of Educational Process ITOS. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 4, p. 20–30. (in Russ.) DOI 10.25205/1818-7900-2018-16-4-20-30

**М. А. Боронина, В. А. Вшивков**

*Институт вычислительной математики и математической геофизики СО РАН  
пр. Академика Лаврентьева, 6, Новосибирск, 630090, Россия*

*boronina@ssd.sgcc.ru, vsh@ssd.sgcc.ru*

## **ПАРАЛЛЕЛЬНЫЙ АЛГОРИТМ ДЛЯ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ ВСТРЕЧНЫХ ПУЧКОВ УЛЬТРАРЕЛЯТИВИСТСКИХ ЧАСТИЦ С УЧЕТОМ УГЛА ВСТРЕЧИ \***

Представлен алгоритм для численного моделирования динамики пучков в суперколлайдерах. При взаимодействии высокоэнергетичных пучков возможна их деформация и разрушение, и изучение устойчивости пучка является актуальной задачей. Рассматривается движение пучков в самосогласованных электромагнитных полях с учетом угла встречи в полностью трехмерном случае. Задача решается методом частиц-в-ячейках. Распараллеливание по области и по частицам позволяет проводить численные эксперименты с  $10^9$  модельных частиц. Приведены результаты моделирования и сравнение с существующим аналитическим решением.

*Ключевые слова:* метод частиц-в-ячейках, эффекты встречи, коллайдеры, численное моделирование.

### **Введение**

Ускорители на встречных пучках остаются одним из важных инструментов изучения фундаментальной физики. Постоянный интерес к высоким энергиям и светимостям сталкивается с научными и технологическими сложностями. В проекте Международного Линейного Коллайдера (International Linear Collider, ILC) энергия пучков в системе центра масс достигает 500 ГэВ, а ожидаемые светимости  $\sim 2 \cdot 10^{34} \text{ с}^{-1} \text{ см}^{-2}$  [1]. Однако критические плотности частиц в пучке приводят к высоким силам отталкивания между частицами пучка и возможному быстрому разрушению пучка, приводящему к понижению светимости. Перераспределение электромагнитных полей может вывести дорогую установку из строя. Угол встречи пучков (2–25 мрад) определяется системой «финальная линза – вывод пучка», при этом параметры пучка тесно связаны с углом встречи [2–4]. Изучение нелинейной динамики пучков является значимой задачей конструирования коллайдеров и их использования, в то время как их сложность и стоимость являются ограничивающими факторами их создания и проведения численных экспериментов. Численное моделирование может помочь избежать неожиданных проблем и оптимизировать параметры.

В работе представлен алгоритм для численного моделирования однопролетного взаимодействия встречных ультрарелятивистских пучков ( $\gamma \sim 10^3\text{--}10^5$ ) при наличии угла встречи. Существующие коды для моделирования динамики пучков основаны на разделении пучков в продольном направлении на слои частиц [5; 6]. Трехмерная динамика моделируется перестановкой двумерных слоев частиц: двумерные поля каждого слоя действуют на частицы слоя

---

\* Работа выполнена в рамках государственного задания ИВМиМГ СО РАН (проект 0315-2016-0009).

встречного пучка с теми же продольными координатами [7–9]. Сведение задачи к двумерной основано на допущении, что коллективное движение частиц происходит вдоль одной оси и является оптимальным для моделирования многопролетных режимов взаимодействия пучков в циклических коллайдерах, где деформации пучков малы на каждом обороте, а различия существенны лишь после тысяч оборотов. В случае резкого разрушения пучков с высокой плотностью и при наличии угла встречи численные эксперименты должны проводиться на основе трехмерной модели / алгоритма. Предлагаемый нами алгоритм учитывает трехмерность задачи. Смешанная лагранжево-эйлерова декомпозиция позволяет проводить численные эксперименты для гауссовых пучков с нелинейной фокусировкой. В работе представлены результаты взаимодействия пучков и сравнение с аналитическим решением в докритическом случае.

### Математическая модель

Рассматривается динамика пучков электронов / позитронов в лабораторной системе координат. Расчетная область представляет собой параллелепипед, пучки движутся в вакууме в самосогласованных электромагнитных полях. В модели используется кинетическое уравнение Власова для функции распределения электронов  $f_-$  и позитронов  $f_+$  и уравнения Максвелла:

$$\begin{aligned}\frac{\partial f_{+,-}}{\partial t} + \vec{v}_{+,-} \frac{\partial f_{+,-}}{\partial \vec{r}} + \vec{F}_{+,-} \frac{\partial f_{+,-}}{\partial \vec{p}} &= 0 \\ \operatorname{rot} \vec{E} &= -\frac{1}{c} \frac{\partial \vec{B}}{\partial t} \\ \operatorname{rot} \vec{B} &= \frac{4\pi}{c} \vec{j} + \frac{1}{c} \frac{\partial \vec{E}}{\partial t} \\ \operatorname{div} \vec{E} &= 4\pi(n_- e^- + n_+ e^+) \\ \operatorname{div} \vec{B} &= 0.\end{aligned}$$

Здесь

$$\vec{F}_{+,-} = q_{+,-} \left( \vec{E} + \frac{1}{c} [\vec{v}_{+,-}, \vec{B}] \right)$$

есть сила Лоренца. Скорости частиц обозначены как  $\vec{v}_{+,-}$ , импульсы –  $\vec{p}_{+,-}$ , координаты –  $\vec{r}_{+,-}$ , при этом релятивистские факторы  $\gamma_{+,-} = 1 / \sqrt{1 - |\vec{v}_{+,-}|^2 / c^2}$ ,  $\vec{p}_{+,-} = \gamma_{+,-} m_{+,-} \vec{v}_{+,-}$ . Уравнения характеристик уравнения Власова совпадают с уравнениями движения:

$$\begin{aligned}\frac{d\vec{p}_{+,-}}{dt} &= \vec{F}_{+,-}, \\ \frac{d\vec{r}_{+,-}}{dt} &= \vec{v}_{+,-}.\end{aligned}$$

Плотности токов и заряда вычисляются как моменты функции распределения:

$$\begin{aligned}\rho &= \int_{V_p} (q_+ f_+ + q_- f_-) d\vec{p}, \\ \vec{j} &= \int_{V_p} (q_+ f_+ \vec{v}_+ + q_- f_- \vec{v}_-) d\vec{p}.\end{aligned}$$

Уравнения автоматически учитывают трехмерность задачи. Начальные и граничные условия также должны ее учитывать. В начальный момент времени задаются координаты  $\vec{r}_{+,-}(0) = \vec{r}_{0+,-}$  и импульсы частиц  $\vec{p}_{+,-}(0) = \vec{p}_{0+,-}$ , поле в области определяется как поле пучков, сфокусированных полем коллайдера  $E_{x,y}(0) = E_{x,y+}(0) + E_{x,y-}(0)$ ,  $B_{x,y}(0) = B_{x,y+}(0) + B_{x,y-}(0)$ . Граничные условия предполагаются открытыми, границы находятся в ближней волновой зоне, и в поперечном направлении эффектом релятивистского запаздывания можно пренебречь. На границах в продольном направлении поля равны нулю в любой момент времени, так как частицы до них не долетают.

В качестве безразмерных величин взяты характерная скорость движения частиц пучка – скорость света  $c = 2.9979 \cdot 10^{10}$  см/с и размер пучка  $L_0 = 1$  см. Время измеряется в единицах  $T^* = L_0/c$ , импульсы  $p^* = m_e c$ , силы  $F^* = m_e c^2/L_0$ , поля  $E^* = m_e c^2/eL_0$  и  $H^* = m_e c^2/eL_0$ , плотности заряда  $\rho^* = E/4\pi L_0$  и плотности тока  $j^* = c\rho^*$ .

Для решения уравнений используется метод частиц-в-ячейках со схемой Бориса [11] и схемой Лэнгдона – Лазински [12] на сдвинутых сетках [13]:

$$\frac{B^{m+\frac{1}{2}} - B^{m-\frac{1}{2}}}{\tau} = -rot_h E^m,$$

$$\frac{E^{m+1} - E^m}{\tau} = j^{m+\frac{1}{2}} + rot_h B^{m+\frac{1}{2}}.$$

Схемы обладают вторым порядком точности по времени и пространству. Применение алгоритма Бунемана [14] для расчета токов позволяет удовлетворить разностный аналог уравнения неразрывности.

В работе рассматриваются пучки с одинаковым распределением, различие состоит в знаке заряда и углах с продольной осью  $z$ . Плотности заряда невозмущенных пучков определяются фокусировкой квадрупольной линзой:

$$\rho_{+,-}(x, y, z, p_x, p_y) = \frac{N_{+,-} e}{(2\pi)^{\frac{3}{2}} \sigma_x^* \sigma_y^* \sigma_z} \times$$

$$\times \exp \left[ -\frac{(x \pm (x_{IP} - x_{c+,-}))^2}{2\sigma_x^2} - \frac{(y \pm (y_{IP} - y_{c+,-}))^2}{2\sigma_y^2} - \frac{(z \pm (z_{IP} - z_{c+,-}))^2}{2\sigma_z^2} \right] \times$$

$$\times \frac{1}{2\pi \sigma_x^* \sigma_y^*} \exp \left[ -\frac{p_x^2}{2p_z^2 \sigma_x'^2} - \frac{p_y^2}{2p_z^2 \sigma_y'^2} \right].$$

Поперечные размеры пучков с центрами  $(x_{c+,-}, y_{c+,-}, z_{c+,-})$  определяются значениями  $\sigma_{x,y}^* = \sqrt{\beta_{x,y} \varepsilon_{x,y}}$ ,  $\beta_{x,y} = \beta_{x,y}^* + (z - z_{IP})^2 / \beta_{x,y}^*$ , где  $\varepsilon_x, \varepsilon_y$  – эмиттансы,  $\beta_x^*, \beta_y^*$  – значения бета-функции в месте встречи. Поперечные импульсы также имеют гауссово распределение с  $\sigma_{x',y'}^* = \sqrt{\varepsilon_{x,y} / \beta_{x,y}^*}$ , пучки моноэнергетические  $|\vec{p}_{+,-}| = (p_{x+,-}, p_{y+,-}, p_{z+,-}) = \sqrt{\gamma_{+,-}^2 - 1}$  и  $\gamma_+ = \gamma_- = \gamma$ . Ось первого пучка направлена под углом  $\theta$  к оси  $z$  в плоскости  $(x, z)$ . Огибающие пучков определяются значениями  $\beta_x^*, \beta_y^*$  и обозначены соответствующим пучку цветом для  $3.5\sigma_{x,y}$  на рис. 1 (эффект песочных часов).

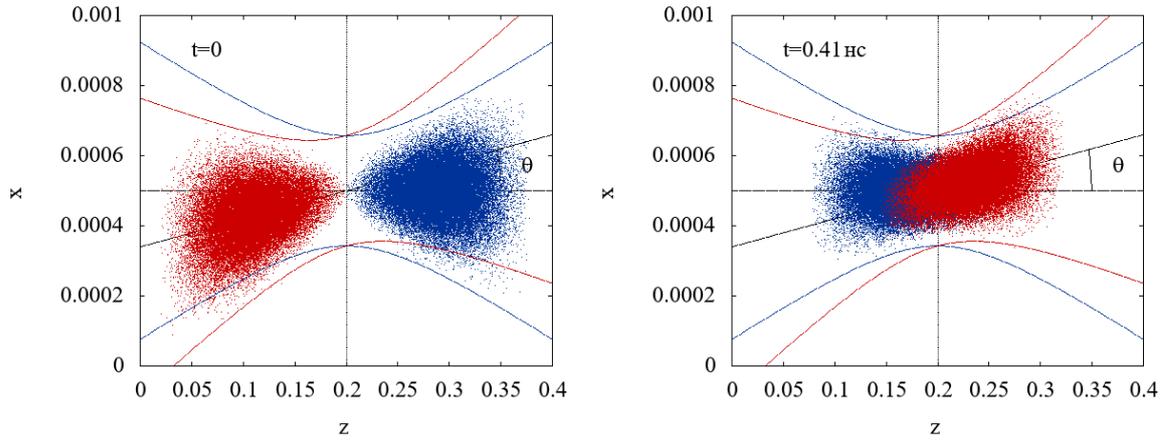


Рис. 1. Координаты пучков  $(z, x)$  при  $t = 0$  (слева) и  $t = 0,41$  нс (справа)

Электрическое поле пучка может быть вычислено по формуле для движущегося релятивистского заряда

$$\vec{E} = \frac{1}{\gamma^2} \frac{q}{(1 - \beta^2 \sin^2 \alpha)^{\frac{3}{2}}} \frac{\vec{R}}{R^3},$$

где  $\vec{R}$  – радиус-вектор точки пространства, где вычисляется поле, из местоположения заряда,  $\alpha$  – угол между направлением скорости  $\vec{v}$  и  $\vec{R}$  [15]. Электрическое поле в перпендикулярном движении направлении увеличивается в  $\gamma$  раз, в продольном направлении сокращается в  $\gamma^2$  раз по сравнению с нерелятивистским случаем. Для задания начальных и граничных условий проводится суммирование вкладов полей от элементов плотности заряда с помощью специального вида модельных частиц [16]. Для пучков, двигающихся вдоль оси  $z$  релятивистские факторы  $\sim 10^3$ – $10^6$  позволяют суммировать вклады в поле лишь от узлов плотности с той же продольной координатой, так как поле сосредоточено в экваториальной плоскости. Поскольку углы встречи в суперколлайдерах составляют 2–20 мрад, а соотношение размеров пучков  $\sigma_x : \sigma_z \sim 10$  нм : 1 мм, то даже хорошее разрешение сетки в продольном направлении (1000 узлов) не является достаточным для описания различия в поле с учетом угла встречи и без его учета. Таким образом, электрическое поле пучка, двигающегося под углом к оси  $z$ , вычисляется также по [16]. Начальные и граничные поля для магнитного поля вычисляются по формуле

$$\vec{B}_{+,-} = [\vec{v}_{+,-}, \vec{E}_{+,-}].$$

Светимость описывает количество актов взаимодействия между частицами в сечении за единицу времени и может быть вычислена как

$$L = K \iiint \rho_{+x}(x) \rho_{+y}(y) \rho_{+z}(s - s_0) \rho_{-x}(x) \rho_{-y}(y) \rho_{-z}(s + s_0) dx dy ds ds_0,$$

где  $K = \sqrt{(\vec{v}_1 - \vec{v}_2)^2 - (\vec{v}_1 \times \vec{v}_2)^2}$  – кинематический фактор,  $s_0 = ct$  – расстояние до места встречи [17; 18]. Численно этот интеграл вычисляется суммированием:

$$L = KN_+ N_- \sum_m \sum_{i,l,k} \rho_{+x}(x) \rho_{+y}(y) \rho_{+z}(z - ct) \rho_{-x}(x) \rho_{-y}(y) \rho_{-z}(z + ct) h_x h_y h_z \tau,$$

где  $x = (i - 1.5)h_x$ ,  $y = (l - 1.5)h_y$ ,  $z = (k - 1.5)h_z$ ,  $t = m\tau$ , а  $c = 1$  в безразмерных переменных.

В общем случае аналитических решений задачи динамики пучков не существует, но для пучков с малыми зарядами взаимодействием частиц можно пренебречь, и светимость может быть вычислена по формуле [19]:

$$L = \frac{N_1 N_2 \cos(\theta/2)}{2\pi \sigma_x^* \sigma_y^* \sigma_z} \int \frac{s^{-s^2 A} ds}{\sqrt{\frac{s^2 \cos^2(\theta/2)}{\beta_x^{*2}} + 1} \sqrt{\frac{s^2 \cos^2(\theta/2)}{\beta_y^{*2}} + 1}}, \quad (1)$$

где

$$A = \frac{\cos^2(\theta/2)}{\sigma_z^2} + \frac{\sin^2(\theta/2)}{\sigma_x^{*2} \left( 1 + \frac{s^2 \cos^2(\theta/2)}{\beta_x^{*2}} \right)}.$$

Поскольку светимость – интегральная величина, то ее можно использовать в качестве теста на качество работы алгоритма при моделировании докритических режимов динамики пучков.

Пучок в начальный момент времени находится далеко от места фокусировки, его размеры в поперечном направлении достаточно велики, и расчетная область должна быть соответствующего размера. Размеры в месте встречи определяются малыми величинами  $\sigma_{x,y}^*$ , на них должно приходиться не менее 5–10 узлов сетки. Поэтому для расчетной области большого размера должна использоваться сетка с достаточно большим количеством узлов [20]. При этом распределение частиц в пучке существенно неравномерно, за счет эффекта песочных часов в малой области вокруг места встречи может скапливаться более 70 % частиц, и вместить все данные частиц в память одного процессора не представляется возможным [21]. Точность метода частиц зависит от количества частиц в ячейке, поэтому измельчение сетки должно сопровождаться соответствующим увеличением общего числа частиц. Условная устойчивость схемы Лэнгдона – Лазински приводит к необходимости измельчать временной шаг при измельчении минимального пространственного шага. Таким образом, для проведения численных экспериментов с  $10^7$  модельными частицами на сетках  $80^3$  уже требуется распараллеливание. Для этого применяется смешанная декомпозиция: область делится на подобласти, каждая из которых обрабатывается группой процессоров. Частицы, находящиеся в этой области, распределены равномерно между процессорами группы, число процессоров в группе определяется количеством частиц в области в начальный момент времени. Также допускается ручная настройка, поскольку при критических плотностях пучков неизвестно, как будет деформироваться пучок и где будут находиться частицы. Обмены сеточными значениями полей, плотностей, токов требуют двух вспомогательных узлов для каждой подобласти [22]. При пересылках данных частиц выбирается случайный процессор из соседней группы, что позволяет сохранить равномерность распределения частиц внутри групп. Количество модельных частиц определяется количеством используемых процессоров [23–25].

## Результаты численных экспериментов

Рассмотрим два встречных пучка с  $\beta_x = 1.6$  см,  $\beta_y = 0.034$  см,  $\varepsilon_x = 2.03 \cdot 10^{-8}$  см·рад,  $\varepsilon_y = 1.19 \cdot 10^{-9}$  см·рад,  $\sigma_z = 0.03$  см,  $\gamma = 6850$  в расчетной области размерами  $L_x = 0.005$  см,  $L_y = 0.0006$  см,  $L_z = 0.4$  см, пучки сталкиваются в  $(L_x/2, L_y/2, L_z/2)$ . На рис. 2, а показана зависимость увеличения светимости  $H_D$  от параметра разрушения  $D_y$ . Здесь увеличение светимости сравнивается с ее значением для  $10^6$  частиц  $H_D = L(D_y)/L(D_{y0}) = (L(Q)/Q^2)/(L(Q_0)/Q_0^2)$ ,  $L(Q)$  учитывает эффект песочных часов, заряд пучков  $Q = \pm Ne$ ,  $N_+ = N_- = N$ ,  $Q_0 = \pm N_{0e}$ ,  $N_0 = 10^6$ , параметр разрушения определяется как

$$D_{x,y} = \frac{2Nr_e \sigma_z}{\gamma \sigma_{x,y} (\sigma_x + \sigma_y)}.$$

Для некритических режимов  $D_{x,y} \ll 1$  пучки не испытывают возмущений, и светимость совпадает с геометрической.

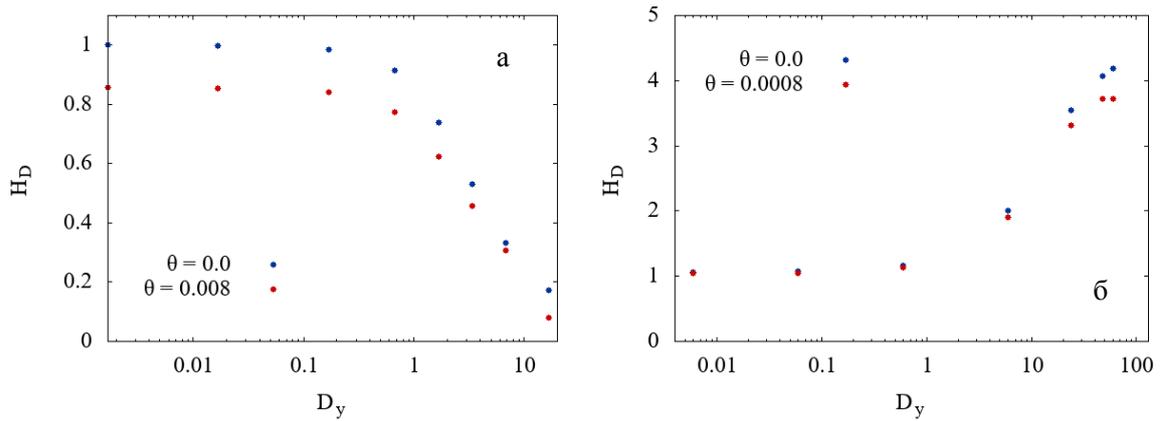


Рис. 2. Зависимость приращения светимости от параметра разрушения

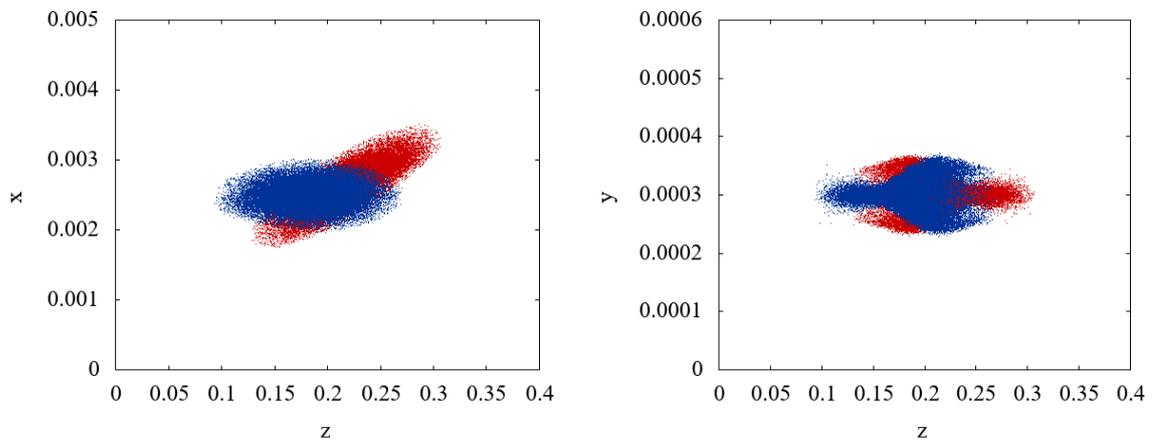


Рис. 3. Координаты пучков  $(z, x)$  (слева) и  $(z, y)$  (справа) при  $t = 0.36$  нс для первого набора параметров

Зависимость  $H_D$  демонстрирует падение светимости, что связано с деформацией пучков. На рис. 3. показаны координаты пучков с зарядами при  $N = 4 \cdot 10^9$ , угле встречи  $\theta = 0.08$  в момент времени  $t = 0.36$  нс.

Выбор параметров может привести к другим эффектам. Например, при  $\beta_x = 0.08$  см,  $\beta_y = 0.034$  см,  $\epsilon_x = 2.53 \cdot 10^{-8}$  см·рад,  $\epsilon_y = 1.79 \cdot 10^{-9}$  см·рад,  $\sigma_z = 0.03$  см,  $\gamma = 6850$  размеры области  $L_x = 0.001$  см,  $L_y = 0.0004$  см,  $L_z = 0.4$  см. На рис. 2, б представлены зависимости  $H_D$  от  $D_y$  при  $\theta = 0$  (синий цвет) и  $\theta = 0.0008$  (красный цвет). Углы встречи снижают устойчивость пучка,  $H_D$  выше при лобовом столкновении пучков. Также наблюдается повышение светимости в 4 раза по сравнению с некритическим режимом за счет сильного сжатия пучков.

На рис. 1 показаны координаты пучков с малыми зарядами ( $N = 10^6$ ), которые отвечают некритическому режиму. На рис. 4 показаны координаты пучков для зарядов с  $N = 10^{10}$  ( $D_x = 10$ ,  $D_y = 60$ ) в момент времени  $t = 0.41$  нс. Плотность пучков высока, пучки разрушаются вследствие сильной фокусировки полем встречного пучка, и светимость в этом случае примерно такая же, что и для случая  $D_y = 48$ .

Сравнение светимости может быть сделано для некритических режимов. На рис. 5, а для первого случая показано приращение светимости  $L(\theta)/L(0)$  в зависимости от угла встречи для пучков с  $N = 10^6$  (красные точки) и для режима с разрушением пучка ( $N = 4 \cdot 10^9$ , синие точки) и приращение  $L_{an}(\theta)/L_{an}(0)$  (черная линия), где интеграл (1) был получен численно методом прямоугольников. На рис. 5, б аналогично показаны  $L(\theta)/L(0)$  для второго случая с  $N = 10^6$  (красные точки) и с  $N = 10^{10}$  (синие точки), а также светимость  $L_{an}(\theta)/L_{an}(0)$ . Для меньших бета-функций угол встречи играет существенную роль в понижении устойчивости пучка.

В обоих случаях зависимость светимости от угла в некритическом режиме совпадает с теоретически ожидаемой.

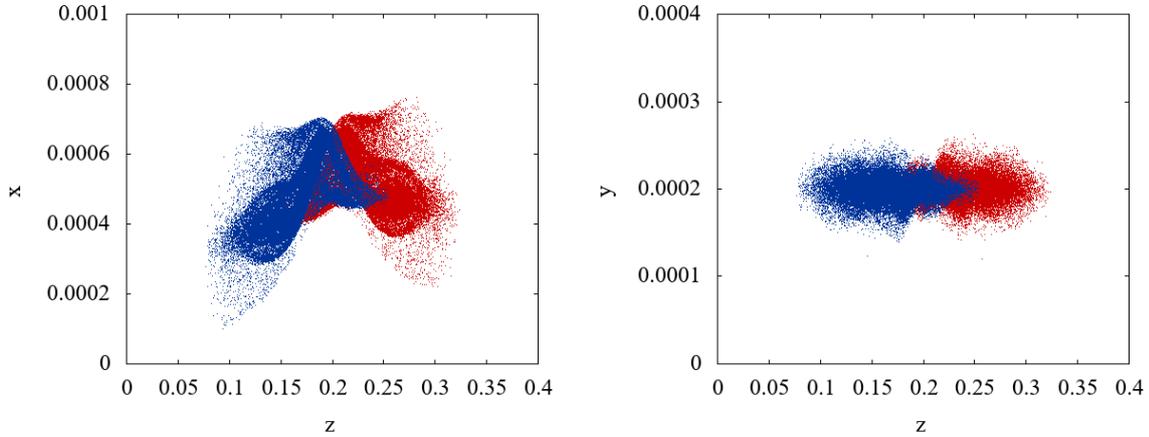


Рис. 4. Координаты пучков  $(z, x)$  (слева) и  $(z, y)$  (справа) для второго набора параметров

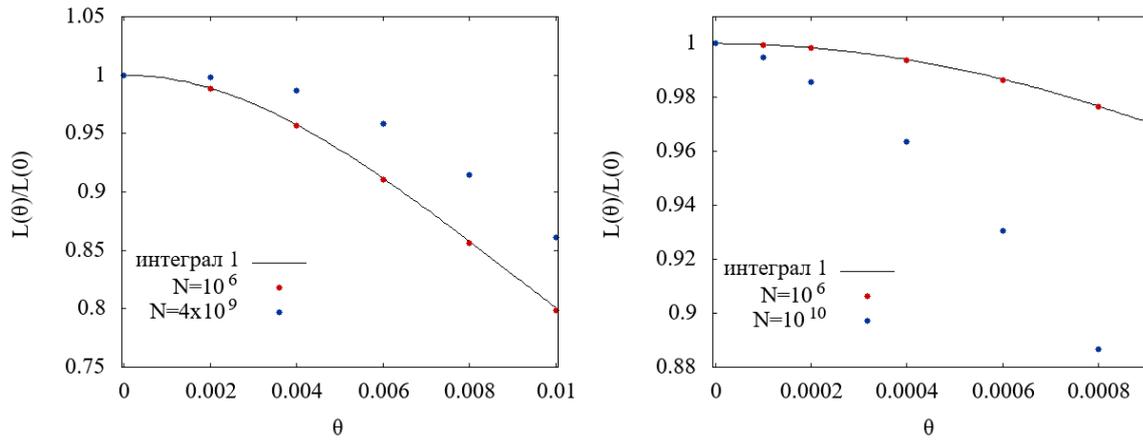


Рис. 5. Зависимость приращения светимости от угла встречи:  
 $a$  – при  $\beta_x^* = 1,6$ ;  $b$  – при  $\beta_x^* = 0,08$

Для вычислений использовалась сетка с  $100 \times 100 \times 100$  узлами,  $10^8$  модельных частиц,  $56 \cdot 10^3$  временных шагов  $\tau = 3.33 \cdot 10^{-6}$  для первого случая и  $62 \cdot 10^3$  временных шагов  $\tau = 3 \cdot 10^{-6}$  для второго; 224 процессорных ядра были распределены между 25 группами, в центральной группе было 136 ядер для первого случая и 104 для второго. В обоих случаях вычисления длились  $\sim 27$  часов. Для меньших сеток, например,  $50 \times 50 \times 50$  при  $10^7$  модельных частиц соответственно более крупный временной шаг был использован, и при вычислениях использовалось 56 ядер в 11 группах, на каждый расчет потребовалось  $\sim 5$  часов.

Расчеты проводились на суперкомпьютере Ломоносов (МГУ, Москва), Политехник (СПбПУ, Санкт-Петербург) и кластере ССКЦ (ИВМиМГ СО РАН, Новосибирск).

**Заключение**

Предложен новый параллельный алгоритм для численного моделирования взаимодействия пучков с высокими энергиями в суперколлайдерах. Алгоритм позволяет проводить вычисления для однопролетного взаимодействия пучков в полностью трехмерном ультрарелятивистском случае с учетом угла встречи и эффекта песочных часов, когда возможна сильная нелинейная деформация и разрушение пучков. Алгоритм основан на решении уравнения Власова методом частиц-в-ячейках и уравнений Максвелла. Распараллеливание по области

и по частицам позволяет решить проблему неравномерного распределения частиц в пространстве и времени. Представлены результаты численного моделирования и анализ эволюции пучка в трехмерном случае. Результаты демонстрируют качественное совпадение с теоретическими ожиданиями для критических режимов и количественное совпадение для некритических режимов. Таким образом, показана работоспособность и перспективность метода для математического моделирования эффектов взаимодействия пучков при наличии угла встречи. Код может быть использован при исследовании проблем, связанных с устойчивостью пучка и оптимизацией параметров, а также для подготовки студентов соответствующей специальности.

### Список литературы

1. *Adolphsen C. et al.* The International Linear Collider Technical Design Report – Volume 3.II: Accelerator Baseline Design (2013) arXiv:1306.6328 [physics.acc-ph] .
2. *Appleby R., Angal-Kalinin D., Bambade P., Cavalier S., Meur G. Le, Touze F., Iwashita Y.* The 2 mrad Crossing Angle Scheme for the International Linear Collider // Proc. of EPAC08 (Genoa, Italy, 2008), p. MOPP005.
3. *Nosochkov Y., Mo\_eit K., Seryi A., Woods M., Arnold R., Oliver W., Parker B., Torrence E.* Design of ILC Extraction Line for 20 mrad Crossing Angle // Proc. of the 2005 Particle Accelerator Conference (Knoxville, USA, 2005), p. RPPP030.
4. *Telnov V. I.* Journal of Instrumentation 13, p. P03020 (2018).
5. *Hirata K.* Phys. Rev. Lett. 74, 2228–2231 (1994).
6. *Buffat X., Barraud L., Barranco J., Florio A. and Pieloni T.* Numerical and experimental studies of coherent beam-beam modes: stability and decoherence // Workshop on Beam-Beam Effects in Circular Colliders (Berkley, USA, 2018), p. RPPP030.
7. *Schulte D.* Study of Electromagnetic and Hadronic Background in the Interaction Region of the TESLA Collider. PhD thesis, DESY 1997.
8. *Terzic B. et al.* Long-Term Simulations of Beam-Beam Dynamics on GPUs // Proc. of IPAC2017 (Copenhagen, Denmark, 2017), p. THPAB086.
9. *Qiang J., Furman M., Ryne R.* Phys. Rev. ST Accel. Beams 5, p. 104402 (2002).
10. *Hockney R. W., Eastwood J. W.* Computer Simulation Using Particles (CRC Press, Boca Raton, Florida, USA, 1988).
11. *Boris J. P.* Relativistic plasma simulation – optimization of a hybrid code (Fourth Conference on numerical Simulation of Plasmas. Washington, 1970), p. 3–67.
12. *Langdon A. B., Lasinski B. F.* Controlled Fusion, Methods in Computational Physics: Advances in Research and Applications, vol. 16 (Elsevier, 1976), p. 327–366.
13. *Yee K. S.* IEEE Trans. Antenn. Propagat. AP-14, 302–307 (1966).
14. *Villasenor J., Buneman O.* Comput. Phys. 69, 306–316 (1992).
15. *Landau L. D., Lifshitz E. M.* The Classical Theory of Fields. A Course of Theoretical Physics. Pergamon Press, 1971, vol. 2.
16. *Вшивков В. А., Боронина М. А.* Трехмерное моделирование динамики ультрарелятивистских пучков заряженных частиц: особенности вычисления начальных и граничных условий // Матем. моделирование. 2012. Т. 24 (2). С. 67–83.
17. *Sands M.* The physics of electron storage rings: an introduction, SLAC-121 (1971).
18. *Herr W.* Particle Colliders and Concept of Luminosity // CERN Accelerator School. Granada, Spain, 2012.
19. *Peng Y., Zhang Y.* Luminosity Reduction with Hourglass Effect and Crossing Angle in an e-p Collider // Proc. of IPAC2015 (Richmond, USA, 2015), p. TUPTY010.
20. *Lotov K. V., Timofeev I. V., Mesyats E. A., Snytnikov A. V., Vshivkov V. A.* Physics of Plasmas 22, p. 024502 (2015).
21. *Boronina M., Vshivkov V., Levichev E., Nikitin S., Snytnikov V.* 3D PIC Method Development for Simulation of Beam-Beam Effects in Supercolliders // Proceedings of PAC07 (Albuquerque, USA, 2007), p. THPAN060.
22. *Boronina M. A., Korneev V. D.* Bull. Nov. Comp. Center 16, 15–22 (2013).
23. *Kireev S.* Parallel Computing Technologies, 5698, 406–413 (2009).

24. Berendeev E., Dudnikova G., Efimova A., Vshivkov V. Numerical Analysis and Its Applications, 10187, 227–234 (2017).
25. Vshivkov V. A., Boronina M. A. Journal of Plasma Physics 81, p. 495810605 (2015).
26. Chen P., Yokoya K. Phys. Rev. D38, p. 987 (1988).
27. Birdsall C. K., Langdon A. B. Plasma Physics via Computer Simulation (Institute of Physics Publishing, Bristol, UK, 1991).
28. Березин Ю. А., Вишивков В. А. Метод частиц в динамике разреженной плазмы. Новосибирск: Наука, 1980.

Материал поступил в редколлегию 15.10.2018

**M. A. Boronina, V. A. Vshivkov**

*Institute of Computational Mathematics and Mathematical Geophysics SB RAS  
6 Academician Lavrentiev Ave., Novosibirsk, 630090, Russian Federation*

*boronina@ssd.sccc.ru, vsh@ssd.sccc.ru*

### PARALLEL ALGORITHM FOR NUMERICAL SUMULATIONS OF ULTRARELATIVISTIC PARTICLE BEAMS COLLIDING WITH CROSSING ANGLE

We present an algorithm for numerical modelling of beam dynamics in supercolliders. When the high-energy beams interact their deformation and disruption are possible and the study of the beam stability is an important problem. We consider the beams motion in self-consistent electromagnetic fields and we take into account the crossing angle of the beams in the fully three-dimensional case. We use particle-in-cell method to solve the problem. The domain and particle parallelization allows performing of the numerical experiments with  $10^9$  model particles. The results of the simulations and the comparison with the existing analytical solution are presented.

*Keywords:* particle-in-cell method, beam-beam effects, colliders, numerical modelling

#### References

1. Adolphsen C. et al. The International Linear Collider Technical Design Report – Volume 3.II: Accelerator Baseline Design. Office of Scientific and Technical Information (OSTI), 2013. DOI 10.2172/1347940
2. Appleby R., Angal-Kalinin D., Bambade P., Cavalier S., Meur G. Le, Touze F., Iwashita Y. The 2 crossing angle scheme for the international linear collider. *EPAC 2008 – Contributions to the Proceedings*. Genoa, 2008, p. 556–558.
3. Nosochkov Y., Moffeit K., Seryi A., Woods M., Arnold R., Oliver W. et al. Design of ILC Extraction Line for 20 Mrad Crossing Angle. *Proceedings of the 2005 Particle Accelerator Conference*. Knoxville, 2005. DOI 10.1109/pac.2005.1591034
4. Telnov V. I. Optimization of the Beam Crossing Angle at the ILC for E+e- and  $\gamma\gamma$  Collisions. *Journal of Instrumentation*, 2018, vol. 13, no. 03, p. P03020–P03020. DOI 10.1088/1748-0221/13/03/p03020
5. Hirata K. Analysis of Beam-Beam Interactions with a Large Crossing Angle. *Physical Review Letters*, 1995, p. 2228–2231. DOI 10.1103/physrevlett.74.2228
6. Buffat X., Barraud L., Barranco J., Florio A., Pieloni T. Numerical and experimental studies of coherent beam-beam modes: stability and decoherence. *Workshop on Beam-Beam Effects in Circular Colliders*. Berkeley, 2018.
7. Schulte D. Study of Electromagnetic and Hadronic Background in the Interaction Region of the TESLA Collider. PhD thesis. DESY, 1997.
8. Terzić B. et al. Long-Term Simulations of Beam-Beam Dynamics on GPUs. *Proc. 8<sup>th</sup> Int. Particle Accelerator Conf. (IPAC'17)*. Copenhagen, 2017, p. 3918–3920. DOI <https://doi.org/10.18429/JACoW-IPAC2017-THPAB086>
9. Qiang J., Furman M., Ryne R. Strong-strong beam-beam simulation using a Green function approach. *Physical Review Special Topics – Accelerators and Beams*, 2002. DOI 10.1103/physrevstab.5.104402
10. Hockney R. W., Eastwood J. W. Computer Simulation Using Particles, 1988. DOI 10.1887/0852743920

11. Boris J. P. Relativistic plasma simulation-optimization of a hybrid code. *Proc. 4<sup>th</sup> Conf. Num. Sim. Plasmas*. Washington, 1970, p. 3–67.
12. Langdon A. B., Lasinski B. F. Electromagnetic and Relativistic Plasma Simulation Models. Controlled Fusion. *Methods in Computational Physics: Advances in Research and Applications*, 1976, vol. 16, p. 327–366. DOI 10.1016/B978-0-12-460816-0.50014-2
13. Yee K. Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, 1966, vol. 14, no. 3, p. 302–307. DOI 10.1109/TAP.1966.1138693
14. Villasenor J., Buneman O. Rigorous charge conservation for local electromagnetic field solvers. *Computer Physics Communications*, 1992, vol. 69, no. 2–3, p. 306–316. DOI 10.1016/0010-4655(92)90169-y
15. Landau L. D., Lifshitz E. M. The Classical Theory of Fields, Vol. 2 of A Course of Theoretical Physics, 1971.
16. Vshivkov V., Boronina M. Three-dimensional simulation of ultrarelativistic charged beams dynamics: study of initial and boundary conditions. *Matem. Mod.*, 2012, vol. 24, no. 2, p. 67–83.
17. Sands M. Physics of electron storage rings: an introduction, 1970. DOI 10.2172/4064201
18. Herr W. Particle Colliders and Concept of Luminosity. CERN Accelerator School. Granada, 2012.
19. Peng Y. M., Zhang Y. The Luminosity Reduction with Hourglass Effect and Crossing Angle in an e-p Collider. *Proc. 6<sup>th</sup> Int. Particle Accelerator Conf. (IPAC'15)*. Richmond, 2015, p. 2016–2018. DOI 10.18429/JACoW-IPAC2015-TUPTY010
20. Lotov K. V., Timofeev I. V., Mesyats E. A., Snytnikov A. V., Vshivkov V. A. Note on quantitatively correct simulations of the kinetic beam-plasma instability. *Physics of Plasmas*, 2015, vol. 22, no. 2. DOI 10.1063/1.4907223
21. Boronina M., Vshivkov V., Levichev E., Nikitin S., Snytnikov V. 3D PIC Method Development for Simulation of Beam-Beam Effects in Supercolliders. *Proc. of PAC-07*. Albuquerque, 2007. DOI 10.1109/pac.2007.4440427
22. Boronina M. A., Korneev V. D. The parallel three-dimensional PIC code for the numerical modeling of ultrarelativistic beams. *Bulletin of the Novosibirsk Computing Center*, 2013, no. 16, p. 15–22.
23. Kireev S. A Parallel 3D Code for Simulation of Self-gravitating Gas-Dust Systems. *Parallel Computing Technologies. PaCT 2009. Lecture Notes in Computer Science*, 2009, vol. 5698, p. 406–413. DOI 10.1007/978-3-642-03275-2\_40
24. Berendeev E., Dudnikova G., Efimova A., Vshivkov V. Computer Simulation of Plasma Dynamics in Open Plasma Trap. *Numerical Analysis and Its Applications*, 2017, p. 227–234. DOI 10.1007/978-3-319-57099-0\_23
25. Boronina M. A., Vshivkov V. A. Parallel 3-D particle-in-cell modelling of charged ultrarelativistic beam dynamics. *Journal of Plasma Physics*, 2015, vol. 81, no. 6. DOI 10.1017/S0022377815001178
26. Chen P., Yokoya K. Disruption effects from the interaction of round e-beams. *Physical Review D*, 1988, vol. 38, no. 3, p. 987–1000. DOI 10.1103/physrevd.38.987
27. Birdsall C. K., Langdon A. B. Plasma Physics via Computer Simulation, 1991. DOI 10.1887/0750301171
28. Berezin Y. A., Vshivkov V. A. Particles method in the dynamics of a rarefied plasma. Novosibirsk, Nauka, 1980. (in Russ.)

Received 15.10.2018

*For citation:*

Boronina M. A., Vshivkov V. A. Parallel Algorithm for Numerical Simulations of Ultra Relativistic Particle Beams Colliding with Crossing Angle. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 4, p. 31–40. (in Russ.) DOI 10.25205/1818-7900-2018-16-4-31-40

**В. Н. Глинских<sup>1-3</sup>, М. Н. Никитенко<sup>1</sup>**

<sup>1</sup> *Институт нефтегазовой геологии и геофизики им. А. А. Трофимука СО РАН  
пр. Академика Коптюга, 3, Новосибирск, 630090, Россия*

<sup>2</sup> *Новосибирский государственный университет  
ул. Пирогова, 1, Новосибирск, 630090, Россия*

<sup>3</sup> *Новосибирский государственный технический университет  
пр. К. Маркса, 20, Новосибирск, 630090, Россия*

*GlinskikhVN@ipgg.sbras.ru, NikitenkoMN@ipgg.sbras.ru*

**ТЕОРЕТИКО-АЛГОРИТМИЧЕСКАЯ БАЗА  
И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ДАННЫХ  
ДИЭЛЕКТРИЧЕСКОГО КАРОТАЖА ДЛЯ ИЗУЧЕНИЯ ЧАСТОТНОГО СПЕКТРА  
ЭЛЕКТРОФИЗИЧЕСКИХ ПАРАМЕТРОВ ГЕОЛОГИЧЕСКОЙ СРЕДЫ \***

Статья посвящена разработке теоретической базы компьютерного моделирования данных диэлектрического каротажа с целью обоснования нового импортозамещающего оборудования для каротажа нефтяных и газовых скважин. Разработаны алгоритмы и реализованы быстрые программы компьютерного моделирования высокочастотных электромагнитных сигналов в рамках слоисто-однородных моделей сред. По результатам масштабного моделирования амплитудно-фазовых характеристик высокочастотного поля установлена их зависимость от частотного спектра электрофизических параметров. Проведенные исследования позволяют выполнить научное обоснование и оптимальное проектирование новой электромагнитной зондирующей системы.

*Ключевые слова:* компьютерное моделирование, прямая задача, геоэлектрическая модель, электрофизические параметры, частотная дисперсия, диэлектрический каротаж.

## **Введение**

Электромагнитные методы играют важную роль в комплексе геофизических методов исследования геологической среды в поисковых, разведочных и эксплуатационных скважинах. Современной тенденцией в развитии аппаратурно-методического обеспечения электромагнитного каротажа является создание многочастотных, многоэлементных, многокомпонентных зондовых систем. На этом пути уже преодолены многие технические и методические сложности, такие как глубокая компенсация «прямого поля», коммутация пространственно-направленного условия «возбуждение – приём», оперативная обработка данных, быстрая передача сигналов, в том числе беспроводным способом, и многое другое. Модификация зондирующих установок, а зачастую создание приборов следующего поколения на новой

---

\* Исследования выполнены в рамках проекта «Идентификация математических моделей акустики, электродинамики и теории упругости» комплексной программы фундаментальных исследований СО РАН «Междисциплинарные интеграционные исследования».

микропроцессорной и элементной базе есть следствие прорыва в информационно-вычислительных технологиях, в частности для каротажных исследований.

Каротаж на переменном электрическом токе в области высоких частот (20–500 МГц) называется диэлектрическим каротажем (ДК). ДК предназначен для исследования пространственного распределения диэлектрической проницаемости (ДП) горных пород. Отечественная аппаратура ДК разработана в 60–70-е годы прошлого столетия [1]. Развитие теоретической и аппаратурной базы связано с волновым [2–5] и индуктивным [6] ДК, использующими мега- и гигагерцовый диапазоны частот соответственно. Диэлектрический каротаж на предельно высоких частотах в условиях слабопроводящих сред оперирует приближениями лучевого ряда и кинематической картиной волнового поля [7–11].

В настоящее время в компаниях Halliburton и Schlumberger применяется аппаратура волнового ДК. Она успешно используется в различных геологических условиях, где традиционные комплексы геофизических исследований и подходы к интерпретации недостаточно эффективны [12–15]. Наряду с аппаратурой ДК изучение ДП горных пород проводят по данным электромагнитного каротажа в диапазоне частот от сотен килогерц до единиц мегагерц. При этом основные результаты интерпретации данных ЭМК связывают с выявлением эффекта частотной дисперсии комплексной УЭП, определением ДП для оценки характера насыщения пород-коллекторов и изучением глинистых отложений [16–26].

В связи с повсеместным вовлечением в изучение новых типов геологических залежей, в том числе с трудноизвлекаемыми запасами, существенно усложняется и расширяется круг задач скважинной геоэлектрики. Это приводит к необходимости создания новых электромагнитных методов исследования геологической среды, разработки новых подходов в решении прямых и обратных задач электродинамики, а также использования новых реалистичных интерпретационных моделей. Текущее развитие теоретико-методической базы электромагнитных зондирований геологических сред связано с вовлечением математического аппарата, учитывающего как сложную геометрическую структуру объектов, так и различные эффекты взаимодействия и распространения электромагнитного поля в горных породах. Чрезвычайно актуальными в последние годы стали задачи геоэлектрики в геологических средах, имеющих высокие значения удельного электрического сопротивления, характерные для флюидонасыщенных карбонатных и высокобитуминозных терригенных коллекторов.

В последнее десятилетие наблюдается большой интерес к исследованию механизмов и установлению зависимостей электромагнитных свойств горных пород от возбуждаемого поля на основе измерений в скважинах, лабораторных исследований на образцах керна и математического моделирования [23; 27–34]. Дисперсия электрофизических параметров горных пород наблюдается и при изучении нефтегазовых скважин с применением высокочастотного электромагнитного каротажа. Здесь используется промежуточный диапазон частот (от 875 кГц до 14 МГц), где влияние на сигнал оказывает не только диффузионные, но и волновые процессы в среде [19–22]. Современное состояние и развитие ДК детально проанализировано С. М. Аксельродом [27]. Однако в связи со сложной структурой объекта на микро- и макромасштабах исследования, а именно его пространственной неоднородностью и контрастностью физических характеристик в областях моделирования, в теории электромагнетизма до настоящего времени не существует единого математически обоснованного подхода к решению этого класса задач.

Большие перспективы вовлечения в разработку глубокозалегающих юрских и палеозойских залежей обуславливают настоятельную необходимость опережающего развития геофизических технологий для этих слабо изученных объектов. В настоящей работе выполняется изучение возможностей новых методов электромагнитных зондирований с использованием математического моделирования в реалистичных постановках применительно к изучению новых типов залежей с трудноизвлекаемыми запасами.

### **Измеряемые сигналы и конструктивные характеристики прибора**

Для восстановления частотного спектра электрофизических параметров предлагается использовать широкий диапазон частот от 20 до 500 МГц, где влияние как удельного электри-

ческого сопротивления (УЭС), так и относительной диэлектрической проницаемости (ОДП), на измеряемые электромагнитные сигналы значительно.

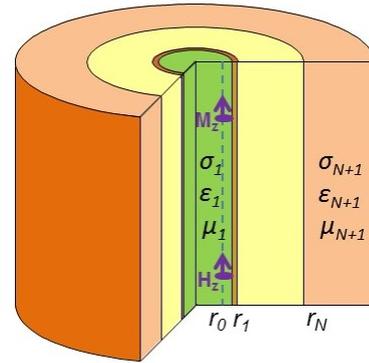
Датчики прибора прижаты к стенке скважины для того, чтобы увеличить чувствительность измеренных данных к параметрам пласта, особенно на высоких частотах и в проводящем буровом растворе, где сигналы быстро затухают. Генераторная и приемные катушки расположены соосно скважине; в двух приемных катушках измеряется разность фаз и затухание амплитуды э.д.с. (или магнитного поля). Длины зондов составляют от 0,1 до 0,8 м, база трехкатушечного зонда равна 0,2. Длины выбраны исходя из возможности технической реализации и с учетом минимально допустимого уровня э.д.с. в приемных катушках.

Данная конструкция допускает следующие упрощения при моделировании: источник электромагнитного поля можно считать точечным магнитным диполем, вычисление электромагнитного поля для дальнейшей трансформации в разность фаз и затухание амплитуды можно также производить в точке, характеризующей положение измерительной катушки.

### Решение прямой задачи

Рассмотрим задачу о поле вертикального магнитного диполя в цилиндрически-слоистой среде (рис. 1). Диполь находится в первом слое (скважине), плотность тока в нем изменяется по закону  $I(t) = I_0 e^{-i\omega t}$ .

Рис. 1. Модель среды: цилиндрические слои имеют радиусы  $r_j$ , проводимости  $\sigma_j$ , диэлектрические проницаемости  $\epsilon_j$  и магнитные проницаемости  $\mu_j$ . Зонд  $M_z H_z$  смещен с оси скважины, совпадающей с осью  $z$  цилиндрической или декартовой системы координат



Для решения задачи используется метод разделения переменных [35; 36]. В первом слое с источником электромагнитное поле представляется в виде суммы нормального и аномального поля:

$$\begin{cases} \vec{E} = \vec{E}_1 + \vec{E}_0, \\ \vec{H} = \vec{H}_1 + \vec{H}_0. \end{cases} \quad (1)$$

Поля  $\vec{E}_1, \vec{H}_1, \vec{E}_0, \vec{H}_0$ , а также поля в  $j$ -м слое  $\vec{E}_j, \vec{H}_j$  подчиняются уравнениям Максвелла:

$$\begin{cases} \text{rot } \vec{H}_0 = \gamma_1 \vec{E}_0, \\ \text{rot } \vec{E}_0 = i\omega\mu_1 \vec{H}_0 - \vec{j}^m. \end{cases} \quad (2)$$

$$\begin{cases} \text{rot } \vec{H}_j = \gamma_j \vec{E}_j, \\ \text{rot } \vec{E}_j = i\omega\mu_j \vec{H}_j. \end{cases} \quad (3)$$

Здесь  $\gamma_j = \sigma_j - i\omega\epsilon_j$  – комплексная проводимость,  $\vec{j}^m = (0, 0, j_z^m)$  – магнитный ток,  $j_z^m = -i\omega\mu_1 \cdot M_z \cdot U(P, P_0)$ ,  $M_z = I_0 \cdot S \cdot n_t$  – момент диполя,  $S$  – площадь генераторной катуш-

ки,  $n_i$  – число ее витков,  $U(P, P_0)$  – функция источника,  $P_0$  – координата источника,  $P$  – координата приемника,  $P_0 = (x_0, y_0, z_0)$ ,  $P = (x, y, z)$ .

На границах цилиндрических слоев тангенциальные компоненты электромагнитного поля непрерывны:

$$\begin{cases} [E_{r1}^*]_{r=r_1} = -E_{r0}(r_1), \\ [H_{r1}]_{r=r_1} = -H_{r0}(r_1), \\ [E_{\phi j}^*]_{r=r_j} = 0, \\ [H_{\phi j}]_{r=r_j} = 0, \quad j = 2, N. \end{cases} \quad (4)$$

Индексом « $t$ » обозначена  $\phi$ - или  $z$ -компонента поля. Квадратные скобки в формуле (4) означают скачок функции. Уравнения (2)–(4), а также условия убывания поля на бесконечности и конечности поля на оси скважины однозначно определяют электромагнитное поле в произвольной точке модели среды.

Будем искать аномальные поля  $\vec{E}_j, \vec{H}_j$ , предполагая, что нормальные поля  $\vec{E}_1, \vec{H}_1, \vec{E}_0, \vec{H}_0$  известны. Определим прямое и обратное преобразование Фурье:

$$\begin{cases} A^*(r, \phi, \lambda) = \int_{-\infty}^{\infty} A(r, \phi, z) e^{-i\lambda z} dz, \\ A(r, \phi, z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} A^*(r, \phi, \lambda) e^{i\lambda z} d\lambda. \end{cases} \quad (5)$$

Из уравнений (3) получаем, что Фурье-трансформации тангенциальных компонент выражаются через Фурье-трансформации вертикальных компонент следующим образом:

$$\begin{cases} E_{rj}^* = \frac{1}{p_j^2} \left( -i\lambda \frac{dE_{zj}^*}{dr} - \frac{i\omega\mu_j}{r} \frac{dH_{zj}^*}{d\phi} \right), \\ E_{\phi j}^* = \frac{1}{p_j^2} \left( i\omega\mu_j \frac{dH_{zj}^*}{dr} - \frac{i\lambda}{r} \frac{dE_{zj}^*}{d\phi} \right), \\ H_{rj}^* = \frac{1}{p_j^2} \left( -i\lambda \frac{dH_{zj}^*}{dr} - \frac{\gamma_j}{r} \frac{dE_{zj}^*}{d\phi} \right), \\ H_{\phi j}^* = \frac{1}{p_j^2} \left( \gamma_j \frac{dE_{zj}^*}{dr} - \frac{i\lambda}{r} \frac{dH_{zj}^*}{d\phi} \right), \end{cases} \quad (6)$$

$p_j^2 = \lambda^2 + k_j^2$ ,  $k_j^2 = -i\omega\mu_j\gamma_j$ , а последние подчиняются уравнениям

$$\begin{cases} \frac{1}{r} \frac{d}{dr} \left( r \frac{dE_{zj}^*}{dr} \right) + \frac{1}{r^2} \frac{d^2 E_{zj}^*}{d\phi^2} - p_j^2 E_{zj}^* = 0, \\ \frac{1}{r} \frac{d}{dr} \left( r \frac{dH_{zj}^*}{dr} \right) + \frac{1}{r^2} \frac{d^2 H_{zj}^*}{d\phi^2} - p_j^2 H_{zj}^* = 0. \end{cases} \quad (7)$$

Уравнения (4) преобразуются к виду  $r = r_1$ :

$$\left\{ \begin{array}{l} [E_{zj}^*]_{r=r_1} = -E_{z0}^*, \\ [H_{zj}^*]_{r=r_1} = -H_{z0}^*, \\ \left[ \frac{1}{p_j^2} \left( i\omega\mu_j \frac{dH_{zj}^*}{dr} - \frac{i\lambda}{r} \frac{dE_{zj}^*}{d\phi} \right) \right]_{r=r_1} = -\frac{1}{p_1^2} \left( i\omega\mu_1 \frac{dH_{z0}^*}{dr} - \frac{i\lambda}{r} \frac{dE_{z0}^*}{d\phi} \right), \\ \left[ \frac{1}{p_j^2} \left( \gamma_j \frac{dE_{zj}^*}{dr} - \frac{i\lambda}{r} \frac{dH_{zj}^*}{d\phi} \right) \right]_{r=r_1} = -\frac{1}{p_1^2} \left( \gamma_1 \frac{dE_{z0}^*}{dr} - \frac{i\lambda}{r} \frac{dH_{z0}^*}{d\phi} \right). \end{array} \right. \quad (8)$$

$r = r_j, j = 2, N :$

$$\left\{ \begin{array}{l} [E_{zj}^*]_{r=r_j} = 0, \\ [H_{zj}^*]_{r=r_j} = 0, \\ \left[ \frac{1}{p_j^2} \left( i\omega\mu_j \frac{dH_{zj}^*}{dr} - \frac{i\lambda}{r} \frac{dE_{zj}^*}{d\phi} \right) \right]_{r=r_j} = 0, \\ \left[ \frac{1}{p_j^2} \left( \gamma_j \frac{dE_{zj}^*}{dr} - \frac{i\lambda}{r} \frac{dH_{zj}^*}{d\phi} \right) \right]_{r=r_j} = 0. \end{array} \right. \quad (9)$$

Разложим  $E_{zj}^*, H_{zj}^*$  и  $E_{z0}^*, H_{z0}^*$  в ряды по угловой координате:

$$\left\{ \begin{array}{l} E_{zj}^* = \sum_{n=0}^{\infty} (e_{nj}^c \cos n\bar{\phi} + e_{nj}^s \sin n\bar{\phi}), \\ H_{zj}^* = \sum_{n=0}^{\infty} (h_{nj}^c \cos n\bar{\phi} + h_{nj}^s \sin n\bar{\phi}), \\ E_{z0}^* = \sum_{n=0}^{\infty} (e_{n0}^c \cos n\bar{\phi} + e_{n0}^s \sin n\bar{\phi}), \\ H_{z0}^* = \sum_{n=0}^{\infty} (h_{n0}^c \cos n\bar{\phi} + h_{n0}^s \sin n\bar{\phi}). \end{array} \right. \quad (10)$$

Здесь и далее  $\bar{\phi} = \phi - \phi_0$ ,  $x = r \cos \phi$ ,  $y = r \sin \phi$ ,  $x_0 = r_0 \cos \phi_0$ ,  $y_0 = r_0 \sin \phi_0$ ,  
 $R = \sqrt{r^2 + r_0^2 - 2rr_0 \cos \bar{\phi}}$ .

Подставляя уравнение (10) в (7)–(9), получаем уравнения (11)–(13), определяющие неизвестные функции  $e_{nj}^{c,s}, h_{nj}^{c,s}$  (угловые гармоники):

$$\left\{ \begin{array}{l} \left( \frac{1}{r} \frac{d}{dr} \left( r \frac{de_{nj}^{c,s}}{dr} \right) - \left( \frac{n^2}{r^2} + p_j^2 \right) e_{nj}^{c,s} \right) = 0, \\ \left( \frac{1}{r} \frac{d}{dr} \left( r \frac{dh_{nj}^{c,s}}{dr} \right) - \left( \frac{n^2}{r^2} + p_j^2 \right) h_{nj}^{c,s} \right) = 0. \end{array} \right. \quad (11)$$

$$\begin{cases} \left[ e_{nj}^{c,s} \right]_{r=r_1} = -e_{n0}^{c,s}, \\ \left[ h_{nj}^{c,s} \right]_{r=r_1} = -h_{n0}^{c,s}, \\ \left[ \frac{1}{p_j^2} \left( i\omega\mu_j \frac{dh_{nj}^{c,s}}{dr} \mp \frac{i\lambda n}{r} e_{nj}^{s,c} \right) \right]_{r=r_1} = -\frac{1}{p_1^2} \left( i\omega\mu_1 \frac{dh_{n0}^{c,s}}{dr} \mp \frac{i\lambda n}{r} e_{n0}^{s,c} \right), \\ \left[ \frac{1}{p_j^2} \left( \gamma_j \frac{de_{nj}^{c,s}}{dr} \mp \frac{i\lambda n}{r} h_{nj}^{s,c} \right) \right]_{r=r_1} = -\frac{1}{p_1^2} \left( \gamma_1 \frac{de_{n0}^{c,s}}{dr} \mp \frac{i\lambda n}{r} h_{n0}^{s,c} \right). \end{cases} \quad (12)$$

$$\begin{cases} \left[ e_{nj}^{c,s} \right]_{r=r_j} = 0, \\ \left[ h_{nj}^{c,s} \right]_{r=r_j} = 0, \\ \left[ \frac{1}{p_j^2} \left( i\omega\mu_j \frac{dh_{nj}^{c,s}}{dr} \mp \frac{i\lambda n}{r} e_{nj}^{s,c} \right) \right]_{r=r_j} = 0, \\ \left[ \frac{1}{p_j^2} \left( \gamma_j \frac{de_{nj}^{c,s}}{dr} \mp \frac{i\lambda n}{r} h_{nj}^{s,c} \right) \right]_{r=r_j} = 0, \quad j=2, N. \end{cases} \quad (13)$$

Решение уравнений (11) есть линейная комбинация модифицированных функций Бесселя [37]:

$$\begin{cases} e_{nj}^{c,s} = C_{nj}^{c,s} I_n(p_j r) + P_{nj}^{c,s} K_n(p_j r), \\ h_{nj}^{c,s} = D_{nj}^{c,s} I_n(p_j r) + Q_{nj}^{c,s} K_n(p_j r), \quad \text{Re } p_j > 0. \end{cases} \quad (14)$$

Используя условия  $E_{z1}^*, H_{z1}^* \Big|_{r \rightarrow 0} \rightarrow 0$  и  $E_{zN+1}^*, H_{zN+1}^* \Big|_{r \rightarrow \infty} \rightarrow 0$ , получаем  $P_1^{c,s} = Q_1^{c,s} = C_{N+1}^{c,s} = D_{N+1}^{c,s} = 0$ . Оставшиеся коэффициенты  $C_{nj}^{c,s}, D_{nj}^{c,s}, j=1, N$  и  $P_{nj}^{c,s}, Q_{nj}^{c,s}, j=2, N+1$  могут быть определены из граничных условий (12), (13). Способ определения угловых гармоник для нормального поля  $e_{n0}^{c,s}, h_{n0}^{c,s}$ , входящих в граничные условия, следующий.

Функция источника из уравнения (2) записывается в виде

$$U(P, P_0) = \delta(x - x_0) \delta(y - y_0) \delta(z - z_0).$$

К уравнению (2) применяется преобразование Фурье по всем координатам, в результате чего получается алгебраическое уравнение для  $H_{z0}^+$ , в то время как  $E_{z0}^+ = 0$ . Выполняя обратное преобразование Фурье по координатам  $(x, y)$  для полученного выражения для магнитного поля  $H_{z0}^+$  и разлагая в ряд определенное таким образом  $H_{z0}^*$ , имеем

$$\begin{cases} h_{n0}^c = -\frac{e^{-i\lambda z_0}}{2\pi} \cdot l_n \cdot p_1^2 \cdot M_z \cdot I_n(p_1 r_0) K_n(p_1 r), \quad l_0 = 1, l_n = 2, n \neq 0 \\ h_{n0}^s = 0, \\ e_{n0}^{c,s} = 0. \end{cases} \quad (15)$$

Учитывая (15), получаем  $C_{nj}^c = D_{nj}^s = 0, j=1, N$  и  $P_{nj}^c = Q_{nj}^s = 0, j=2, N+1$ . Чтобы определить ненулевые коэффициенты, используется уравнение (14), а (12) и (13) представляются в матричной форме:

$$\begin{aligned}\vec{X} &= \hat{V} \vec{\Psi}_{N+1}, \\ \hat{V} &= \hat{W}_1^{-1}(r_1) \hat{W}_2(r_1) \hat{W}_2^{-1}(r_2) \hat{W}_3(r_2) \dots \hat{W}_N^{-1}(r_N) \hat{W}_{N+1}[1,2](r_N).\end{aligned}\quad (16)$$

$$\hat{W}_j(r) = \begin{bmatrix} K_n(p_j r) & 0 & I_n(p_j r) & 0 \\ 0 & K_n(p_j r) & 0 & I_n(p_j r) \\ -\alpha_j K_n(p_j r) & \beta_j K_n'(p_j r) & -\alpha_j I_n(p_j r) & \beta_j I_n'(p_j r) \\ \varsigma_j K_n'(p_j r) & \alpha_j K_n(p_j r) & \varsigma_j I_n'(p_j r) & \alpha_j I_n(p_j r) \end{bmatrix}, \quad (17)$$

$$\hat{W}_j^{-1}(r) = p_j r \cdot \begin{bmatrix} I_n'(p_j r) & \frac{\alpha_j}{\varsigma_j} I_n(p_j r) & 0 & -\frac{1}{\varsigma_j} I_n(p_j r) \\ -\frac{\alpha_j}{\beta_j} I_n(p_j r) & I_n'(p_j r) & -\frac{1}{\beta_j} I_n(p_j r) & 0 \\ -K_n'(p_j r) & -\frac{\alpha_j}{\varsigma_j} K_n(p_j r) & 0 & \frac{1}{\varsigma_j} K_n(p_j r) \\ \frac{\alpha_j}{\beta_j} K_n(p_j r) & -K_n'(p_j r) & \frac{1}{\beta_j} K_n(p_j r) & 0 \end{bmatrix}. \quad (18)$$

$$\alpha_j = \frac{i\lambda n}{p_j^2 r}, \beta_j = \frac{i\omega\mu_j}{p_j}, \varsigma_j = \frac{\gamma_j}{p_j}.$$

$\hat{W}_{N+1}[1,2]$  – это первый и второй столбцы матрицы  $\hat{W}_{N+1}$ .

$$\vec{\Psi}_{N+1} = \begin{bmatrix} P_{nN+1}^s \\ Q_{nN+1}^c \end{bmatrix}, \quad \vec{X} = \begin{bmatrix} 0 \\ -M_z \frac{p_1^2}{\pi} I_n(p_1 r_0) \\ C_{n1}^s \\ D_{n1}^c \end{bmatrix}.$$

Множитель  $e^{-i\lambda z_0}$  учтен в последующем выражении (20).

Из уравнений (16) с использованием (17) и (18) определяются неизвестные коэффициенты  $D_{n1}^c$ , а соотношения (10) и (14) дают выражение для Фурье-трансформации магнитного поля в первом слое:

$$H_{z1}^* = \sum_{n=0}^{\infty} D_{n1}^c I_n(p_1 r) \cos n\bar{\phi}. \quad (19)$$

Применяя обратное преобразование Фурье (5) к  $H_{z1}^*$  и  $H_{z0}^*$ , а также используя четность подынтегральных функций, окончательно имеем:

$$\begin{aligned}H_z &= H_{z0} + \frac{M_z}{\pi} \int_0^{\infty} \left( \sum_{n=0}^{\infty} D_{n1}^c I_n(p_1 r) \cos n\bar{\phi} \right) \cos \lambda(z - z_0) d\lambda, \\ H_{z0} &= \frac{M_z e^{-k_1 \bar{R}}}{2\pi \bar{R}^3} \left[ 1 + k_1 \bar{R} - \frac{R^2}{\bar{R}^2} \left( 3 + 3k_1 \bar{R} + (k_1 \bar{R})^2 \right) \right], \quad \bar{R}^2 = R^2 + (z - z_0)^2, \operatorname{Re} k_j > 0.\end{aligned}\quad (20)$$

Отметим, что решение задачи об электромагнитном поле, возбуждаемом смещенным с оси скважины магнитным диполем, выполнено в [38].

На основе полученного решения разработан вычислительный алгоритм для быстрого моделирования электромагнитных сигналов диэлектрического каротажа и выполнена его программная реализация. Высокая точность вычислений и быстродействие компьютерной программы достигается путем представления полученного решения в виде, сокращающем число вычислительных операций, использования необходимой нормировки, предотвращающей экспоненциальный рост / затухание компонентов решения, схем эффективного интегрирования быстро осциллирующих слабозатухающих функций и суммирования медленно сходящихся рядов Фурье. Разработка программы сопровождается оценками точности и обязательным внутренним и внешним тестированием. Компьютерная программа написана на языке программирования Fortran 77 и работает под операционной системой Microsoft Windows.

### Моделирование сигналов диэлектрического зонда

Цель данного исследования – выяснить, как ведут себя электромагнитные сигналы в интересующем диапазоне значений электрофизических параметров, установить чувствительность сигналов к изменению параметров, что обеспечит в последующем возможность определения частотного спектра УЭС и ОДП в результате инверсии данных диэлектрического прибора.

Для анализа поведения сигналов при изменении УЭС и ОДП выбрана модель среды «скважина – пласт». УЭС скважины  $\rho_c = 2$  Ом·м, ОДП  $-\epsilon_c^* = 1$ , радиус  $-r_c = 0.108$  м. Учитывая конечный размер катушек, при моделировании источник и приемник электромагнитного поля были расположены на расстоянии 1 см от стенки скважины.

На рис. 2 приведены зависимости разности фаз и затухания амплитуд от УЭС и ОДП пласта для зонда 0,5 м на частотах 20–500 МГц.

На низкой частоте 20 МГц (рис. 2, а, б) преобладает зависимость измеряемых сигналов от сопротивления, чувствительность к диэлектрической проницаемости повышается при увеличении УЭС, особенно для разности фаз.

На частоте 100 МГц (рис. 2, в, г) для сопротивлений меньше 10 Ом·м разность фаз зависит как от УЭС, так и от ОДП, при увеличении сопротивления зависимость от УЭС практически исчезает, исключая только малые значения ОДП. Для затухания амплитуд зависимость от УЭС и ОДП наблюдается во всем диапазоне параметров, но для высоких значений УЭС и ОДП – на фоне небольших отрицательных значений сигнала.

Видимые артефакты (резкая смена цвета на рисунках, отражающая уровень сигнала) связаны с переходом значения э.д.с. через ноль.

На высокой частоте 500 МГц (рис. 2, д, е) зависимость разности фаз от УЭС еще более ослабевает. Затухание амплитуд чувствительно и к УЭС, и к ОДП. В целом на высоких частотах (100–500 МГц) разность фаз и затухание амплитуд ведут себя качественно по-разному, что видно при сравнении рисунков: разность фаз слабо зависит от сопротивления, а для амплитуды сохраняется зависимость от обоих параметров.

При уменьшении длины зонда снижается уровень сигнала, а качественное поведение сигналов сохраняется и соответствует описанному выше. Также на высокой частоте количество переходов э.д.с. через ноль сокращается – картина распределения уровня сигналов становится более гладкой.

При увеличении длины зонда уровень сигналов увеличивается, на высокой частоте количество переходов э.д.с. через ноль возрастает, а основные выводы о поведении сигналов остаются в силе.

Так как определение УЭС и ОДП происходит на одной частоте по измеренным разности фаз и затуханию амплитуд в нескольких трехкатушечных зондах, то, возможно, целесообразно иметь больше рабочих частот (пять-восемь) в диапазоне от 20 до 500 МГц для обеспечения устойчивости подбора электрофизических параметров с использованием сглаживания частотных спектров.

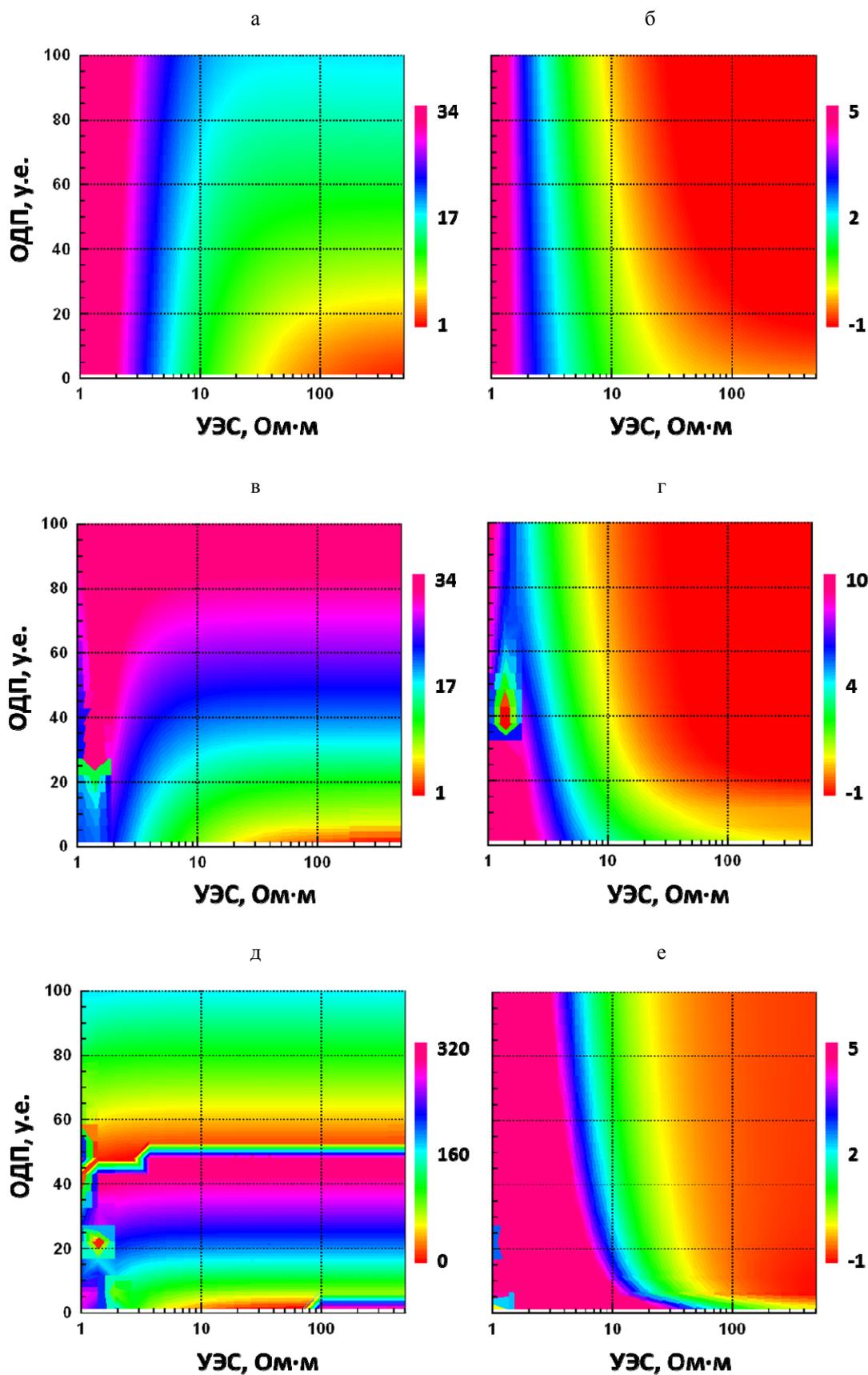


Рис. 2. Разность фаз (а, в, д) и затухание амплитуд (б, г, е) в зависимости от УЭС и ОДП пласта: а, б – зонд 0,5 м, частота 20 МГц; в, г – зонд 0,5 м, частота 100 МГц; д, е – зонд 0,5 м, частота 500 МГц

Таким образом, в выбранном частотном и геометрическом диапазонах конструктивных параметров прибора существует устойчивая зависимость измеряемых характеристик от электрофизических параметров пласта, пересекаемого скважиной. Разность фаз и затухание амплитуд демонстрируют качественно разное поведение, являясь, таким образом, независимыми измерениями, что важно для надежной инверсии данных диэлектрического зонда.

### Заключение

Для развития методов и создания новых приборов электромагнитных зондирований в скважинах выполнена разработка теоретической базы математического моделирования данных диэлектрического каротажа, необходимая для установления возможностей его применения к изучению залежей углеводородов с трудноизвлекаемыми запасами. Разработаны быстрые программно-алгоритмические средства численного моделирования высокочастотных электромагнитных сигналов на основе полученного решения прямой задачи диэлектрического каротажа в рамках слоисто-однородных моделей сред. Масштабное моделирование и сравнительный анализ амплитудно-фазовых характеристик высокочастотного электромагнитного поля позволили установить их устойчивую зависимость от частотного спектра электрофизических параметров в широком диапазоне частот – от 20 до 500 МГц. Расчеты показали, что на измеряемые электромагнитные сигналы значительное влияние оказывают УЭС, и ОДП. Разработанная теоретическая база и всесторонний анализ результатов численного моделирования в дальнейшем позволит выполнить научное обоснование параметров и оптимальное проектирование конфигурации новой электромагнитной зондирующей системы, предназначенной для изучения электрофизических свойств залежей углеводородов с трудноизвлекаемыми запасами.

### Список литературы

1. Аксельрод С. М. К измерению диэлектрической проницаемости горных пород в условиях скважины // Прикладная геофизика. 1968. Вып. 52.
2. Брылкин Ю. Л., Дубман Л. И. О диэлектрической проницаемости влажных песчаных пород // Электромагнитные методы исследования скважин. Новосибирск: Наука, 1979. С. 233-242.
3. Брылкин Ю. Л., Дубман Л. И. О диэлектрической проницаемости горных пород осадочного происхождения // Геология и геофизика. 1972. № 1. С. 117-121.
4. Даев Д. С., Денисов С. Б. О высокочастотном индукционном каротаже // Разведочная геофизика. М.: Недра, 1970. Вып. 42. С. 106-123.
5. Даев Д. С., Денисов С. Б., Костин А. И. Измерение удельного сопротивления пород методом волнового каротажа проводимости // Проблемы нефти и газа Тюмени: Науч.-техн. сб. Тюмень: ЗапСибНИГНИ, 1977. Вып. 33. С. 3-7.
6. Кауфман А. А., Антонов Ю. Н. Диэлектрический индуктивный каротаж. Новосибирск: Наука, 1971. 140 с.
7. Даев Д. С. Высокочастотные электромагнитные методы исследования скважин. М.: Недра, 1974. 192 с.
8. Денисов С. Б. Определение коллекторских параметров пластов по данным диэлектрического каротажа в карбонатных разрезах // Нефтегазовая геология и геофизика. 1981. Вып. 11. С. 29-32.
9. Денисов С. Б., Макагонова С. К. Интерпретация кривых волнового диэлектрического каротажа в пластах ограниченной мощности // Прикладная геофизика. М.: Недра, 1978. Вып. 92. С. 189-196.
10. Денисов С. Б., Макагонова С. К. Принципы интерпретации данных диэлектрического каротажа при измерении нескольких характеристик электрического поля // Нефтегазовая геология и геофизика. 1981. Вып. 2. С. 38-42.

11. *Панич И. М.* Исследование возможностей импедансных зондов и зондов с частотно-геометрической фокусировкой токов проводимости при диэлектрическом каротаже скважин: Автореф. дис. ... канд. техн. наук. Новосибирск, 1979. 19 с.
12. *Bittar M., Li J., Kainer G., Cherry R., Torres D., McCoy D.* A modern microwave formation evaluation sensor and its applications in reservoir evaluation // Transactions of the SPWLA 51th Annual Logging Symposium (June 19–23, 2010). Perth, Australia, 2010. Paper B.
13. *Hizem M., Budan H., Deville B., Faivre O., Mossé L., Simon M.* Dielectric dispersion: a new wireline petrophysical measurement // Proceedings of the SPE Annual Technical Conference and Exhibition (September 21–24, 2008). Denver, Colorado, USA, 2008. Paper 116130.
14. *Seleznev N., Habashy T., Boyd A., Hizem M.* Formation properties derived from a multi-frequency dielectric measurement // Transactions of the SPWLA 47th Annual Logging Symposium (June 4–7, 2006). Veracruz, Mexico, 2006. Paper VVV.
15. *Toumelin E., Torres-Verdin C., Bona N.* A new pore-scale framework for the simulation and interpretation of wide-band dielectric measurements // Abstracts of the 19th Annual International Symposium (August 21–25, 2005). Toronto, Canada, 2005. Paper SCA2005-21.
16. *Аксельрод С. М.* Влияние частотной дисперсии электрических свойств горных пород на результаты определения удельного сопротивления пластов (по материалам зарубежной литературы) // Каротажник. 2007. Вып. 10 (163). С. 103–126.
17. *Глинских В. Н., Никитенко М. Н., Эпов М. И.* Линеаризованные решения прямых и обратных двумерных задач высокочастотного электромагнитного каротажа в проводящих средах с учетом токов смещения // Геология и геофизика. 2013. Т. 54, № 12. С. 1942–1951.
18. *Глинских В. Н., Никитенко М. Н., Эпов М. И.* Моделирование и инверсия данных электромагнитного каротажа в пластах конечной мощности, вскрытых на биополимерных и нефтяных буровых растворах // Геология и геофизика. 2013. Т. 54, № 11. С. 1803–1813.
19. *Эпов М. И., Глинских В. Н.* Анализ пространственной чувствительности относительных характеристик в задачах высокочастотного электромагнитного каротажа // Геология и геофизика. 2005. Т. 46, № 11. С. 1168–1175.
20. *Эпов М. И., Глинских В. Н.* Быстрое двумерное моделирование высокочастотного электромагнитного поля для задач каротажа // Геология и геофизика. 2003. Т. 44, № 9. С. 942–952.
21. *Эпов М. И., Глинских В. Н.* Линеаризация относительных характеристик высокочастотного магнитного поля в двумерных проводящих средах // Геология и геофизика. 2004. Т. 45, № 2. С. 266–274.
22. *Эпов М. И., Глинских В. Н.* Электромагнитный каротаж: моделирование и инверсия. Новосибирск: Гео, 2005. 98 с.
23. *Anderson B., Barber T., Luling M., Sen P.* Observation of large dielectric effects on induction logs, or, can source rocks be detected with induction measurements // Transactions of the SPWLA 47th Annual Logging Symposium (June 4–7, 2006). Veracruz, Mexico, 2006. Paper OOO.
24. *Anderson B., Barber T., Luling M., Sen P., Taherian R., Klein J.* Identifying potential gas-producing shales from large dielectric permittivities measured by induction quadrature signals // Transactions of the SPWLA 49th Annual Logging Symposium (May 25–28, 2008). Edinburgh, Scotland, 2008. Paper HHHH.
25. *Anderson B., Barber T., Luling M., Rasmus J., Sen P., Tabanou J., Haugland M.* Observations of large dielectric effects on LWD propagation resistivity logs // Transactions of the SPWLA 48th Annual Logging Symposium (June 3–6, 2007). Austin, Texas, United States, 2007. Paper BB.
26. *Епов М. И., Сухорукова С. В., Глинских В. Н., Никитенко М. Н., Нецаев О. В., Суродина И. В.* Effective electromagnetic log data interpretation in realistic reservoir models // Open Journal of Geology. 2013. Vol. 3. No. 2B. P. 81–86.
27. *Аксельрод С. М.* Новые тенденции в диэлектрическом каротаже (по материалам зарубежной печати) // Каротажник. 2012. Вып. 4 (214). С. 78–112.
28. *Ельцов Т. И., Доровский В. Н., Ганеев Д. Н.* Низкочастотные диэлектрические спектры пород, насыщенных водонефтяной смесью // Геология и геофизика. 2014. Т. 55, № 8. С. 1270–1281.

29. Эпов М. И., Бобров П. П., Миронов В. Л., Репин А. В. Диэлектрическая релаксация в глинистых нефтесодержащих породах // Геология и геофизика. 2011. Т. 52, № 9. С. 1302–1309.
30. Эпов М. И., Миронов В. Л., Бобров П. П., Савин И. В., Репин А. В. Исследование диэлектрической проницаемости нефтесодержащих пород в диапазоне частот 0,05–16 ГГц // Геология и геофизика. 2009. Т. 50, № 5. С. 613–618.
31. Эпов М. И., Савин И. В., Миронов В. Л. Диэлектрическая температурная модель влажных нефтесодержащих пород в диапазоне частот от 0.5 до 15 ГГц // Геология и геофизика. 2012. Т. 53, № 7. С. 912–919.
32. Garrouch A., Alsafran E., Garrouch K. A classification model for rock typing using dielectric permittivity and petrophysical data // Journal of Geophysics and Engineering. 2009. Vol. 6. No. 3. P. 311–323.
33. Morales J., Trejo A., Cabrera M., Decoster E., Gea I. Schlumberger. Multifrequency dielectric tool identifies oil/water contacts within complex mexican carbonates in oil-based mud // Transactions of the SPWLA 56th Annual Logging Symposium (June 18–22, 2015). Long Beach, California, USA, 2015. Paper J.
34. Toumelin E., Torres-Verdin C. Pore-scale simulation of kHz-GHz electromagnetic dispersion of rocks: effects of rock morphology, pore connectivity, and electrical double layers // Transactions of the SPWLA 50th Annual Logging Symposium (June 21–24, 2009). The Woodlands, Texas, USA, 2008. Paper RRR.
35. Кауфман А. А., Каганский А. М., Кривоуцкий В. С. Радиальные характеристики индукционных зондов, смещенных с оси скважины // Геология и геофизика. 1974, № 7. С. 102–106.
36. Табаровский Л. А. Применение метода интегральных уравнений в задачах геоэлектрики. Новосибирск: Наука, 1975. 144 с.
37. Градштейн И. С., Рыжик И. М. Таблицы интегралов, рядов и произведений. М.: Физматгиз, 1963. 1100 с.
38. Nikitenko M., Itskovich G., Seryakov A. Fast electromagnetic modeling in cylindrically layered media excited by eccentric magnetic dipole // Radio Science. 2016. Vol. 51. P. 573–588.

Материал поступил в редколлегию 25.09.2018

V. N. Glinskikh<sup>1-3</sup>, M. N. Nikitenko<sup>1</sup>

<sup>1</sup> Trofimuk Institute of Petroleum Geology and Geophysics SB RAS  
3 Academician Koptuyug Ave., Novosibirsk, 630090, Russian Federation

<sup>2</sup> Novosibirsk State University  
1 Pirogov Str., Novosibirsk, 630090, Russian Federation

<sup>3</sup> Novosibirsk State Technical University  
20 K. Marx Ave., Novosibirsk, 630090, Russian Federation

GlinskikhVN@ipgg.sbras.ru, NikitenkoMN@ipgg.sbras.ru

## THEORETICAL-ALGORITHMIC BASE AND COMPUTER SIMULATION OF DIELECTRIC LOGGING DATA TO STUDY FREQUENCY SPECTRUM OF ELECTROPHYSICAL PARAMETERS OF GEOLOGICAL ENVIRONMENT

The article is devoted to the development of theoretical base for computer simulation of dielectric logging data with the aim of substantiating new import-substituting equipment for logging oil and gas wells. We have developed algorithms and implemented fast computer programs for the

simulation of high-frequency electromagnetic signals within the framework of homogeneously layered models of environment. Based on the results of large-scale simulation of the amplitude-phase characteristics of a high-frequency field, their dependence on the frequency spectrum of electrophysical parameters was established. The carried out research allows executing scientific substantiation and optimal design of a new electromagnetic sounding system.

*Keywords:* computer simulation, direct problem, geoelectric model, electrophysical parameters, frequency dispersion, dielectric logging.

## References

1. Akselrod S. M. Dielectric constant logging. *Prikladnaya Geofizika*, 1968, issue 52. (in Russ.)
2. Brylkin Yu. L., Dubman L. I. Dielectric constants of water-saturated sands. In: Antonov, Yu. N. (ed.). *Electromagnetic Logging Methods*. Novosibirsk, Nauka, 1979, p. 233–242. (in Russ.)
3. Brylkin Yu. L., Dubman L. I. Dielectric constants of sedimentary rocks. *Geologiya i Geofizika*, 1972, vol. 13, no. 1, p. 117–121. (in Russ.)
4. Daev D. S., Denisov S. B. High-frequency electromagnetic logging. *Razvedochnaya Geofizika*, 1970, vol. 42, p. 106–123. (in Russ.)
5. Daev D. S., Denisov S. B., Kostin A. I. Measurement of electrical resistivity of rocks by the method of wave conductivity logging. *Problems of oil and gas in Tyumen: scientific-technical transactions*. Tyumen, ZapSibNIGNI, 1977, issue 33, p. 3–7. (in Russ.)
6. Kaufman A. A., Antonov Yu. N. *Dielectric Constant Electromagnetic Logging*. Novosibirsk, Nauka, 1971, 140 p. (in Russ.)
7. Daev D. S. High-frequency electromagnetic methods of well logging. Moscow, Nedra, 1974, 192 p. (in Russ.)
8. Denisov S. B. Determination of reservoir parameters from dielectric logging data in carbonate cross sections. *Neftegazovaya geologiya i geofizika*, 1981, issue 11, p. 29–32. (in Russ.)
9. Denisov S. B., Makagonova S. K. Interpretation of wave dielectric logs in layers of finite thickness. *Prikladnaya geofizika*. Moscow, Nedra, 1978, issue 92, p. 189–196. (in Russ.)
10. Denisov S. B., Makagonova S. K. Principles of dielectric logging data interpretation when measuring several electric field characteristics. *Neftegazovaya geologiya i geofizika*, 1981, issue 2, p. 38–42. (in Russ.)
11. Panich I. M. Investigation of the capabilities of impedance tools and tools with frequency-geometric focusing of conduction currents during dielectric logging of wells. Abstract of dissertation of candidate in technical sciences. Novosibirsk, 1979, 19 p. (in Russ.)
12. Bittar M., Li J., Kainer G., Cherry R., Torres D., McCoy D. A modern microwave formation evaluation sensor and its applications in reservoir evaluation. *Transactions of the SPWLA 51th Annual Logging Symposium (June 19–23, 2010)*. Perth, Australia, 2010, paper B.
13. Hizem M., Budan H., Deville B., Faivre O., Mossé L., Simon M. Dielectric dispersion: a new wireline petrophysical measurement. *Proceedings of the SPE Annual Technical Conference and Exhibition (September 21–24, 2008)*. Denver, Colorado, USA, 2008, paper 116130.
14. Seleznev N., Habashy T., Boyd A., Hizem M. Formation properties derived from a multi-frequency dielectric measurement. *Transactions of the SPWLA 47th Annual Logging Symposium (June 4–7, 2006)*. Veracruz, Mexico, 2006, paper VVV.
15. Toumelin E., Torres-Verdín C., Bona N. A new pore-scale framework for the simulation and interpretation of wide-band dielectric measurements. *Abstracts of the 19th Annual International Symposium (August 21–25, 2005)*. Toronto, Canada, 2005, paper SCA2005-21.
16. Akselrod S. M. Frequency dependence of electric properties in rocks: effects on resistivity estimates (according to published evidence from outside Russia). *Karotazhnik*, 2007, no. 10, p. 103–126. (in Russ.)
17. Glinskikh V. N., Nikitenko M. N., Epov M. I. Processing high-frequency electromagnetic logs from conducting formations: linearized 2D forward and inverse solutions with regard to eddy currents. *Russian Geology and Geophysics*, 2013, vol. 54, no. 12, p. 1515–1521. (in Russ.)

18. Glinskikh V. N., Nikitenko M. N., Epov M. I. Numerical modeling and inversion of electromagnetic logs in the wells drilled with biopolymer and oil-mud. *Russian Geology and Geophysics*, 2013, vol. 54, no. 11, p. 1409–1416. (in Russ.)
19. Epov M. I., Glinskikh V. N. Space sensitivity of relative parameters in high-frequency electromagnetic logging. *Russian Geology and Geophysics*, 2005, vol. 46, no. 11, p. 1168–1175. (in Russ.)
20. Epov M. I., Glinskikh V. N. Fast two-dimensional simulation of high-frequency electromagnetic field in induction logging. *Russian Geology and Geophysics*, 2003, vol. 44, no. 9, p. 904–915. (in Russ.)
21. Epov M. I., Glinskikh V. N. Linearization of relative parameters of a high-frequency magnetic field in two-dimensional conducting media. *Russian Geology and Geophysics*, 2004, vol. 45, no. 2, p. 247–257. (in Russ.)
22. Epov M. I., Glinskikh V. N. Induction Logging: Forward Modeling and Inversion. Novosibirsk, Geo Publ., 2005, 98 p. (in Russ.)
23. Anderson B., Barber T., Luling M., Sen P. Observation of large dielectric effects on induction logs, or, can source rocks be detected with induction measurements. *Transactions of the SPWLA 47th Annual Logging Symposium (June 4–7, 2006)*. Veracruz, Mexico, 2006, paper OOO.
24. Anderson B., Barber T., Lüling M., Sen P., Taherian R., Klein J. Identifying potential gas-producing shales from large dielectric permittivities measured by induction quadrature signals. *Transactions of the SPWLA 49th Annual Logging Symposium (May 25–28, 2008)*. Edinburgh, Scotland, 2008, paper HHHH.
25. Anderson B., Barber T., Lüling M., Rasmus J., Sen P., Tabanou J., Haugland M. Observations of large dielectric effects on LWD propagation resistivity logs. *Transactions of the SPWLA 48th Annual Logging Symposium (June 3–6, 2007)*. Austin, Texas, USA, 2007, paper BB.
26. Epov M. I., Suhorukova C. V., Glinskikh V. N., Nikitenko M. N., Nechaev O. V., Surodina I. V. Effective electromagnetic log data interpretation in realistic reservoir models. *Open Journal of Geology*, 2013, vol. 3, no. 2B, p. 81–86.
27. Akselrod S. M. New trends in dielectric constant logging (according to published evidence from outside Russia). *Karotazhnik*, 2012, no. 4, p. 78–112. (in Russ.)
28. Eltsov T. I., Dorovsky V. N., Gapeev D. N. Dielectric spectra of water – oil – saturated porous media for the khz range and determination of volume fractions of system components. *Russian geology and geophysics*, 2014, vol. 55, no. 8, p. 1009–1018. (in Russ.)
29. Epov M. I., Bobrov P. P., Mironov V. L., Repin A. V. Dielectric relaxation in oil-bearing clayey rocks. *Russian Geology and Geophysics*, 2011, vol. 52, no. 9, p. 1028–1034. (in Russ.)
30. Epov M. I., Mironov V. L., Bobrov P. P., Savin I. V., Repin A. V. Dielectric spectroscopy of oil-bearing rocks at 0.05–16 GHz. *Russian Geology and Geophysics*, 2009, vol. 50, no. 5, p. 470–474. (in Russ.)
31. Epov M. I., Savin I. V., Mironov V. L. A temperature-dependent dielectric model for water- and oil-containing rocks in the frequency range from 0.5 to 15 GHz. *Russian Geology and Geophysics*, 2012, vol. 53, no. 7, p. 698–703. (in Russ.)
32. Garrouch A., Alsafran E., Garrouch K. A classification model for rock typing using dielectric permittivity and petrophysical data. *Journal of Geophysics and Engineering*, 2009, vol. 6, no. 3, p. 311–323.
33. Morales J., Trejo A., Cabrera M., Decoster E., Gea I., Schlumberger. Multifrequency dielectric tool identifies oil/water contacts within complex mexican carbonates in oil-based mud. *Transactions of the SPWLA 56th Annual Logging Symposium (June 18–22, 2015)*. Long Beach, California, USA, 2015, Paper J.
34. Toumelin E., Torres-Verdin C. Pore-scale simulation of kHz-GHz electromagnetic dispersion of rocks: effects of rock morphology, pore connectivity, and electrical double layers. *Transactions of the SPWLA 50th Annual Logging Symposium (June 21–24, 2009)*. The Woodlands, Texas, USA, 2008, Paper RRR.
35. Kaufman A. A., Kagansky A. M., Krivoputskiy V. S. Radial characteristics of induction tools displaced from the borehole axis. *Geology and geophysics*, 1974, no. 7, p. 102–106. (in Russ.)
36. Tabarovskiy L. A. Application of the integral equations method to geoelectric problems. Novosibirsk, Nauka, 1975, 144 p. (in Russ.)

37. Gradshtein I. S., Ryzhik I. M. Tables of integrals, series and products. Moscow, Fizmatgiz, 1963, 1100 p. (in Russ.)
38. Nikitenko M., Itskovich G., Seryakov A. Fast electromagnetic modeling in cylindrically layered media excited by eccentric magnetic dipole. *Radio Science*, 2016, vol. 51, p. 573–588.

*Received 25.09.2018*

*For citation:*

Glinskikh V. N., Nikitenko M. N. Theoretical-Algorithmical Base and Computer Simulation of Dielectric Logging Data to Study Frequency Spectrum of Electrophysical Parameters of Geological Environment. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 4, p. 41–55. (in Russ.) DOI 10.25205/1818-7900-2018-16-4-41-55

УДК 519.686, 519.876.5  
DOI 10.25205/1818-7900-2018-16-4-56-67

**С. С. Журавлев, С. В. Рудометов, В. В. Окольников, С. Р. Шакиров**

*Институт вычислительных технологий СО РАН  
пр. Академика Лаврентьева, 6, Новосибирск, 630090, Россия*

*s-zhur@yandex.ru, rsw@inbox.ru, okoln@mail.ru, ShakirovSR@ict.nsc.ru*

## **ПРИМЕНЕНИЕ МОДЕЛЬНО-ОРИЕНТИРОВАННОГО ПРОЕКТИРОВАНИЯ К СОЗДАНИЮ АСУ ТП ОПАСНЫХ ПРОМЫШЛЕННЫХ ОБЪЕКТОВ \***

Представлена реализация подхода «модельно-ориентированного проектирования» для решения задачи создания автоматизированных систем управления технологическими процессами (АСУ ТП) горнодобывающих предприятий.

Подход «модельно-ориентированного проектирования» заключается в организации контура тестирования, в котором испытуемый объект функционирует в среде, идентичной реальному процессу по входным сигналам. Подход основан на использовании математических моделей, имитирующих протекание технологического процесса под управлением АСУ ТП. Математические модели используются для формирования наборов тестовых данных.

Разработан имитационный программно-аппаратный комплекс. Основными частями комплекса являются: рабочая станция оператора SCADA, рабочая станция оператора модели, среда передачи данных, блок формирования физических сигналов, блок дублирования физических сигналов и программируемые логические контроллеры АСУ ТП. В качестве среды имитационного моделирования применена система MTSS.

Предложен и реализован вариант развития способа Hardware-in-the-loop, обеспечивающего создание совместных компонентов систем автоматизации.

Приведено описание методики тестирования АСУ ТП. Данная методика успешно применена для разработки, отладки и тестирования АСКУ ТО М (Автоматизированная система контроля и управления технологическим объектом во взрывозащищенном исполнении).

*Ключевые слова:* модельно-ориентированное проектирование, автоматизированная система управления технологическими процессами, промышленные системы управления, тестирование, прикладное программное обеспечение.

### **Введение**

Создание автоматизированных систем управления технологическими процессами для опасных промышленных объектов требует не только строгого соблюдения стандартов в этой области, но и принятия дополнительных мер со стороны изготовителя для обеспечения повышенной надежности систем<sup>1</sup>. Ошибки, допущенные на этапе проектирования системы и разработки прикладного программного обеспечения, могут приводить: на этапах пусконаладки и опытной эксплуатации к увеличению сроков и стоимости ввода системы автоматизации в работу; на этапе эксплуатации к возникновению нештатных и аварийных ситуаций.

---

\* Работа выполнена при частичной поддержке РФФИ (проект № 16-07-01179).

<sup>1</sup> О промышленной безопасности опасных производственных объектов // Федеральный закон № 116-ФЗ от 21.07.1997 (с изм. и доп. от 07.03.2017); Приказ федеральной службы по экологическому, технологическому и атомному надзору от 19 ноября 2013 г. № 550 «Об утверждении Федеральных норм и правил в области промышленной безопасности «Правила безопасности в угольных шахтах» (с изм. от 8.08.2017)».

*Журавлев С. С., Рудометов С. В., Окольников В. В., Шакиров С. Р.* Применение модельно-ориентированного проектирования к созданию АСУ ТП опасных промышленных объектов // Вестн. НГУ. Серия: Информационные технологии. 2018. Т. 16, № 4. С. 56–67.

В связи с этим возникает задача проверки надежности прикладного программного обеспечения АСУ ТП на предприятии изготовителе на всех этапах разработки. Это позволяет: повысить безотказность системы и, соответственно, безопасность технологического процесса; снизить расходы на пуско-наладку и опытно-промышленную эксплуатацию; упростить сопровождение, модернизацию и оптимизацию прикладного программного обеспечения.

### Подход МОП

Для тестирования прикладного программного обеспечения на всех этапах создания систем применяют подход МОП (модельно-ориентированное проектирование, англ.: model-based design approach)<sup>2</sup>, который основан на использовании моделей проектируемой системы и модели контролируемого процесса [1; 2]. Эти модели являются спецификацией проектируемой системы, которая актуализируется во время выполнения разработки. Данная спецификация позволяет выполнить анализ корректности и возможности выполнения требований посредством моделирования<sup>3</sup>.

Применение подхода МОП обеспечивает повышение надежности создаваемых программных, аппаратных и программно-аппаратных систем<sup>4</sup>. Появлению этого подхода предшествовали этапы становления общей теории систем, создание CASE (Computer-Aided Software Engineering) – совокупность средств автоматизации разработки программного обеспечения, разработка языка UML (Unified Modeling Language) в 1990-х гг., обеспечивающего проектирование систем на разных уровнях абстракций и появление в 2001 г. MDA (Model Driven Architecture)<sup>5</sup>.

Подход МОП основан на применении концепции «in-the-loop testing» [2]. Существуют следующие основные способы тестирования на базе подхода МОП<sup>6</sup>, в которых в качестве объекта тестирования может выступать: модель – Model-in-the-loop (MiL), программное обеспечение – Software-in-the-loop (SiL), прототип устройства – Processor-in-the-loop (PiL), готовое устройство – Hardware-in-the-loop (HiL). В способе HiL используется система целиком или ее компонент (данный способ применяется для организации приемочных тестов оборудования). Как правило, подход МОП используется совместно с «V-подходом» разработки программного обеспечения<sup>7</sup>.

Подход МОП получил широкое распространение, особенно в сфере автоматизации, его применяют известные во всем мире компании: Ford<sup>8</sup>, Tesla<sup>9</sup>, Festo<sup>10</sup> и др. Также подход МОП или его элементы (MiL, SiL, PiL, HiL, отдельно или совместно в разных комбинациях) активно применяется исследователями, например, в разработке генератора случайных чисел [3]; одновременном применении способов SiL и HiL для разработки программного обеспечения манипуляторов для роботов [4]; моделировании канала связи командно-измерительной

---

<sup>2</sup> Erkkinen T., Conrad M. Verification, Validation, and Test with Model-Based Design. 2008. URL: [https://www.mathworks.com/tagteam/53246\\_COMVEC%2008%20-%20VVnT%20with%20MBD.pdf](https://www.mathworks.com/tagteam/53246_COMVEC%2008%20-%20VVnT%20with%20MBD.pdf) (дата обращения 17.03.2018). DOI 10.4271/2008-01-2709

<sup>3</sup> Ефремов А. А., Сорокин С. С., Зенков С. М. Модельно-ориентированное проектирование – международный стандарт инженерных разработок. 2018. URL: <https://matlab.ru/upload/resources/EDU%20Conf/pp%2040-43%20Sorokin.pdf> (дата обращения 30.07.2018).

<sup>4</sup> Murphy B., Wakefield A. Early verification and validation using model-based design. The MathWorks. 2009. URL: <https://www.edn.com/design/integrated-circuit-design/4313371/Early-verification-and-validation-using-model-based-design> (дата обращения 28.07.18).

<sup>5</sup> Object Management Group (OMG). Model Driven Architecture. 2018. URL: <http://www.omg.org/mda/> (дата обращения 30.07.2018).

<sup>6</sup> Maibaum O. Comparison of Approaches to the Test of Control Unit Software. 2007. URL: [http://elib.dlr.de/46744/1/ics\\_maibaum.pdf](http://elib.dlr.de/46744/1/ics_maibaum.pdf) (дата обращения 14.09.2018).

<sup>7</sup> Donald Firesmith. Using V Models for Testing. Carnegie Mellon University Software Engineering Institute. URL: [https://insights.sei.cmu.edu/sei\\_blog/2013/11/using-v-models-for-testing.html](https://insights.sei.cmu.edu/sei_blog/2013/11/using-v-models-for-testing.html) (дата обращения 30.07.2018).

<sup>8</sup> Ranville S., Black P. Automated Testing Requirements – Automotive Perspective. *The Second Int. Workshop on Automated Program Analysis, Testing and Verification*, 2001. URL: <http://hissa.nist.gov/~black/Papers/autoTestReqs/WAPATV.rtf> (дата обращения 30.07.2018).

<sup>9</sup> Использование модельно-ориентированного проектирования для разработки родстера от Tesla. URL: <https://matlab.ru/success-story/tesla> (дата обращения 01.08.2018).

<sup>10</sup> Festo использует Модельно-Ориентированное Проектирование в разработке инновационных манипуляторов. URL: <https://matlab.ru/success-story/festo> (дата обращения 01.08.2018).

системы космического аппарата [5]; модельно-ориентированном проектировании системы автоматического управления температурой с циркуляцией промежуточного теплоносителя [6]; разработке системы управления электроприводом на основе метода модельно-ориентированного программирования [7]; быстром прототипировании систем управления как части модельно-ориентированной разработки теплового насоса сушилки [8]; управлении встроенной моделью в полунатурном моделировании для промышленных беспроводных сетей предприятий с применением WirelessHART стандарта [9], разработке сервиса задания сценариев предъявления стимулов с использованием модельно-ориентированного подхода [10] и др.

### **Развитие подхода МОП**

В Институте вычислительных технологий СО РАН (ИВТ СО РАН) разрабатываются автоматизированные системы управления технологическими процессами для шахт, рудников, объектов нефтегазовой отрасли и др. [11].

Функционально системы автоматизации, создаваемые в институте [12] для объектов горнодобывающей отрасли, можно разделить на две части – подземную (нижний уровень), обеспечивающую контроль, управление и сбор информации от различных технологических объектов шахты, и наземную (верхний уровень), выполняющую функции центрального вычислительного комплекса и рабочего места диспетчера. С развитием систем автоматизации в ИВТ СО РАН [12–14] на контроллеры АСУ ТП стали возлагаться все более сложные алгоритмические и функциональные задачи. Разработка новой системы автоматизации, в которую планируется включить компоненты, разработанные ранее для других систем или функционирующие на объекте, требует проведения проверки совместимости компонентов в рамках созданной единой системы автоматизации.

Интеграция новой системы автоматизации или ее компонентов требует внесения изменений в программное или аппаратное обеспечение существующей АСУ ТП, что во время пуско-наладочных работ на объекте замедляет процесс внедрения этих систем, а также может привести к аварийным ситуациям. В связи с этим на этапе проверки АСУ ТП необходимы не только отладка и тестирование, но и контроль совместимости созданных компонентов системы автоматизации с компонентами других версий и исполнений, функционирующих на действующем объекте.

### **Имитационный программно-аппаратный комплекс**

Разработка систем автоматизации требует выполнения тестов и проверок уже на этапе проектирования и разработки элементов системы. Выявленные на этих этапах недочеты и ошибки могут быть исправлены с минимальными дополнительными затратами. Сложность комплексной отладки и тестирования программ управления АСУ ТП заключается в трудоемкости искусственного формирования полного набора согласованных сигналов технологического оборудования.

Для разработки, отладки и тестирования прикладного программного обеспечения АСУ ТП применяют такие средства, как программные и физические имитаторы сигналов и интерфейсов. Но более эффективно применять специализированные комплексы, содержащие в своей структуре проблемно-ориентированные математические модели автоматизируемых технологических процессов, обеспечивающих согласованную генерацию сигналов датчиков и управляющих сигналов.

Существуют различные специализированные комплексы, например: имитационная модель гидроагрегата для тестирования АСУ ТП [15], тренажерно-управляющий программно-технический комплекс для объектов химической технологии [16], стенд для отладки ПО бортовой авиационной аппаратуры [17], продукт xPC Target – Matlab<sup>11</sup>, продукт WinMOD<sup>12</sup>

<sup>11</sup> xPC Target. 2017. URL: <http://matlab.ru/products/xpc-target> (дата обращения 27.01.2017).

<sup>12</sup> WinMOD system platform. Hennigsdorf: Mewes & Partner GmbH. 1995. URL: <http://winmod.com> (дата обращения 10.03.2015).

и др. К специализированным комплексам относится и разработка ИВТ СО РАН – имитационный программно-аппаратный комплекс для тестирования прикладного программного обеспечения АСУ ТП шахт и рудников [18]. Комплекс осуществляет тестирование программ управления прикладного программного обеспечения по принципу замещения реальных сигналов от датчиков существующего технологического оборудования и диспетчерских команд управления на виртуальные сигналы, сформированные на основе моделируемых параметров технологического оборудования.

Комплекс имеет двухуровневую структуру, содержащую программный и аппаратный уровни. Внешний вид комплекса изображен на рис. 1. Программный уровень содержит: SCADA (от англ. Supervisory Control And Data Acquisition – диспетчерское управление и сбор данных) систему БЛАКАРТ [19], среду имитационного моделирования MTSS [20], менеджер связи, модели технологических процессов, модели технологического оборудования, программы управления АСУ ТП.



Рис. 1. Внешний вид имитационного программно-аппаратного комплекса

Аппаратный уровень комплекса содержит: рабочую станцию оператора модели, рабочую станцию оператора SCADA, среду передачи данных, блок формирования физических сигналов и оборудование АСУ ТП (контроллеры, устройства ввода и вывода сигналов, датчики и т. п.). Блок формирования физических сигналов генерирует на основе модельных данных соответствующие аналоговые или дискретные электрические сигналы.

Имитационная модель учитывает особенности предметной области, а именно: параметры технологического оборудования и его составных частей (насосов, двигателей и т. п.), проектные данные о конфигурации технологического оборудования (настройки и уставки, набор датчиков и сигналов управления), параметры реальных технологических процессов.

Процесс разработки модели технологического оборудования (МТО) для среды моделирования MTSS заключается в создании концептуальной модели и ее переводе в машинное представление, а именно набор Java классов.

В общем виде МТО представлена набором множеств (взаимосвязь между ними изображена на рис. 2):

$$\text{МТО} = \{\text{SUBM}, \text{KTO}, \text{KACS}, \text{PIC}, \text{PORT}, \text{PARAM}, \text{VAR}, \text{M}, \text{COM}, \text{SIG}, \text{MS}, \text{I}, \text{STAT}\},$$

где МТО – объект, представляющей собой модель технологического оборудования с интегрированным прикладным программным обеспечением АСУ ТП; SUBM – набор моделей составных частей технологического оборудования (например, двигатель, редуктор и гидромуфта приводной станции конвейера и т. п.), детализирующий функционирование моделируемого объекта; KTO – алгоритм функционирования технологического оборудования, представляющий собой конечный автомат, описывающий состояния, в которых МТО может находиться в процессе имитации, KACS – конечный автомат программы управления АСУ ТП; PIC – условное графическое обозначение МТО; PORT – входные и выходные порты (каналы передачи данных, имеющие графическую часть на условном графическом обозначении модели), связывающую данную модель с другими МТО, входящими в имитационную модель системы; PARAM – параметры моделируемого объекта; VAR – переменные состояния объекта; M – математическое описание зависимостей между переменными и параметрами моделируемого объекта; COM – список команд управления технологическим оборудованием, при выполнении которых МТО переходит из одного состояния в другое, и сигналов управления составными частями технологического оборудования (например, двигатель приводной станции конвейера); SIG – хранилище сигналов датчиков и команд управления; MS – модели аналоговых датчиков, выполняющие функцию преобразования модельных сигналов к сигналам датчиков; I – интерфейсная часть, обеспечивающая взаимодействие между МТО и интерфейсной частью комплекса, где выполняется формирование цифровых тестовых сигналов в соответствии с конфигурацией АСУ ТП; STAT – интерфейс сбора статистики о функционировании МТО.

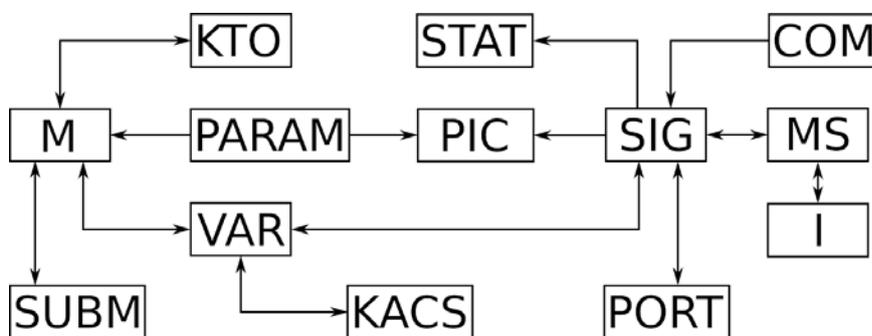


Рис. 2. Взаимосвязь между элементами структуры МТО

Значения сигналов датчиков формируются в процессе взаимодействия блоков конечных автоматов, описывающих алгоритмы функционирования технологического оборудования и АСУ ТП, и математического описания взаимосвязей между параметрами и переменными. Физические сигналы формируются алгоритмом интерфейсной части МТО, обеспечивающим сопряжение между моделью и преобразователями блока формирования физических сигналов. Этот алгоритм преобразует виртуальное значение переменной (например, скорость конвейерной ленты) из физической величины в кодовое значение, нормированное в диапазон значений датчика. Нормированное значение впоследствии преобразуется соответствующим преобразователем в физический сигнал, поступающий на вход тестируемой АСУ ТП.

Основная задача имитационного программно-аппаратного комплекса – это создание среды функционирования контроллеров полностью идентичной реальной по внешним сигналам.

Это достигается путем замены сигналов датчиков технологического оборудования на физические аналоги, численные значения которых формируются имитационной моделью.

Тестирование осуществляется подачей последовательности сигналов с заданными порядком, временными задержками, длительностями и амплитудами. Тестовую последовательность можно формализовать выражением

$$T_i = \{S_i, C_i, \tau_i\},$$

где  $T_i$  –  $i$ -я тестовая последовательность сигналов;  $S_i$  –  $i$ -е множество состояний сигналов датчиков;  $C_i$  –  $i$ -е множество команд управления;  $\tau_i$  – длительность  $i$ -го состояния;  $i$  – порядковый номер тестового набора данных, в котором  $i = 0 \dots n$ ;  $n$  – количество тестовых наборов данных;  $T_{i=0}$  – набор сигналов, задающий начальное состояние системы,  $T_{i=1 \dots (n-1)}$  – наборы сигналов, приводящие систему к тестовому состоянию,  $T_{i=n}$  – набор сигналов на входе системы, находящейся в тестовом состоянии.

Адекватность сгенерированной тестовой последовательности сигналов обуславливается точностью формирования этого набора сигналов и его временных параметров. Точность формирования сигналов определяется погрешностью вычисления виртуальных сигналов, преобразованием значений виртуальных сигналов в код, преобразованием кода в реальный сигнал в соответствующем блоке формирования физических сигналов и влиянием схемы коммутации комплекса и контроллера АСУ ТП. Применительно к токовому сигналу 0–20 мА, относительная погрешность генерации тестового сигнала составила не более 2%. Точность задания временных параметров важна с точки зрения выполняемых проверок АСУ ТП, которые в рассматриваемой предметной области сводятся к контролю пороговых значений (например, для параметра «время разгона конвейера» необходимо задавать значения задержки меньше и больше уставки).

Описанный выше комплекс применяется в ИВТ СО РАН при тестировании прикладного программного обеспечения систем автоматизации угольных шахт.

На созданном комплексе поставленная в статье задача была решена путем модификации способа HiL. Способ HiL<sup>13</sup> заключается в автономном тестировании оборудования АСУ ТП сигналами, числовые значения которых формируются с использованием математической модели, а преобразование модельных сигналов в физические осуществляется соответствующими преобразователями [21].

Способ модифицирован таким образом, что для тестирования контроллера наряду с математической моделью используется параллельно подключенный контроллер другой версии или модели. Это позволяет выполнить проверку и добиться идентичного функционирования тестируемого и эталонного контроллеров на одних и тех же входных сигналах и командах управления. На рис. 3 приведена схема усовершенствованного имитационного программно-аппаратного комплекса.

Модель системы генерирует виртуальные сигналы, которые поступают в блок формирования физических сигналов по интерфейсу RS-485 в формате строкового протокола. Блок, в свою очередь, формирует электрические аналоги сигналов датчиков и преобразует в цифровую форму сигналы управления тестируемого оборудования. Модуль дублирования физических сигналов обеспечивает дублирование сигналов аналоговых и дискретных датчиков. После чего сигналы датчиков синхронно подаются в Контроллер 1 (тестируемый) и Контроллер 2 (эталонный). Управляющие команды верхнего уровня АСУ ТП передаются в контроллеры через интерфейс RS-485 по протоколу Modbus. Преобразованные в цифровую форму сигналы управления передаются в модель системы, где в специализированном модуле осуществляется сравнение сигналов от Контроллера 1 и Контроллера 2, который в данном случае играет роль эталонного. Модель технологического процесса интерпретирует сигналы управления, полученные от тестируемого контроллера (Контроллер 1).

<sup>13</sup> Anders Frederiksen. Model-Based Design of Advanced Motor Control System. Analog Devices, Inc. Technical Article, MS-2577. 2017. URL: <http://www.analog.com/media/en/technical-documentation/technical-articles/Model-Based-Design-of-Advanced-Motor-Control-Systems-MS-2577.pdf> (дата обращения 19.07.2018).

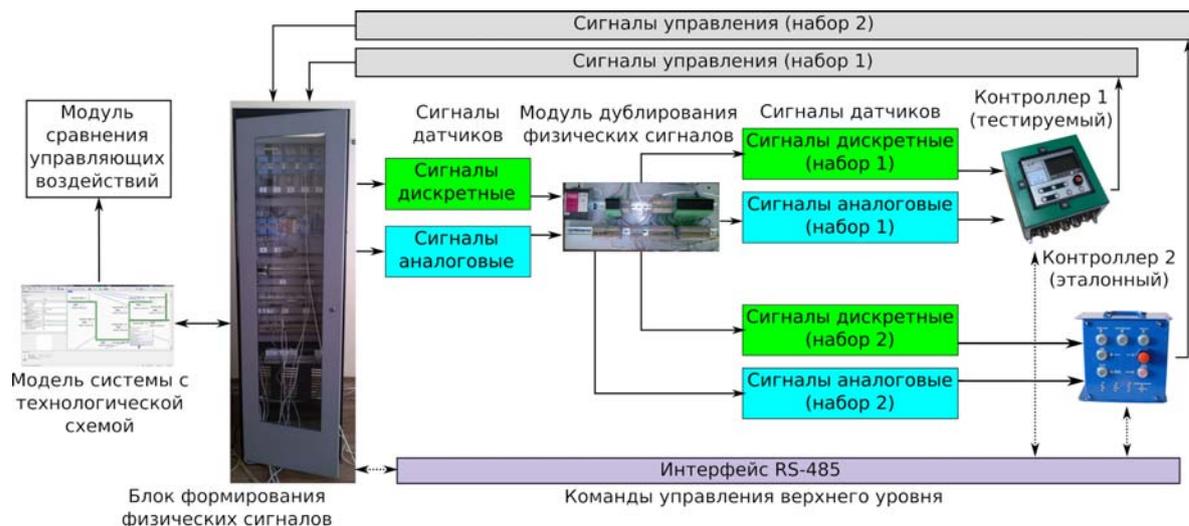


Рис. 3. Модифицированная структура комплекса, развивающая подход МОП

В отличие от классической реализации HiL в данной схеме добавлены модули дублирования физических сигналов и модуль сравнения управляющих воздействий, формируемых тестируемым и эталонным контроллерами.

Анализ идентичности функциональных возможностей контроллеров состоит из следующих шагов: запуск имитационной модели, генерирующей управляющие команды для контроллеров и сигналы датчиков; формирование управляющих сигналов, которые поступают непосредственно в контроллеры, и сигналов датчиков, которые дублируются и после этого поступают в контроллеры; сопоставление управляющих сигналов контроллеров в модуле сравнения управляющих воздействий. При идентичных входных воздействиях и различных выходных управляющих сигналах фиксируется ошибочная ситуация.

Осуществляется дублирование дискретных (кабель-троссовый выключатель, контроль схода ленты, устройство автоматического пожаротушения и т. п.) и аналоговых (датчик скорости ленты или барабана, датчик температуры, датчик заштыбовки и т. п.) сигналов. Реализовано дублирование частотных и токовых сигналов.

### Методика применения имитационного программно-аппаратного комплекса

Проверки при создании АСУ ТП делятся на несколько этапов, соответствующих стадиям проектирования АСУ ТП<sup>14</sup>: контроль проектных данных, отладка прототипов программно-аппаратных средств на уровне разработчика, первичное тестирование (крупномасштабное), приемочное тестирование (проверка серийного образца перед выпуском продукции), пусконаладочные работы, опытно-промышленная эксплуатация. Стадии создания АСУ ТП – проектирование, разработка, отладка и тестирование включают в себя создание модели функционирования системы автоматизации и модели технологического процесса, выполнение отладки и первичных тестов опытных образцов оборудования.

Контроль проектных данных – это проверка перед началом разработки прототипа на соответствие описания программно-аппаратных средств требованиям задачи – достаточна ли функциональность для выполнения задач автоматизации, корректно ли заданы параметры сигналов, корректно ли сформулированы требования для выполнения первичного и приемочного тестирований и др.

<sup>14</sup> ГОСТ 34.601-90 Комплекс стандартов на автоматизированные системы. Автоматизированные системы. Стадии создания. Дата введения 01.01.1992.

Отладка прототипов программно-аппаратных средств на уровне разработчика осуществляется с помощью локальных тестов, нацеленных на проверку корректности алгоритма управления.

Первичное тестирование (крупномасштабное) осуществляется в соответствии с разработанной программой и методикой испытаний. Выявленные ошибки или недочеты устраняются, и проводится повторное испытание.

Программа и методика испытаний разработана в соответствии с требованиями нормативных стандартов<sup>15</sup> и учетом условий эксплуатации АСУ ТП. Созданный документ регламентирует содержание тестовых испытаний и содержит следующие основные пункты.

1. Общесистемное тестирование, включающее проверку цепей электропитания, прохождения отдельных сигналов, исправное функционирование информационных линий, корректную работу индикации.

2. Проверка карты регистров протокола MODBUS [22], в том числе корректное отображение данных у диспетчера SCADA.

3. Контроль выполнения команд, заданных локально (с пульта контроллера) и удаленно (со SCADA или модели).

4. Контроль корректной идентификации контроллером состояний дискретных и аналоговых сигналов в каждом режиме работы, формирования соответствующих выходных управляющих сигналов и соответствующей записи состояния контроллера в карту MODBUS.

5. Проверка технологических ситуаций и достижимость состояний контроллера, включающая контроль типовых ситуаций (например, пуск и останов конвейера), контроль уставок контроллера (затянувшиеся разгон и торможение, пробуксовка, минимальная и максимальная скорости и др.), подача нештатных (например, сброс датчиков контроля двигателя и тормоза, разрешающего сигнала от предыдущего конвейера и др.) и некорректных сигналов (например, при рабочем значении скорости ленты заданы нулевые показания датчика скорости приводного барабана).

6. Проверка циклической подачи тестовых наборов сигналов в разных режимах и состояниях контроллера.

7. Тестирование интерфейса пользователя.

8. Технологический прогон (72 часа).

Приемочное тестирование направлено на контрольную проверку серийных изделий перед выпуском продукции, выявление сбоев программного обеспечения, вызванных причинами, не зависящими от разработчиков. Такие тесты в рамках текущей работы выполняются по сокращенной программе, содержащей пункты 1, 3, 4 и 8 разработанной методики.

По разработанной программе и методике испытаний проводилось тестирование совместимости созданной системы АСКУ ТО М (Автоматизированная система контроля и управления технологическим объектом во взрывозащищенном исполнении) и АСКУ ТО 2 (Автоматизированная система контроля и управления технологическим оборудованием).

В рамках работ по созданию автоматизированной системы конвейерной линии шахты Грамотеинская выявлено более 30 ошибок и замечаний к созданному прикладному программному обеспечению АСУ ТП. Выявленные ошибки в прикладном программном обеспечении исправлены до начала пуско-наладочных работ на объекте.

## Заключение

Подход МОП получил широкое распространение, особенно в сфере автоматизации. Наиболее часто он применяется для узкоспециализированных задач, таких как стендовые испытания, верификация программного обеспечения, проектирование алгоритмов и др.

Применение совокупности классических способов разработки, отладки и тестирования (MiL, SiL, PiL и HiL) подхода МОП наиболее эффективно для крупномасштабных систем. Такой вариант комплексного применения подхода МОП в основном используется крупными предприятиями.

---

<sup>15</sup> ГОСТ 19.301-79 Единая система программной документации. Программа и методика испытаний. Требования к содержанию и оформлению (с изм. № 1, 2) от 01.01.1981.

Применение подхода МОП позволяет упростить контроль над созданием программно-аппаратных систем на протяжении всего жизненного цикла разработки. Это сокращает время выпуска новой продукции, повышает ее качество и безопасность.

Предложен вариант развития способа HiL, позволяющий выполнять не только проверку создаваемых систем на наличие ошибок проектирования и разработки, но также обеспечить контроль совместимости компонентов двух разных систем одинакового назначения.

Программы тестов, входящих в созданную методику, разработаны с учетом охвата типовых технологических ситуаций при функционировании АСУ ТП, обеспечивая этим полноту тестирования контроллера.

Наиболее значительной по влиянию на безопасность АСУ ТП является проверка достижимости состояний контроллера, так как она позволяет выяснить, все ли состояния и переходы между ними заданы корректно.

Созданный комплекс с модифицированной структурой применен при разработке и тестировании системы АСКУ ТО М. Контроллеры нового поколения внедрены в состав действующей системы АСКУ ТО 2 и успешно эксплуатируются на шахте Грамотеинская.

Применение функции дублирования также позволяет при соответствующем расширении аппаратной части комплекса обеспечить серийное тестирование нескольких контроллеров АСУ ТП, что, в свою очередь, ускорит выполнение приемочных тестов.

### Список литературы

1. Деменков Н. П. Модельно-ориентированное проектирование систем управления // Промышленные АСУ и контроллеры. 2008. № 11. С. 66–69.
2. Ryssel U., Ploennigs J., Kabitzsch K., Folie M. Generative Design of Hardware-in-the-Loop Models // APGES'07. Salzburg, Austria, 2007. P. 4–11.
3. Бартенева В., Бартенева А. Модельно-ориентированное проектирование генератора случайных чисел // Современная электроника, 2014. № 2. С. 64–67.
4. Mossaad Ben Ayed, Lilia Zouari, Mohamed Abid. Software in the Loop Simulation for Robot Manipulators // Engineering, Technology & Applied Science Research. 2017. Vol. 7. No. 5. P. 2017–2021.
5. Арыков С. Б., Судьбин А. А., Шатров В. А., Рябушкин С. А., Вильданов А. И. Моделирование канала связи командно-измерительной системы космического аппарата // Решетневские чтения, 2015. Т. 2, № 19. С. 203–205.
6. Щербина Ю. В. Модельно-ориентированное проектирование системы автоматического управления температурой с циркуляцией промежуточного теплоносителя // Cloud of Science. 2015. Т. 2, № 4. С. 562–576.
7. Полющенко И. С. Разработка системы управления электропривода на основе метода модельно-ориентированного программирования // Вестник МЭИ. 2016. № 6. С. 87–95.
8. Jens Holtkötter, Jan Michael, Christian Henke, Ansgar Trächtler, Marcos Bockholt, Andreas Möhlenkamp, Michael Katter. Rapid-Control-Prototyping as part of Model-Based Development of Heat Pump Dryers // Procedia Manufacturing. 4th International Conference on System-Integrated Intelligence: Intelligent, Flexible and Connected Systems in Products and Production. 2018. Vol. 24. P. 235–242.
9. Chung Duc Tran, Rosdiazli Ibrahim, Vijanth Sagayan Asirvadam, Nordin Saad, Hassan Sabo Miya. Internal model control for industrial wireless plant using WirelessHART hardware-in-the-loop simulator // ISA Transactions. 2018. Vol. 75. P. 236–246.
10. Брак И. В., Сазонова Ю. И. Разработка сервиса задания сценариев предъявления стимулов с использованием модельно-ориентированного подхода // Вестн. НГУ. Серия: Информационные технологии. 2018. Т. 16, № 2. С. 31–40.
11. Голушко С. К., Меркулов И. В., Михальцов Э. Г., Чейдо Г. П., Шакиров Р. А., Шакиров С. Р. Индустриальные информационно-управляющие системы: от проектирования и разработки до практической реализации // Вычислительные технологии. 2013. Спец. вып. Т. 18. С. 4–11.
12. Благодарный А. И., Гусев О. З., Журавлев С. С., Золотухин Е. П., Каратышева Л. С., Колодей В. В., Михальцов Э. Г., Чейдо, Г. П., Шакиров Р. А., Шакиров С. Р. Автоматизиро-

ванная система контроля и управления ленточными конвейерами на угольных шахтах // Горная промышленность. 2008. № 5 (81). С. 38–44.

13. Патент на полезную модель № 100669 Российская Федерация, МПК7 H01H3/00. Пульт группового управления для работы в условиях подземной выработки угля / Гусев О. З., Иванов А. Е., Лесков С. К., Нарымский Б. В., Меркулов И. В., Шакиров Р. А.; заявитель и патентообладатель КТИ ВТ СО РАН. № 2010130020/07; заявл. 19.07.2010; опубл. 20.12.2010. Бюл. № 35. 1 с.

14. Патент на полезную модель № 133951 Российская Федерация, МПК7 G06F 9/00, G05B 19/00. Блок контроля и управления в условиях опасного производства / Гусев О. З., Колодей В. В., Мамаев А. С., Михальцов Э. Г., Шакиров С. Р.; заявитель и патентообладатель КТИ ВТ СО РАН. № 2013126010/08; заявл. 05.06.2013; опубл. 27.10.2013. Бюл. № 30. 1 с.

15. *Захарченко В. Е.* Имитационная модель гидроагрегата для тестирования алгоритмов АСУ ТП // Автоматизация в промышленности. 2007. № 7. С. 37–40.

16. *Сухарев А. В., Головушкин Б. А., Лабутин А. Н., Ерофеева Е. В.* Тренажерно-управляющий программно-технический комплекс для объектов химической технологии // Автоматизация в промышленности. 2011. № 7. С. 25–28.

17. *Бабаян Р. Р., Осипов А. Б.* Стенд для отладки ПО бортовой авиационной аппаратуры // Автоматизация в промышленности. 2013. № 5. С. 55–56.

18. *Журавлев С. С.* Программно-аппаратный комплекс для тестирования программ управления АСУ ТП шахт и рудников // Вычислительные технологии. 2013. Спец. вып. Т. 18. С. 150–155.

19. *Благодарный А. И.* Универсальная SCADA-система БЛАКАРТ под управлением операционной системы QNX // Проблемы информатики. 2009. № 3. С. 62–67.

20. *Рудометов С. В.* Визуально-интерактивная система имитационного моделирования технологических систем // Вестн. СибГУТИ. 2011. № 3. С. 14–27.

21. *Tatiana Kelemenová, Michal Kelemen, Lubica Miková, Vladislav Maxim, Erik Prada, Tomáš Lipták, František Menda.* Model Based Design and HIL Simulations // American Journal of Mechanical Engineering. 2013. Vol. 1. No. 7. P. 276–281. DOI 10.12691/ajme-1-7-25

22. *Томас Дж.* Введение в протокол Modbus. Часть 1 // СТА. 2009. № 2. С. 52–57.

*Материал поступил в редколлегию 06.09.2017*

**S. S. Zhuravlev, S. V. Rudometov, V. V. Okolnishnikov, S. R. Shakirov**

*Institute of Computational Technologies SB RAS  
6 Academician Lavrentiev Ave., Novosibirsk, 630090, Russian Federation*

*s-zhur@yandex.ru, rsw@inbox.ru, okoln@mail.ru, ShakirovSR@ict.nsc.ru*

## **MODEL-BASED DESIGN APPROACH FOR DEVELOPMENT PROCESS CONTROL SYSTEMS OF HAZARDOUS INDUSTRIAL FACILITIES**

The implementation of the "model-oriented design" approach for solving the problem of creating process control systems of mining enterprises is presented.

The "model-oriented design" approach is to organize a test loop in which the test object functions in an environment identical to the actual process, in terms of input signals. The approach is based on the use of mathematical models that simulate the functioning of the technological process under the control of the process control system. Mathematical models are used to form sets of test data.

A simulation software and hardware complex was developed. The main parts of the complex are: SCADA operator workstation, model operator workstation, data transmission layer, physical signal generation unit, physical signal duplication unit and programmable logic controllers of the process control system. As a simulation environment, the MTSS system is used.

A variant of the development of the Hardware-in-the-loop method, which provides the creation of compatible components of automation systems, is proposed and implemented.

The description of the method of testing the process control system is given. This technique has been successfully applied to the development, debugging and testing of the automated system for monitoring and controlling a technological object in explosion-proof design.

*Keywords:* model-based design, process control system, industrial control system, testing, application software.

## References

1. Demenkov N. P. Model-based design of control systems. *Industrial ACS and controllers*, 2008, No. 11, P. 66–69. (in Russ.)
2. Ryssel U., Ploennigs J., Kabitzsch K., Folie M. Generative Design of Hardware-in-the-Loop Models. APGES'07, Salzburg, Austria, 2007, p. 4–11.
3. Bartenev V., Barteneva A. Model-based design of random number generator. *Modern electronics*, 2014, no. 2, p. 64–67. (in Russ.)
4. Mossaad Ben Ayed, Lilia Zouari, Mohamed Abid. Software In the Loop Simulation for Robot Manipulators. *Engineering, Technology & Applied Science Research*, 2017, vol. 7, no. 5, p. 2017–2021.
5. Arikov S. B., Sudbin A. A., Shatrov V. A., Ryabushkin S. A., Vildanov A. I. Modeling of the communication channel for command-measuring systems of the spacecraft. *Reshetnevs reads*, 2015, vol. 2, no. 19, p. 203–205. (in Russ.)
6. Shcherbina Y. V. Model-based design of system of automatic control of the temperature of the circulating intermediate heat transfer. *Cloud of science*, 2015, vol. 2, no. 4, p. 562–576. (in Russ.)
7. Palyshenkov I. S. Development of control system of electric drive on the basis of model-oriented programming. *Vestnik MEI*, 2016, no. 6, p. 87–95. (in Russ.)
8. Jens Holtkötter, Jan Michael, Christian Henke, Ansgar Trächtler, Marcos Bockholt, Andreas Möhlenkamp, Michael Katter. Rapid-Control-Prototyping as part of Model-Based Development of Heat Pump Dryers. *Procedia Manufacturing. 4th International Conference on System-Integrated Intelligence: Intelligent, Flexible and Connected Systems in Products and Production*, 2018, vol. 24, p. 235–242.
9. Chung Duc Tran, Rosdiazli Ibrahim, Vijanth Sagayan Asirvadam, Nordin Saad, Hassan Sabo Miya. Internal model control for industrial wireless plant using WirelessHART hardware-in-the-loop simulator. *ISA Transactions*, 2018, vol. 75, p. 236–246.
10. Brak I. V., Sazonova Y. I. Development of the service of the task scenarios the presentation of stimuli using a model-based approach. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 2, p. 31–40. (in Russ.)
11. Golushko S. K., Merkulov I. V., Mikhaltsov E. G., Cheido G. P., Shakirov R. A., Shakirov S. R. Industrial information and control systems: from design and development to practical realization. *Computational technologies*, 2013, special issue, vol. 18, p. 4–11. (in Russ.)
12. Blagodarniy A. I., Gusev O. Z., Zhuravlev S. S., Zolotukhin E. P., Karatysheva L. S., Kolodey V. V., Mikhaltsov E. G., Cheido G. P., Shakirov R. A., Shakirov S. R. Automated system for monitoring and control of belt conveyors in coal mines. *Mining*, 2008, no. 5 (81), p. 38–44. (in Russ.)
13. The patent for useful model No. 100669 Russian Federation, IPC7 H01H3/00. Remote group control to work in the conditions of the underground workings of the coal / Gusev O. Z., Ivanov A. E., Leskov K. S., Narymskiy B. V., Merkulov I. V., Shakirov R. A.; applicant and patentee, DTI CT SB RAS. No. 2010130020/07; appl. 19.07.2010; publ. 20.12.2010, bul. no. 35, 1 p. (in Russ.)
14. The patent for useful model No. 133951 Russian Federation, IPC7 G06F 9/00, G05B 19/00. Unit monitoring and control in hazardous production / Gusev O. Z., Kolodey V. V., Mamaev A. S., Mihalov E. G., Shakirov S. R.; applicant and patent holder, DTI CT SB RAS. No. 2013126010/08; appl. 05.06.2013; publ. 27.10.2013, bul. no. 30, 1 p. (in Russ.)
15. Zakharchenko V. E. Simulation model of hydraulic unit for testing of ACS TP algorithms. *Automation in industry*, 2007, no. 7, p. 37–40. (in Russ.)

16. Sukharev A. V., Golovushkin B. A., Labutin A. N., Erofeeva E. V. Simulation and control software and hardware for chemical technology objects. *Automation in industry*, 2011, No. 7, p. 25–28. (in Russ.)
17. Babayan R. R., Osipov A. B. Stand for debugging on-board software of aviation equipment. *Automation in industry*, 2013, no. 5, p. 55–56. (in Russ.)
18. Zhuravlev S. S. Software and hardware complex for testing of control programs of ACS TP of mines. *Computational technologies*, 2013, special issue, vol. 18, p. 150–155. (in Russ.)
19. Blagodarniy A. I. Universal SCADA system BLAKART running under QNX operating system. *Problems of Informatics*, 2009, no. 3, p. 62–67. (in Russ.)
20. Rudometov S. V. Visually-interactive system of simulation of technological systems. *Vestnik SibGUTI*, 2011, no. 3, p. 14–27. (in Russ.)
21. Tatiana Kelemenová, Michal Kelemen, Lúbia Miková, Vladislav Maxim, Erik Prada, Tomáš Lipták, František Menda. Model Based Design and HIL Simulations. *American Journal of Mechanical Engineering*, 2013, vol. 1, no. 7, p. 276–281. DOI 10.12691/ajme-1-7-25
22. Thomas G. Introduction to the Modbus Protocol. Part 1. *MTA*, 2009, no. 2, p. 52–57. (in Russ.)

Received 06.09.2018

*For citation:*

Zhuravlev S. S., Rudometov S. V., Okolnishnikov V. V., Shakirov S. R. Model-Based Design Approach for Development Process Control Systems of Hazardous Industrial Facilities. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 4, p. 56–67. (in Russ.)  
DOI 10.25205/1818-7900-2018-16-4-56-67

**Ю. Г. Карин<sup>1</sup>, Е. В. Балков<sup>1</sup>, Д. И. Фадеев<sup>1,2</sup>, А. К. Манштейн<sup>1</sup>  
Г. Л. Панин<sup>1</sup>, А. О. Алымов<sup>2</sup>, Д. Б. Романов<sup>3</sup>**

<sup>1</sup> *Институт нефтегазовой геологии и геофизики им. А. А. Трофимука СО РАН  
пр. Академика Коптюга, 3, Новосибирск, 630090, Россия*

<sup>2</sup> *Новосибирский государственный университет  
ул. Пирогова, 1, Новосибирск, 630090, Россия*

<sup>3</sup> *ООО «Конструкторское бюро электрометрии»  
Москва, Россия*

*karinyg@ipgg.sbras.ru, balkovev@ipgg.sbras.ru, fadeevdi@ipgg.sbras.ru*

## **ЭЛЕКТРОМАГНИТНОЕ ПРОФИЛИРОВАНИЕ КОМПАКТНОЙ АППАРАТУРОЙ: НОВЫЙ ПОДХОД И РЕЗУЛЬТАТЫ ПРИМЕНЕНИЯ \***

Современные портативные приборы для электромагнитного профилирования и малоглубинного зондирования позволяют изучать распределение удельного электрического сопротивления грунта. К типичным сферам применения подобных комплексов можно отнести: исследование археологических объектов – курганы, древние поселения; решение экологических задач – изучение отвалов и хвостохранилищ горнодобывающих предприятий, зон распространения вредных веществ; локализация инженерных объектов – труб и коммуникаций, исследование состояния дамб и плотин. Перспективной сферой являются объекты агропромышленного комплекса – оценка плодородности почв (распределение минерализации), контроль эффективности мелиорации.

Целью проводимых исследований является разработка и создание переносного программно-аппаратурного комплекса электромагнитного профилирования, позволяющего с высокой достоверностью и точностью обнаруживать и дифференцировать по удельному электрическому сопротивлению объекты различного рода. На основании оригинального принципа компенсации первичного поля генераторной катушки была разработана компактная многочастотная аппаратура для электромагнитного профилирования. В результате прототипирования создана оптимальная конструкция корпуса, позволяющая реализовать данный принцип компенсации. Для эффективного использования аппаратного комплекса было создано программное обеспечение QZond, позволяющее управлять процессом сбора информации и в режиме реального времени визуализировать полученные данные. Проведенные полевые испытания на известных объектах электрометрического полигона ИНГТ СО РАН, при поиске локальных и протяженных проводников, а также археолого-геофизические исследования показывают эффективность реализации компактной аппаратуры электромагнитного профилирования.

*Ключевые слова:* геофизические методы поиска, электромагнитное профилирование, аппаратура, археологические исследования, инженерная геофизика, программное обеспечение.

В настоящее время широко известны приборы EM-31 и EM-38 для электромагнитного профилирования канадской фирмы «Geonics Ltd», аппаратура EMP-400 американской фирмы «GSSI» [1]. Малоглубинное частотное зондирование и профилирование позволяют выпол-

---

\* Исследовательская работа выполнена при частичной финансовой поддержке проекта РФФИ «Комплексные исследования археологических памятников Западной Сибири геофизическими методами: новые полевые технологии и способы интерпретации данных» (№ 17-29-04314).

*Карин Ю. Г., Балков Е. В., Фадеев Д. И., Манштейн А. К., Панин Г. Л., Алымов А. О., Романов Д. Б.* Электромагнитное профилирование компактной аппаратурой: новый подход и результаты применения // Вестн. НГУ. Серия: Информационные технологии. 2018. Т. 16, № 4. С. 68–78.

нять прибор GEM2 американской фирмы «Geophex Ltd» [2] и аппаратура ЭМС, разработанная в Институте нефтегазовой геологии и геофизики (ИНГГ) СО РАН [3; 4]. В работе [5] представлена классификация и подробное описание характеристик компактной аппаратуры индукционных исследований.

Перечисленные приборы, как правило, используются для решения малоглубинных задач в таких областях, как археология, поиск сходов и кладов, экология, инженерные изыскания, изучение объектов агропромышленного комплекса. Характерными отличиями искомых объектов в указанных сферах являются контраст по удельному электрическому сопротивлению (УЭС) в сравнении с вмещающей средой и небольшая глубина расположения (до 10 м). В случае археологических объектов их размеры в плане варьируются от нескольких сантиметров, например предметы быта древних людей, до нескольких десятков метров, как в случае с курганами. Чаще грунтовые могильники или остатки древних жилищ имеют заполнение материалом, схожим по УЭС с вмещающей средой, но из-за измененной структуры все равно удается выделять их на фоне нетронутой вмещающей среды [6]. Локальные объекты, такие как трубы, кабели и прочие инженерные конструкции, даже в случае малых размеров дают значительно больший вторичный сигнал при воздействии на них переменным магнитным полем [7]. Отходы горнообогатительных комбинатов чаще всего имеют низкое удельное сопротивление (1–20 Ом·м) из-за большого количества солей и высокоминерализованных растворов [8], а также большое распространение в плане.

### Способ реализации ЭМ профилирования

Компенсация первичного поля в таких устройствах как GEM-2 (разработка американской фирмы Geophex) и Nemfis (разработка ИНГГ СО РАН) [3] выполняется за счет использования компенсационной катушки [2]. Схема и условие компенсации продемонстрированы на рис. 1.

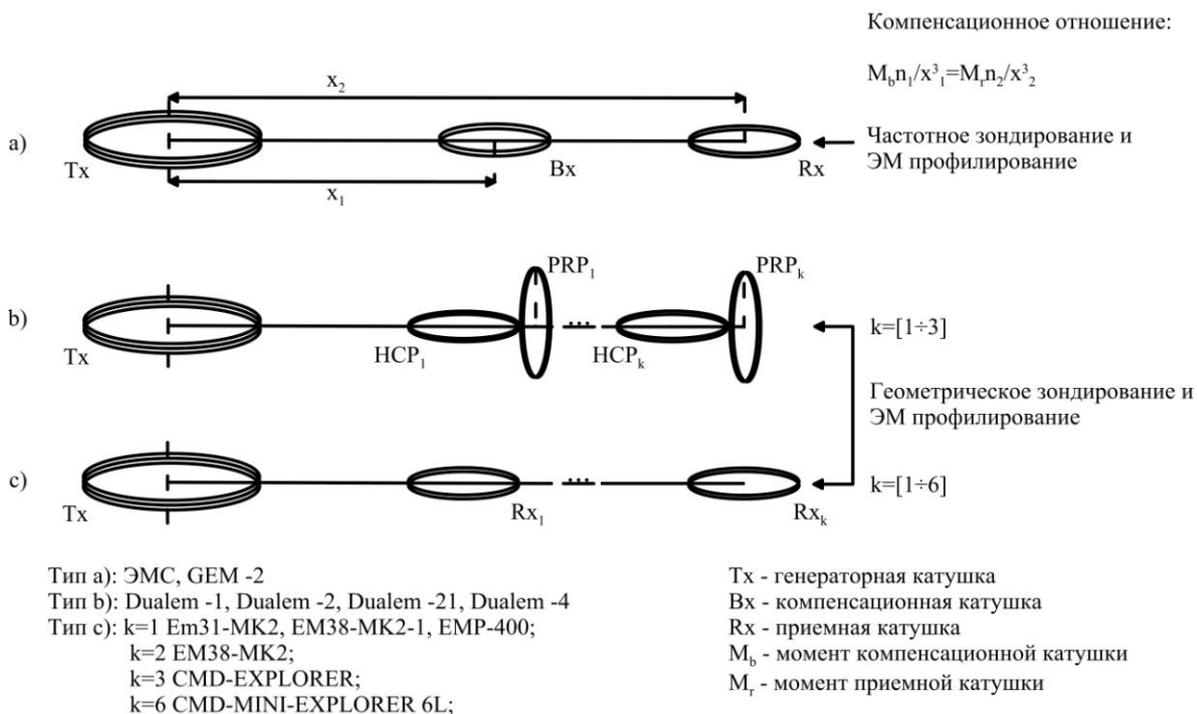


Рис. 1. Различные типы компоновки компактной малоглубинной аппаратуры индукционного зондирования и профилирования

Из-за сложного процесса настройки (патент RU 2 461 850), проблем с ложными аномалиями, характерных для трехкатушечных зондов, потери полезного сигнала при вычитании прямого поля актуальным остался вопрос о нахождении альтернативного метода компенсации поля генератора.

В ИНГГ СО РАН запатентован способ компенсации первичного поля генераторной катушки, особым расположением приемных катушек [9]. Структура поля магнитного диполя в изотропном пространстве имеет характерную линию, на которой вертикальная компонента напряженности магнитного поля меняет знак (рис. 2).

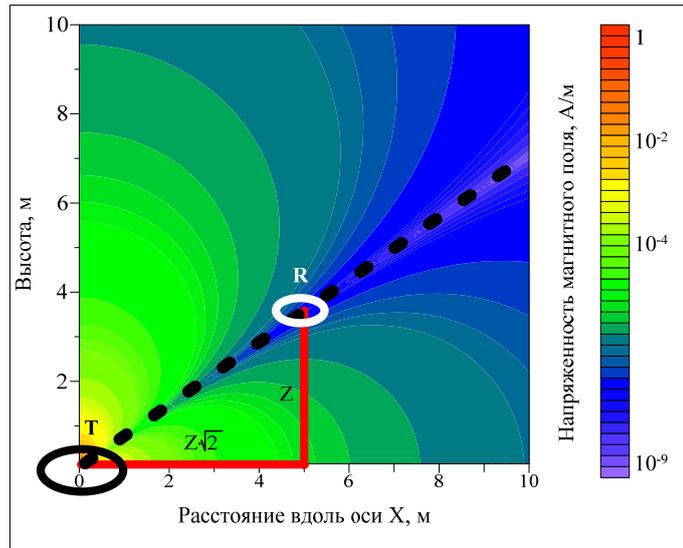


Рис. 2. Карта распределения вертикальной компоненты напряженности ЭМ поля от генератора (Т, приемник – R).

Уравнение этой поверхности легко получить из выражения для расчета напряженности поля магнитного диполя в однородном пространстве:

$$H_z^t = -\frac{M_t}{4\pi R^3} \left( \frac{3r^2}{R^2} + \frac{3kr^2}{R} + k^2 r^2 - 2 - 2kR \right) e^{-kR}, \quad (1)$$

где  $k^2 = i\omega\mu_0\sigma$ ;  $r$  и  $R$  – цилиндрический и сферический радиусы;  $M_t = I \cdot S \cdot n$  – момент генераторной катушки, представляющий собой произведение тока, площади и количества витков в катушке;  $\sigma$  – электропроводность пространства. Из выражения (1) при малых  $kr$  можно найти условие компенсации прямого поля:

$$r \approx \sqrt{2z},$$

где  $r$  – разнос,  $z$  – разность высот между приемной и генераторной катушками.

Это дало предпосылки к созданию прибора, у которого набор приемных катушек расположен в зоне минимального прямого поля (см. рис. 2) и на разном удалении от генератора. В такой реализации параметрами зондирования выступают частота генерируемого поля и расстояние между источником и приемником.

Разнесение генераторной и приемной катушек по вертикали приводит к однозначному расположению точки привязки сигнала – к центру генераторной катушки – и существенно уменьшает эффект возникновения нескольких экстремумов сигнала над приповерхностными объектами [10].

### Прототипирование аппаратуры

В результате макетирования и физического моделирования на электрометрическом полигоне ИНГГ СО РАН [7] были выбраны оптимальный разнос и рабочие частоты, которые по-

звоняют обнаруживать большинство мишеней полигона, имитирующих реальные объекты [11]. Авторами был разработан и изготовлен прототип компактной трехчастотной аппаратуры электромагнитного профилирования.

На рис. 3 приведены различные варианты исполнения новой аппаратуры от первого макета (а) до более современного прототипа (б). Все они выполнены из листового стеклотекстолита, имеют небольшой вес и компактные размеры. Несмотря на внешние отличия, любой из указанных прототипов в полной мере позволяет реализовать запатентованный способ компенсации прямого поля и за счет жесткого каркаса обеспечивает точное расположение генераторной и приемной катушек относительно друг друга.



Рис. 3. Макет аппаратуры электромагнитного профилирования (а).  
Прототип, выполненный из листового стеклотекстолита (б)

Генераторная катушка содержит две обмотки, резонансную и накачивающую, диаметр 270 мм, высота намотки 20 мм. Расстояние между генератором и приемником 70 см. Размер приемной катушки 42 мм, высота намотки 10 мм. Управляющие элементы аппаратуры расположены вдоль поверхности с минимальной напряженностью прямого поля генератора, тем самым минимизируется влияние на измеряемые данные.

### Программное обеспечение QZond

Взаимодействие с аппаратурой осуществляется через беспроводные протоколы Bluetooth и Wi-Fi смартфоном или планшетным компьютером. Для этих целей разработано и реализовано программно-алгоритмическое обеспечение, предназначенное для проектирования системы наблюдений, управления работой и предварительной обработки данных аппаратуры электромагнитных индукционных исследований. QZond (рис. 4) представляет собой приложение диалоговой структуры для ОС Android и сочетает следующие возможности: подготовка и редакция системы наблюдений; управление аппаратурой индукционных исследований; сбор данных GPS приёмника; предварительная обработка измеренных данных; визуализация в режиме реального времени; построение постобработанных карт, разрезов и псевдо трехмерное отображение данных. Программное обеспечение внедрено в пользование и является неотъемлемой частью аппаратурно-программных комплексов отечественного производства (ЭМС, Геовизер).

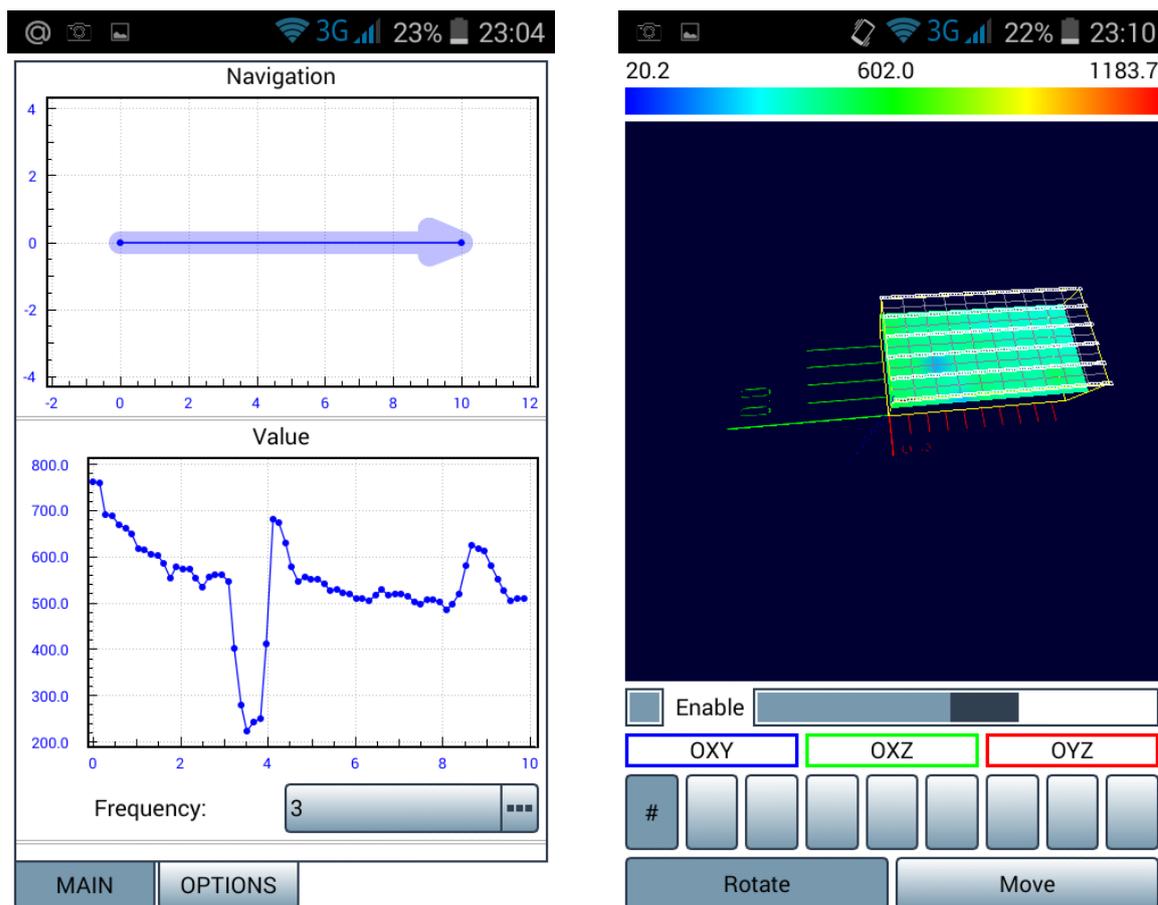


Рис. 4. Пример визуального интерфейса программы управления и первичной обработки данных QZond

## Результаты применения аппаратуры

Прототип аппаратно-программного комплекса был испытан на электрометрическом полигоне ИНГГ СО РАН [7]. Рисунок 5 демонстрирует результаты площадного профилирования на частоте 28 кГц. На исследованном участке на различной глубине заложены три металлические мишени с различной пространственной ориентацией. Все мишени находятся на глубине не более 1 м, поэтому обуславливают несколько экстремумов в сигнале. Горизонтально ориентированные объекты при исследовании прототипом аппаратуры проявляются в единственной (для объекта на глубине 0,8 м) либо одной преобладающей аномалии (объект на глубине 0,2 м). Вертикально ориентированный объект проявляется в виде нескольких аномалий различного «знака».

Результаты профилирования над алюминиевыми флягами и металлическими ёмкостями показаны на рис. 6. Две алюминиевые фляги зарыты на глубине 0,5 и 1 м, расположены на 3-м и 7-м метре соответственно на оси наблюдений. Железные бочки заложены на 12-м и 18-м метре по профилю, на глубинах 1 и 2,5 м соответственно.

Несмотря на общий высокий уровень сигнала (5 000 единиц АЦП – это шестая часть всего рабочего диапазона), все четыре бочки выделяются на профильной кривой. Однако уровень сигнала от бочек на глубине более 2 м приближается к уровню собственных шумов. Вероятнее всего, необходимо увеличивать отношение сигнал / шум в измерительной части аппаратуры.

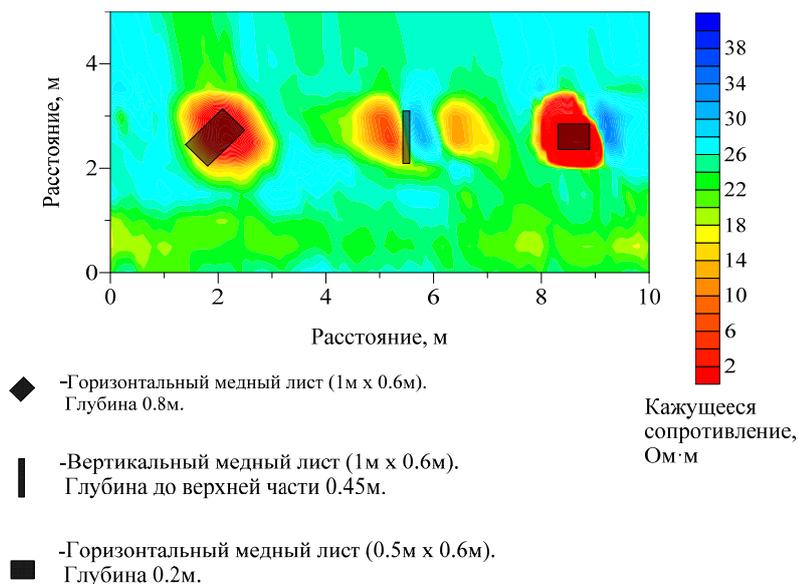


Рис. 5. Карта распределения кажущегося УЭС, полученная при помощи разработанной аппаратуры над тремя локальными проводящими объектами

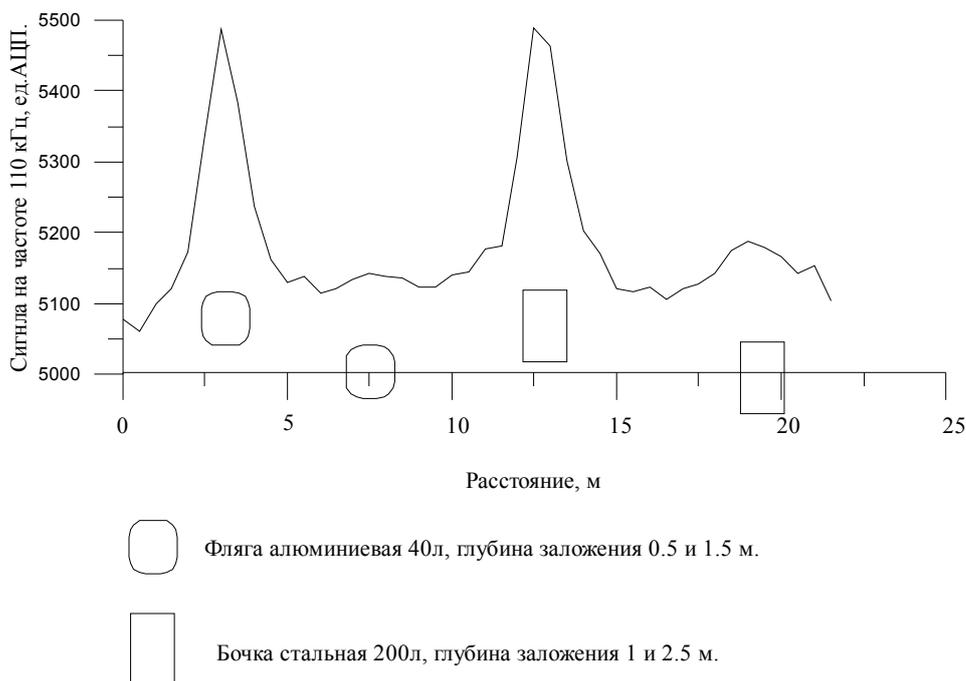


Рис. 6. Диаграмма сигнала аппаратуры, полученная в результате профилирования над металлическими флягами и ёмкостями на полигоне ИНГГ СО РАН

Практические работы по поиску металлических труб большого диаметра (до 1 м) демонстрируют эффективность применения разработки для решения такого типа задач. На рис. 7 приведены графики зависимости сигнала от расстояния вдоль исследуемых профилей.

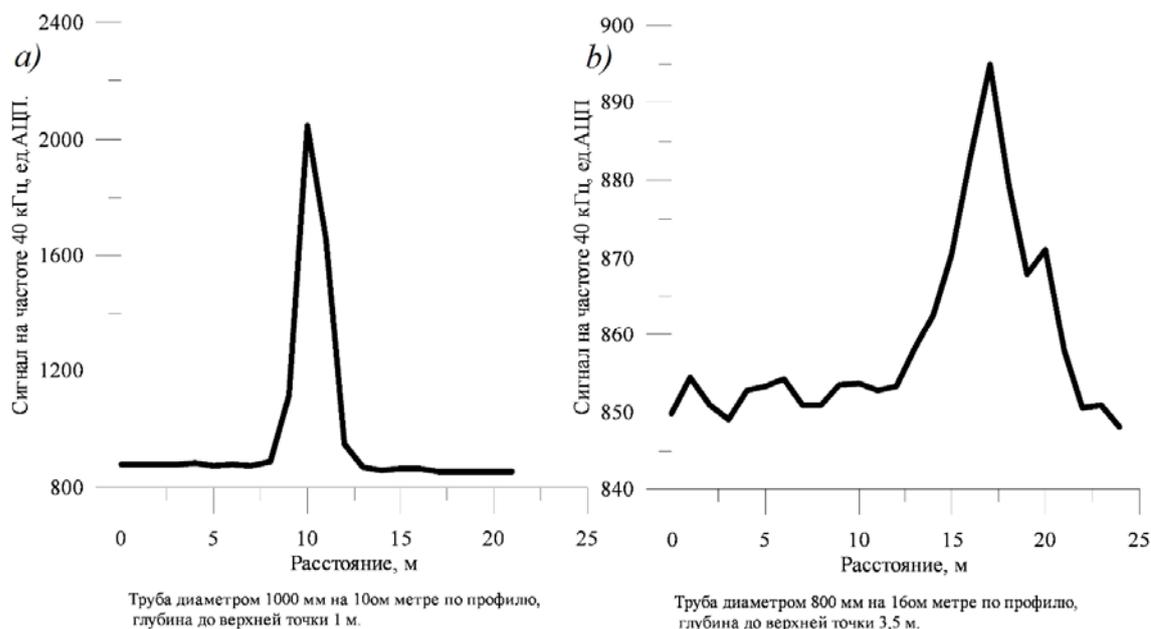


Рис. 7. Диаграммы сигнала разработанной аппаратуры, полученные в результате профилирования с целью поиска металлических труб на глубинах до 3,5 м (а) и 1 м (б). Графики для диаметров труб 1 и 0,8 м на рисунке (а) и (б) соответственно. Рабочая частота 40 кГц

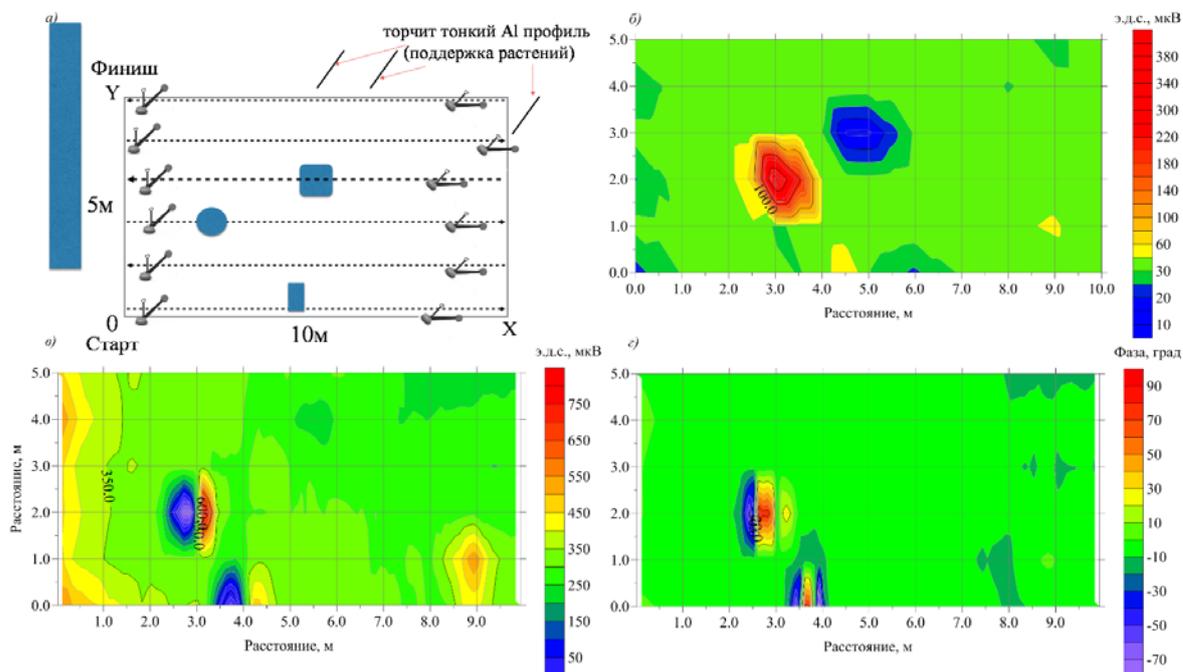


Рис. 8. Схема эксперимента (а) по поиску локальных объектов: квадрат – алюминиевый таз  $d = 30$  см с медной проволокой 4 кг на глубине 1 м; круг – катушка металлоискателя на поверхности; прямоугольник – стальной лист  $25 \times 50$  см толщиной 1,5 мм на поверхности, слева от площадки находится парник, каркас которого выполнен из алюминиевого профиля. Полученные карты распределения модуля э.д.с. (б) на частоте 40 кГц, реальной компоненты э.д.с. (в) и фазы сигнала (г) на частоте 111 кГц

Профильные кривые показывают наличие хорошо проводящих объектов на 10-м (рис. 7, а) и 16-м (рис. 7, б) метрах по измеряемым профилям, об этом свидетельствует резкое увеличение сигнала. В обоих случаях уровень полезного сигнала позволяет локализовать положение искомых объектов на профилях, что говорит о достаточной чувствительности прибора.

К тому же при отсутствии помех (постройки, металлический мусор) трубы являются высококонтрастной и яркой мишенью.

Был проведен ряд экспериментальных измерений над локальными проводящими объектами. На рис. 8 изображена схема эксперимента и заложенные мишени, представлены результаты измерений в виде карт распределения модуля э.д.с., реальной компоненты и фазы сигнала. Наиболее контрастно на всех картах выделяется аномалия, вызванная замкнутой многовитковой катушкой металлоискателя (локальные координаты мишени (3 м, 2 м)). Металлический лист (координаты (4 м, 0,3 м)) выделяется на картах распределения реальной компоненты и фазы сигнала (см. рис. 8, в, з). Аномалия, соответствующая алюминиевому тазу (координаты (5 м, 3 м)) отчетливо выделяется на более низкой частоте (40 кГц, рис. 8, б) в модуле э.д.с. Наибольшее негативное влияние нецелевых объектов (алюминиевый каркас парника слева от исследуемой площадки) наблюдается на карте распределения реальной компоненты э.д.с. (рис. 8, в), влияние на фазу сигнала минимально (рис. 8, з).

Что касается археологических задач, то и в этой области аппаратура показывает неплохие результаты. На рис. 9 показаны карты распределения сигнала от среды. Целью работ являлось картирование остатков древних построек. Работы проводились в Сузунском районе Новосибирской области вблизи р. Слезянка.

Исследовалась площадка размером 50 × 44 м. На рис. 9, а изображена карта распределения сигнала по данным аппаратуры индукционных исследований ЭМС, пунктирными линиями показаны контуры строений, выделяющиеся визуально в рельефе, как углубления грунта на глубину не более 20 см в самом центре. Видно, что аномалии, обозначенные голубым цветом, неплохо совпадают с этими контурами, следовательно, аномалии соответствуют остаткам древних сооружений.

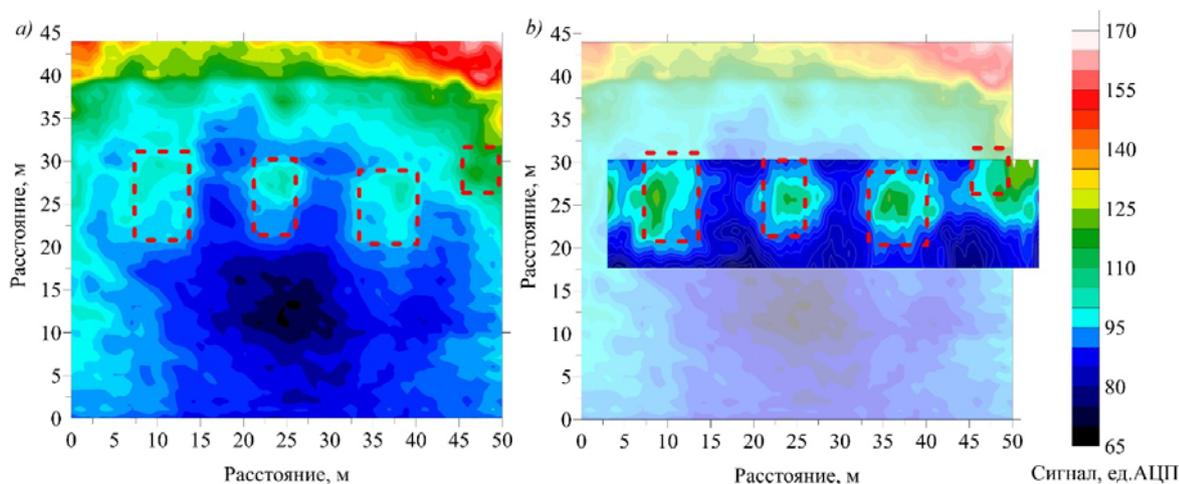


Рис. 9. Карты распределения сигнала по исследуемой площади аппаратурой ЭМС (а), фрагмент карты по данным новой аппаратуры (б)

На рис. 9, б показан фрагмент карты, полученной по данным разработанной аппаратуры. Видно, что аномалии сигнала, зеленые и голубые цвета на фоне синего, также неплохо совпадают с контурами западин. Следовательно, можно сделать вывод, что аппаратура «Геовизер» не уступает по своим характеристикам аппаратуре ЭМС, а наблюдаемые в данном случае аномалии более контрастны.

### Заключение

На основании оригинального запатентованного в ИНГГ СО РАН принципа компенсации первичного поля генераторной катушки в результате цикла прототипирования разработана

компактная многочастотная аппаратура для электромагнитного профилирования. Проведенные полевые испытания на известных объектах электрометрического полигона ИНГГ СО РАН, при поиске локальных и протяженных проводников, а также археолого-геофизические исследования показывают высокую эффективность разработанной аппаратуры.

### Список литературы

1. *Ortuani B., Chiaradia E. A., Priori S., L'Abate G., Canone D., Mele M., Comunian A., Giudici M., Facchi A.* Comparing EM38 and Profiler-EMP400 for the Delineation of Homogeneous Management Zones within Agricultural Fields // *Extended Abstracts, First Conference on Proximal Sensing Supporting Precision Agriculture*. Barcelona, 2015. DOI 10.3997/2214-4609.201413828
2. *Won I. J., Keiswetter D. A., Fields G. R. A., Sutton L. C.* GEM-2: A new multifrequency electromagnetic sensor // *Journal of Environmental and Engineering Geophysics*. 1996. Vol. 2 (1). P. 129–138.
3. *Маништейн А. К., Панин Г. Л., Тикунов С. Ю.* Аппаратура частотного электромагнитного зондирования «ЭМС» // *Геология и геофизика*. 2008. Т. 49, № 6. С. 571–579.
4. Балков Е. В. Технология малоглубинного частотного зондирования // *Геофизика*. 2011. № 6. С. 42–47.
5. Балков Е. В., Фадеев Д. И., Карин Ю. Г., Маништейн А. К., Маништейн Ю. А., Панин Г. Л. Новый подход к малоглубинным электромагнитным зондированиям // *Геология и геофизика*. 2017. Т. 58, № 5. С. 783–791.
6. Эпов М. И. и др. Археогеофизические исследования в СО РАН, 2013–2015 годы // *Мультидисциплинарные методы в археологии: новейшие итоги и перспективы: Материалы международного симпозиума*. Новосибирск, 2017. С. 405–413.
7. Балков Е. В., Стойкин Т. А., Маништейн А. К., Карин Ю. Г. Результаты применения малоглубинного электромагнитного профилирования на электрометрическом полигоне Института нефтегазовой геологии и геофизики СО РАН // *Геофизические исследования*. 2013. Т. 14, № 3. С. 55–63.
8. Эпов М. И. и др. Определение состава горно-рудных отходов геохимическими и геофизическими методами (на примере хвостохранилища Салаирского горно-обогатительного комбината) // *Геология и геофизика*. 2017. Т. 58, № 12. С. 1944–1954.
9. Маништейн А. К., Балков Е. В. Способ и устройство для индукционного частотного зондирования. Патент на изобретение № 2502092 от 20.12.2013.
10. Балков Е. В. Множественные образы в сигнале от локальных объектов при электромагнитном профилировании компактным зондом с разнесенными катушками // *Всероссийская школа-семинар имени М. Н. Бердичевского и Л. Л. Ваньяна по электромагнитным зондированиям Земли*. СПб., 2011. 4 с.
11. Балков Е. В., Халатов С. Ю., Маништейн А. К., Панин Г. Л. Малоглубинное электромагнитное профилирование компактной аппаратурой: теоретические оценки и результаты экспериментальных работ // *Инженерная геофизика – 2014: 10-я юбилейная конференция и выставка (Геленджик, 21–25 апреля 2014 г.)*. Геленджик, 2014.

**Yu. G. Karin**<sup>1</sup>, **E. V. Balkov**<sup>1</sup>, **D. I. Fadeev**<sup>1,2</sup>, **A. K. Manstein**<sup>1</sup>  
**G. L. Panin**<sup>1</sup>, **A. O. Alymov**<sup>2</sup>, **D. B. Romanov**<sup>3</sup>

<sup>1</sup> *Trofimuk Institute of Petroleum Geology and Geophysics SB RAS  
3 Academician Koptyug Ave., Novosibirsk, 630090, Russian Federation*

<sup>2</sup> *Novosibirsk State University  
1 Pirogov Str., Novosibirsk, 630090, Russian Federation*

<sup>3</sup> *Ltd "Design Bureau of Electrometry"  
Moscow, Russian Federation*

*karinyug@ipgg.sbras.ru, balkovev@ipgg.sbras.ru, fadeevdi@ipgg.sbras.ru*

## **ELECTROMAGNETIC PROFILING BY COMPACT APPARATUS: NEW APPROACH AND RESULTS OF APPLICATION**

Modern portable devices for electromagnetic profiling and shallow sounding make it possible to study the distribution resistivity of the soil. Typical areas of application of such complexes: research of archaeological objects – burial mounds, ancient settlements; the solution of environmental problems – the study of dumps and tailings of mining enterprises, zones of distribution of harmful substances; localization of engineering facilities – pipes and communications, research of the state of dams and dams. The perspective area is the objects of the agro-industrial complex – assessment of soil fertility (mineralization distribution), control of melioration efficiency.

The purpose of the conducted research is the development and creation of a portable software and hardware complex of electromagnetic profiling, which allows to detect and differentiate objects of various kinds with high reliability and accuracy with respect to specific electrical resistance. Based on the original principle of compensation of the primary field of the generator coil, a compact multi-frequency equipment for electromagnetic profiling was developed. As a result of the iterative prototyping cycle, the optimal body design is realized, which allows to realize this compensation principle. To effectively use the hardware complex, QZond software was created, which allows to control the process of collecting information and to visualize the data in real time. The conducted field tests on known objects of the electrometric polygon of IPGG SB RAS, when searching for local and extended conductors, as well as archaeological and geophysical studies show the efficiency of the implementation of compact electromagnetic profiling equipment.

*Keywords:* geophysical methods, electromagnetic profiling, equipment, archeology, engineering geophysics, high-precision positioning.

### **References**

1. Ortuani, B., Chiaradia, E.A., Priori, S., L'Abate, G., Canone, D., Mele, M., Comunian, A., Giudici, M., Facchi, A., 2015. Comparing EM38 and Profiler-EMP400 for the Delineation of Homogeneous Management Zones within Agricultural Fields. *Extended Abstracts, First Conference on Proximal Sensing Supporting Precision Agriculture*, 2015, Barcelona. DOI 10.3997/2214-4609.201413828
2. Won, I.J., Keiswetter, D.A., Fields, G.R.A., Sutton, L.C., 1996. GEM-2: A new multifrequency electromagnetic sensor. *Journal of Environmental and Engineering Geophysics*, vol. 2 (1), p. 129–138.
3. Manshtein, A.K., Panin, G.L., Tikunov, S.Iu. 2008. Equipment for frequency electromagnetic sounding "NEMFIS". *Geology and Geophysics*, vol. 49, no. 6, p. 571–579. (in Russ.)
4. Balkov, E.V., 2011. Near surface frequency sounding technology. *Geophysics*, no. 6, p. 42–47. (in Russ.)
5. Balkov, E.V., Fadeev, D.I., Karin, Iu.G., Manshtein, A.K., Manshtein, Yu.A., Panin, G.L., 2017 A new approach to shallow electromagnetic sounding. *Geology and Geophysics*, vol. 58, no. 5, p. 783–791. (in Russ.)

6. Epov M.I. et al. Archeogeophysical Studies in the SB RAS, 2013–2015. *Multidisciplinary Methods in Archeology: Latest Results and Perspectives: Proceedings of the International Symposium*, 2017, p. 405–413. (in Russ.)
7. Balkov, E.V., Stoikin, T.A., Manshtein, A.K., Karin, Iu.G., 2013. Results of using shallow electromagnetic profiling at the electrometric range of the Institute of Petroleum Geology and Geophysics of the SB RAS. *Geophysical research*, vol. 14, no. 3, p. 55–63. (in Russ.)
8. Epov M.I. et al. Determination of the composition of mining waste by geochemical and geophysical methods (on the example of the tailing dump of the Salair mining and processing plant). *Geology and geophysics*, 2017, vol. 58, no. 12, p. 1944–1954. (in Russ.)
9. Manshtein, A.K., Balkov, E.V. Method and device for induction frequency sounding. Patent for invention № 2502092 of 20.12.2013. (in Russ.)
10. Balkov E.V. Multiple images in the signal from local objects with electromagnetic profiling by a compact probe with spaced coils. *All-Russian School-Seminar named after M.N. Berdichevsky and L.L. Vanyana on electromagnetic sounding of the Earth* (St. Petersburg, May 16–21, 2011), ext. theses of doc. St. Petersburg, 2011, 4 p. (in Russ.)
11. Balkov, E.V., Halatov S.Yu., Manshtein, A.K., Panin, G.L. Shallow electromagnetic profiling by compact equipment: theoretical estimates and results of experimental work. *Engineering geophysics 2014: 10th anniversary conference and exhibition* (Gelendzhik, April 21–25, 2014). Gelendzhik, 2014. (in Russ.)

*Received 13.08.2018*

*For citation:*

Karin Yu. G., Balkov E. V., Fadeev D. I., Manshtein A. K., Panin G. L., Alymov A. O., Romanov D. B. Electromagnetic Profiling by Compact Apparatus: New Approach and Results of Application. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 4, p. 68–78. (in Russ.) DOI 10.25205/1818-7900-2018-16-4-68-78

**К. Ф. Лысаков<sup>1,2</sup>, К. К. Облаухов<sup>1</sup>, М. Ю. Шадрин<sup>1,2</sup>**

<sup>1</sup>Новосибирский государственный университет  
ул. Пирогова, 2, Новосибирск, 630090, Россия

<sup>2</sup>Институт автоматики и электрометрии СО РАН  
пр. Академика Коптюга, 1, Новосибирск, 630090, Россия

lysakov@ngs.ru, obko@mail.ru, mikesha@sl.iae.nsk.su

## **ПРИМЕНЕНИЕ FPGA ДЛЯ УСКОРЕНИЯ РАСЧЕТА ВОЛНОВОГО ФРОНТА ЦУНАМИ \***

Рассматривается вопрос ускорения численного расчёта распространения волны цунами от очага до берега в рамках дифференциальной модели мелкой воды с применением аппаратного ускорения на базе FPGA. Разработана программная архитектура для расчета системы мелкой воды без учёта внешних сил, являющейся эквивалентной используемой в пакете MOST для численного моделирования движения волны цунами по водной акватории.

Численная реализация системы мелкой воды осуществлялась по схеме Мак-Кормака. Точность вычислений разработанного решателя была проверена сравнением с точным решением, известным в случае плоского наклонного дна. Достигнутая точность не уступает, а местами и превосходит точность известного программного пакета MOST.

*Ключевые слова:* расчет цунами, ускорение, FPGA, Мак-Кормак, MOST, численные модели.

### **Введение**

Возможность оценки опасности побережий в случае возникновения цунами до прихода туда волны является востребованной и актуальной задачей для прибрежных регионов, находящихся в зоне риска. Основная проблема служб предупреждения цунами – это своевременное оповещение населения и реализация различных эвакуационных сценариев (в зависимости от ожидаемой высоты волны) в тех местах побережья, где это требуется [1; 2]. Работа направлена на ускорение численного расчёта распространения волны цунами от очага до берега в рамках дифференциальной модели мелкой воды с применением аппаратного ускорения на базе FPGA.

Широко известен и применяется подход на основе численного моделирования распространения волны с использованием дифференциальной модели мелкой воды. Наибольшее распространение получили несколько разностных схем, корректно решающих нелинейные уравнения мелкой воды, реализованные в виде алгоритмов пакеты прикладного программного обеспечения MOST, TUNAMI и COMCOT. Численный расчёт распространения волны на персональном компьютере в достаточной области (порядка 10 миллионов расчётных узлов)

---

\* Работа выполнена в рамках бюджетного проекта 0319-2018-0010 «Исследование и развитие методов и технологий построения интегрированных программно-аппаратных комплексов для задач моделирования и управления динамическими системами обработки и отображения данных» (регистрационный номер: АААА-А17-117062110016-4).

длиться 2–3 часа [3]<sup>1</sup>. При этом пространственные шаги сетки, покрывающей область такого размера, в которую входят и очаговая область, и береговая линия, не может быть достаточно детальной для корректного моделирования вблизи берега и в бухтах. При использовании в расчётах шага, равного нескольким метрам (вместо нескольких сот метров), потребуется на 2–3 порядка больше вычислительных ресурсов (оперативной памяти и времени расчётов), что практически исключает использование персональных компьютеров для получения результата в течение имеющихся в запасе десятков минут.

### Математическая модели и расчетная схема

В пакете MOST для численного моделирования движения волны цунами по водной акватории используется следующий эквивалентный вид так называемой системы мелкой воды без учёта внешних сил (донное трение, сила Кориолиса и т. д.) [4]:

$$\begin{aligned}\frac{\partial H}{\partial t} + \frac{\partial(Hu)}{\partial x} + \frac{\partial(Hv)}{\partial y} &= 0 \\ \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + g \frac{\partial \eta}{\partial x} &= 0 \\ \frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + g \frac{\partial \eta}{\partial y} &= 0\end{aligned}$$

где  $H(x, y, t) = \eta(x, y, t) + D(x, y)$  – полная глубина слоя жидкости,  $\eta(x, y, t)$  – глубина возмущенного слоя,  $D(x, y)$  – профиль глубин, считающийся известным (цифровая батиметрия),  $t$  – время,  $u(x, y, t)$ ,  $v(x, y, t)$  – компоненты вектора скорости вдоль осей  $x$  и  $y$  соответственно,  $g$  – ускорение свободного падения.

Моделирование происходит в прямоугольной и не меняющейся со временем области, с применением разностных схем прямоугольной сетке, где  $\Delta x = x_{i+1} - x_i$ ,  $\Delta y = y_{j+1} - y_j$  – шаги сетки по пространственным переменным  $x$  и  $y$ ,  $\tau^n = t^{n+1} - t^n$  – шаг по времени. В данной работе для простоты используется равномерный шаг как по пространственным переменным, так и по времени.

Для аппроксимации уравнений мелкой воды используется явная двух шаговая конечно-разностная схема второго порядка аппроксимации типа Мак-Кормака [5; 6].

1 шаг:

$$\begin{aligned}\frac{\hat{H}_{ij}^{n+1} - H_{ij}^n}{\tau} + \frac{H_{ij}^n u_{ij}^n - H_{i-1j}^n u_{i-1j}^n}{\Delta x} + \frac{H_{ij}^n v_{ij}^n - H_{i-1j}^n v_{i-1j}^n}{\Delta y} &= 0 \\ \frac{\hat{u}_{ij}^{n+1} - u_{ij}^n}{\tau} + u_{ij}^n \frac{u_{ij}^n - u_{i-1j}^n}{\Delta x} + v_{ij}^n \frac{u_{ij}^n - u_{i-1j}^n}{\Delta y} + g \frac{\eta_{ij}^n - \eta_{i-1j}^n}{\Delta x} &= 0 \\ \frac{\hat{v}_{ij}^{n+1} - v_{ij}^n}{\tau} + u_{ij}^n \frac{v_{ij}^n - v_{i-1j}^n}{\Delta x} + v_{ij}^n \frac{v_{ij}^n - v_{i-1j}^n}{\Delta y} + g \frac{\eta_{ij}^n - \eta_{i-1j}^n}{\Delta y} &= 0\end{aligned}$$

2 шаг:

$$\begin{aligned}\frac{H_{ij}^{n+1} - (\hat{H}_{ij}^{n+1} + H_{ij}^n)/2}{\tau/2} + \frac{\hat{H}_{i+1j}^{n+1} \hat{u}_{i+1j}^{n+1} - \hat{H}_{ij}^{n+1} \hat{u}_{ij}^{n+1}}{\Delta x} + \frac{\hat{H}_{ij+1}^{n+1} \hat{v}_{ij+1}^{n+1} - \hat{H}_{ij}^{n+1} \hat{v}_{ij}^{n+1}}{\Delta y} &= 0 \\ \frac{u_{ij}^{n+1} - (\hat{u}_{ij}^{n+1} + u_{ij}^n)/2}{\tau/2} + u_{ij}^n \frac{\hat{u}_{i+1j}^{n+1} - \hat{u}_{ij}^{n+1}}{\Delta x} + v_{ij}^n \frac{\hat{u}_{ij+1}^{n+1} - \hat{u}_{ij}^{n+1}}{\Delta y} + g \frac{\hat{\eta}_{i+1j}^{n+1} - \hat{\eta}_{ij}^{n+1}}{\Delta x} &= 0 \\ \frac{v_{ij}^{n+1} - (\hat{v}_{ij}^{n+1} + v_{ij}^n)/2}{\tau/2} + u_{ij}^n \frac{\hat{v}_{i+1j}^{n+1} - \hat{v}_{ij}^{n+1}}{\Delta x} + v_{ij}^n \frac{\hat{v}_{ij+1}^{n+1} - \hat{v}_{ij}^{n+1}}{\Delta y} + g \frac{\hat{\eta}_{ij+1}^{n+1} - \hat{\eta}_{ij}^{n+1}}{\Delta y} &= 0\end{aligned}$$

<sup>1</sup> См. также: TSUNAMI Samoa, 29 September, 2009. URL: [https://www.preventionweb.net/files/27077\\_tsunami\\_publication2wfbblanks.pdf](https://www.preventionweb.net/files/27077_tsunami_publication2wfbblanks.pdf)

## Реализация алгоритмов на FPGA

Вентильные матрицы, программируемые пользователем (FPGA) для различных задач, используются с 1990-х гг. При этом вначале наибольшее применение нашли вариации программируемых логических матриц (PLD/CPLD), имеющие энергонезависимую память для хранения конфигурации. PLD успешно применялись для решения небольших задач по разработке автономных устройств и контроллеров, имеющих малое энергопотребление и массгабаритные размеры. Увеличение объема логических элементов и возможность неограниченного перепрограммирования для изменения алгоритмов обработки данных стало возможным при появлении FPGA. Такие FPGA позволяют реализовывать трудоемкие математические алгоритмы обработки данных, а также выполнять прототипирование вычислительных устройств для дальнейшего производства серийных заказных микросхем ASIC. Но до недавнего времени существовало 2 основных препятствия для применения FPGA, связанных с необходимостью ручной реализации алгоритмов с точностью до регистров и триггеров – уровень регистровых передач RTL (register transfer level):

- 1) трудоемкость реализации программных модулей, требующая значительно больше времени (до 2 порядков больше, чем реализация для PC на высокоуровневых языках);
- 2) высокие требования к квалификации разработчика.

Создаваемые средства для автоматизации процесса разработки (Simulink MatLab, SystemVerilog, SystemC) не позволяли производить полноценную разработку алгоритмов на высокоуровневом языке: происходило значительное снижение производительности и требования к ресурсам по сравнению с RTL моделью.

Современная технология High-Level Synthesis (HLS) позволяет использовать для описания цифровых схем C-подобный язык [7]. Это является новым витком развития способов создания различных вычислительных устройств, обеспечивающих полный цикл: от описания архитектуры до верификации проекта с использованием различных средств моделирования. HLS представляет собой автоматизированный процесс проектирования, который интерпретирует алгоритмическое описание поведенческой модели и позволяет создавать цифровые устройства, четко реализующие заданные условия. С помощью HLS можно легко изменять параметры конвейера, подстраиваясь под временные (timing) или пространственные (utilization) требования. Также технология HLS позволяет производить верификацию программного кода до его преобразования в схему для конкретной микросхемы FPGA. Именно технология HLS и большие ресурсы современных микросхем FPGA позволяют сегодня применять их для решения новых типов задач в новых областях.

## Архитектура вычислителя

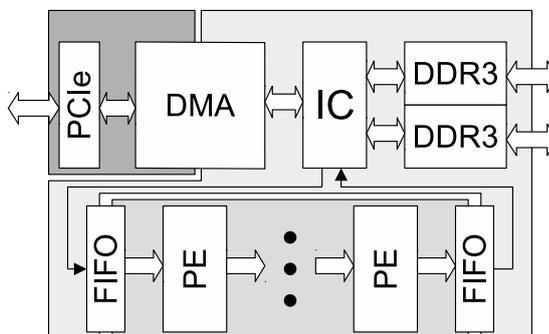
Изложенная вычислительная схема была реализована на платформе вентильных матриц программируемых пользователем FPGA.

Для реализации алгоритма на FPGA была предложена архитектура поточного вычислителя, состоящая из процессорных элементов (ПЭ). ПЭ осуществляет вариант двумерной прогонки – конвейера с последовательным потоком данных. Вспомогательными данными для каждого узла, которые необходимо хранить, являются значения всех используемых функций в 4-х соседних узлах. Особенность архитектуры FPGA позволяет использовать внутреннюю память (BRAM) для организации соответствующего буфера данных и линии задержки, что позволяет реализовать конвейер при количестве памяти до  $2N_x$ .

Математические операции, необходимые для алгоритма Мак-Кормака, реализованы вычислительным конвейером, обеспечивающим производительность обработки одной точки сетки за один такт. Полученное решение позволяет гарантировать производительность с точностью до такта.

Блок-схема предложенной архитектура всего спецпроцессора представлена на рисунке (см. далее). Помимо самого вычислителя, спецпроцессор имеет контроллеры памяти DDR3, контроллер PCIe, модуль DMA, обеспечивающий взаимодействие вычислителя с памятью компьютера-хоста в режиме прямого доступа. Данные поступают в вычислитель из памяти через FIFO, что позволяет легко варьировать частоту отдельных вычислительных блоков,

подстраиваясь под конкретный кристалл FPGA и характеристики внешней памяти и интерфейсов. Сам вычислитель в зависимости от доступных ресурсов FPGA состоит из одного или нескольких процессорных элементов.



Архитектура спецпроцессора

Спроектированный вычислитель на уровне алгоритмов был протестирован с применением технологий HLS. Модули, обеспечивающие работу всего процессора, протестированы с помощью RTL-симуляторов.

Для практической реализации и тестирования были использованы следующие платформы:

- SLEDv7 на базе кристалла семейства Virtex-6 [8];
- VC709 на базе кристалла Virtex-7<sup>2</sup>.

Итоговая производительность и затраты ресурсов зависят не только от количества ПЭ, но также от частоты их работы. В результате тестирования было выяснено, что реализованные конвейеры способны работать на частоте 200 МГц на плате SLEDv7 и до 300 МГц на платах VC709 и KU115.

Далее в таблице представлена производительность вычислителя для двух платформ. Замеры представлены для тестовых данных, используемых NOAA (2580 × 2879 точек, акватория Тихого океана) с временным шагом 10 с. В случае платы SLEDv7 использовалась цепочка из 2 ПЭ одинарной точности на частоте 200 МГц. В результате была получена производительность в 52 итерации в секунду, или час модельного времени за 6,84 с.

Сравнение производительности вычислителей при использовании разных кристаллов

	xc6vsx315-1 (SLEDv7)	xc7vx690-2 (VC709)
Кол-во ПЭ	2	8
Частота (МГц)	200	250
Время одного прохода (мс)	38	31
Время одной итерации (мс)	19	3,8

Платформа VC709 позволила разместить в себе 8 процессорных элементов на частоте 250 МГц. Итоговая производительность составила 263 итерации в секунду, или час модельного времени за 1,368 с.

<sup>2</sup> Xilinx Inc (2016), "Xilinx Virtex-7 FPGA VC709 Connectivity Kit". URL: <https://www.xilinx.com/products/boards-and-kits/dk-v7-vc709-g.html>

## Тестирование предлагаемого решения

Для тестирования описанного численного метода и его реализации было проведено сравнение результатов численного решения модельной задачи распространения цунами от круглого источника над наклонным дном с известным точным решением, а также с решением, полученным пакетом MOST [5]. Сравнение распределения высот и времен прихода волны показало хорошее их соответствие вне зоны шельфа, т. е. там, где глубина превышает 150–200 м.

## Заключение

В целях ускорения расчета распространения волны цунами по глубоководной части океана был спроектирован спецвычислитель на базе архитектуры вентиляльных матриц программируемых пользователем (FPGA). Численная реализация системы мелкой воды осуществлялась по схеме Мак-Кормака. Точность вычислений разработанного решателя была проверена сравнением с точным решением, известным в случае плоского наклонного дна. Достигнутая точность не уступает, а местами и превосходит точность известного программного пакета MOST.

Предложенная архитектура поточного вычислителя на базе FPGA основана на процессорных элементах (ПЭ) для двумерной прогонки (конвейера с последовательным потоком данных), что позволяет масштабировать решение для различных кристаллов, получая максимально возможную производительность. Архитектура предложенного ПЭ позволяет одновременно обрабатывать  $N$  точек входного потока данных, что дает возможность параметризовать ПЭ в зависимости от особенностей используемой вычислительной платформы для получения максимальной производительности.

## Список литературы / References

1. Titov V. V., Gonzalez F. I. Implementation and testing of the method of splitting tsunami (MOST) model // NOAA Technical Memorandum ERL PMEL-112. 1997.
2. Titov V., Gonzalez F., Bezhaev A. Yu., Lavrentiev M. M., Marchuk A. G., Avdeev A. V. Real-Time Tsunami Forecasting: Challenges and Solutions // Proc. of the International Conference "Mathematical Methods in Geophysics". Novosibirsk, 2003. P. 225–228.
3. Gica E., Spillane M., Titov V., Chamberlin C., Newman J. Development of the forecast propagation database for NOAA's short-term inundation forecast for tsunamis (SIFT) // NOAA Technical Memorandum. 2008. URL: <https://www.ndbc.noaa.gov/dart.shtml>
4. Marchuk A. G. Estimating Tsunami Wave Height over a Sloping Bottom in the Ray Approximation // Numerical Analysis and Applications. 2015. Vol. 8 (4). P. 304–313.
5. Lavrentiev M. M., Romanenko A. A., Lysakov K. F. Modern Computer Architecture to Speed-Up Calculation of Tsunami Wave Propagation // Proc. of the Eleventh (2014) Pacific/Asia Offshore Mechanics Symposium. 2014. P. 186–191.
6. Lavrentiev M. M., Romanenko A. A. Modern software architectures facilities to accelerate calculations // Proc. of the XII International workshop «Supercomputing and Computer modelling», RFNC-VNIIEF. Sarov, 2011. P. 251–256.
7. Zwagerman M. D. High Level Synthesis, a Use Case Comparison with Hardware Description Language // Master Theses, 2015. URL: <https://scholarworks.gvsu.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1754&context=theses>
8. Lysakov K. F., Shadrin M. Yu. FPGA Based Hardware Accelerator for High Performance Data-Stream Processing // Pattern Recognition and Image Analysis. 2013. Vol. 23 (1). P. 26–34.

**K. F. Lysakov<sup>1,2</sup>, K. K. Oblaukhov<sup>1</sup>, M. Yu. Shadrin<sup>1,2</sup>**

<sup>1</sup> *Novosibirsk State University*

*2 Pirogov Str., Novosibirsk, 630090, Russian Federation*

<sup>2</sup> *Institute of Automation and Electrometry SB RAS*

*1 Academician Koptuyug Ave., Novosibirsk, 630090, Russian Federation*

*lysakov@ngs.ru, obko@mail.ru, mikesha@sl.iae.nsk.su*

### **USING FPGA TO ACCELERATE COMPUTATION OF TSUNAMI WAVEFRONT**

We consider the problem of speeding up numerical calculation of tsunami wave propagation from source to the coast according to the shallow water system by using FPGA (Field Programmable Gate Array) based hardware acceleration. To compute the solution to shallow water system (neglecting external forces), special software architecture has been developed. This is equivalent to the one used in MOST (Method of Splitting Tsunami) software package to simulate wave movement through the given water area. Numerical solution to the shallow water system was done according to Mac-Cormack scheme. Precision of the developed solver was tested against the known exact solution for the case of plane slope bottom. The achieved accuracy is comparable, or even exceeds, the word recognizable software package MOST.

*Keywords:* tsunami simulation, code acceleration, FPGA, Mac-Cormack scheme, MOST, numerical models.

*Received 12.10.2018*

*For citation:*

Lysakov K. F., Oblaukhov K. K., Shadrin M. Yu. Using FPGA to Accelerate Computation of Tsunami Wavefront. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 4, p. 79–84. (in Russ.) DOI 10.25205/1818-7900-2018-16-4-79-84

**Т. В. Лях<sup>1</sup>, В. Е. Зюбин<sup>1</sup>, Н. О. Гаранина<sup>2</sup>**

<sup>1</sup> *Институт автоматики и электрометрии СО РАН  
пр. Академика Коптюга, 1, Новосибирск, 630090, Россия*

<sup>2</sup> *Институт систем информатики им. А. П. Ершова СО РАН  
пр. Академика Коптюга, 6, Новосибирск, 630090, Россия*

*antsys\_nsu@mail.ru, zyubin@iae.nsk.su, garanina@iis.nsk.su*

## **АВТОМАТИЗИРОВАННАЯ ВЕРИФИКАЦИЯ АЛГОРИТМОВ УПРАВЛЕНИЯ СЛОЖНЫМИ ТЕХНОЛОГИЧЕСКИМИ ОБЪЕКТАМИ НА ПРОГРАММНЫХ ИМИТАТОРАХ \***

Статья посвящена проблеме проверки алгоритмов управления, созданных в рамках процесс-ориентированного подхода, на соответствие входным спецификациям. Представлена общая схема верификации алгоритма управления, приведены ее реализация в автоматизированном варианте и результаты практической апробации в проекте по автоматизации Большого солнечного вакуумного телескопа.

*Ключевые слова:* автоматизация, промышленные алгоритмы управления, верификация, процесс-ориентированные языки программирования.

### **Введение**

Алгоритмы управления сложными технологическими объектами обладают рядом свойств, специфичных для области промышленной автоматизации [1–4]:

- открытость – наличие «окружающей среды», внешнего мира, с которым взаимодействует алгоритм управления;
- событийность – алгоритм управления формирует управляющие воздействия как реакцию на события (значимые изменения во входных данных), в том числе на управляющие команды от оператора;
- неопределенная продолжительность функционирования алгоритма управления;
- синхронизм – необходимость синхронизации реакции алгоритма управления с событиями на объекте управления;
- логический параллелизм – алгоритм управления структурно отражает параллелизм физических процессов на объекте управления, их независимость.

Реализация алгоритмов управления средствами объектно-ориентированных языков общего назначения чревата чрезмерным усложнением программной архитектуры при росте сложности алгоритма [5]. Поэтому в области промышленной автоматизации используются специализированные языковые средства для разработки алгоритмов управления: языки МЭК 61131-3, G (NI LabVIEW), Reflex [6–8].

---

\* Работа выполнена при финансовой поддержке Федерального агентства научных организаций (государственная регистрация № АААА-А17-117060610006-6) и при финансовой поддержке РФФИ (проект № 17-07-01600).

Использование языков МЭК 61131-3 трудоемко из-за низкой выразительности этих языков, а в некоторых случаях и неприемлемо, например при необходимости интеграции кода в сторонние системы [6].

Исследователи альтернативных лингвистических средств для описания алгоритмов управления [2; 4; 7–10] предлагают и практически обосновывают эффективность предметно-ориентированных языков на основе модели конечного автомата, в частности процесс-ориентированного языка Reflex [10].

При использовании конечно-автоматных языков в промышленной автоматизации [7; 8; 11; 12] основную проблему представляет решение задачи тестирования и верификации созданных алгоритмов, поскольку методы, разработанные для тестирования и верификации программного обеспечения в области объектно-ориентированного программирования, слабо применимы [5]. Управляющий алгоритм невозможно тестировать автономно. Тестирование алгоритма на реальном объекте управления может привести к поломке оборудования или аварийной ситуации. Поэтому наиболее распространенный подход – ручная проверка на этапе пуско-наладки: проверяющий контролирует реакцию алгоритма на различные ситуации, постепенно усложняя тесты. Подход очень трудоемкий. Он приводит к серьезным психологическим нагрузкам на разработчиков, не гарантирует полноту верификации, затрудняет контроль качества верификации и в итоге усложняет разработку управляющих алгоритмов [11].

Поэтому разработка методов верификации алгоритмов управления интересна не только с теоретической, но и с практической стороны [5; 13–18]. Современная тенденция – использовать для тестирования и верификации алгоритмов управления программные имитаторы объекта управления [11; 19–21].

В статье предлагается подход к верификации алгоритмов управления сложными технологическими объектами на основе концепции виртуальных объектов управления (ВОУ), включающий в себя создание кода алгоритма управления и программную реализацию объекта управления в виде ВОУ.

Основные положения, изложенные в статье, были представлены на V Международной научной конференции «Математическое и компьютерное моделирование» (Омск, 2017) [22].

### **Общая схема верификации алгоритмов управления технологическими объектами**

Общая схема итерационной разработки алгоритмов управления (рис. 1), предложенная ранее в [11], включает следующие шаги:

- верифицируемый алгоритм управления (его часть) реализуется программно и оформляется в обособленный алгоблок;
- модель технологического объекта (его часть) также реализуется программно и оформляется в обособленный алгоблок, называемый виртуальным объектом управления;
- проводится верификация через создание тестовых ситуаций (сценариев) и контроль реакции алгоритма управления;
- коррекция кода алгоритма управления и (или) виртуального объекта управления по результатам верификации.

Схема позволяет использовать итерационный подход к разработке промышленных алгоритмов управления сложными технологическими объектами. В упрощенном виде схема была опробована при создании набора виртуальных лабораторных стендов для обучения студентов ИТ-специальностей, специализирующихся в области промышленной автоматизации [23]. Упрощение заключалось в неизменности ВОУ. При этом ВОУ выполнены с использованием графики для повышения наглядности и визуального контроля корректности алгоритма управления.

Несмотря на высокую эффективность при использовании в учебном процессе, подход не нашел практического применения в реальных проектах в силу высокой трудоемкости создания графических моделей ВОУ.

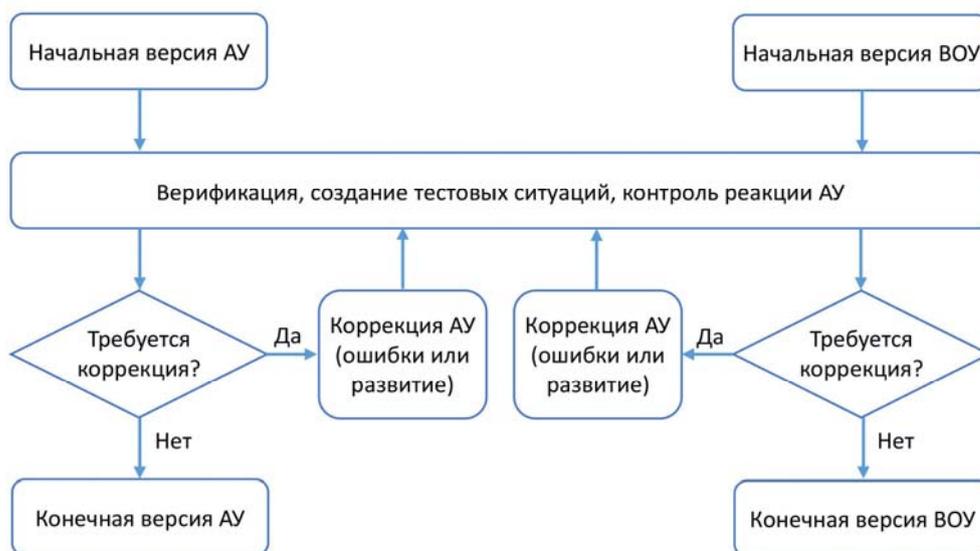


Рис. 1. Общая схема итерационной разработки алгоритмов управления:  
 АУ – алгоритм управления, ВОУ – виртуальный объект управления

### Метод разработки алгоритмов управления на основе концепции виртуальных объектов управления

Для использования в реальных проектах автоматизации предложенная схема была программно реализована в автоматизированном комплексе верификации алгоритмов управления.

В автоматизированном варианте (рис. 2, таблица) управление сценариями работы и контроль реакции алгоритма производится оператором через графический интерфейс (1). Графический интерфейс предоставляет оператору возможность отправлять штатные команды алгоритму управления (6) через очередь сообщений (2) и контролировать сообщения от алгоритма управления через очередь сообщений (4), управлять поведением ВОУ (9) через очередь сообщений (3) и контролировать сообщения от ВОУ через очередь сообщений (5). Также на графическом интерфейсе оператора (ГИО) отображаются состояния входных (7) и выходных (8) дискретных сигналов алгоритма управления. Для имитации входных аналоговых сигналов от датчиков / АЦП организован дополнительный канал связи между ВОУ и АУ (10), а для имитации выходных аналоговых сигналов (ЦАП) – канал связи (11).

Программный комплекс автоматизированной верификации алгоритмов управления был реализован на базе пакета LabVIEW (рис. 3). Комплекс включает ГИО (2), который конфигурируется через модуль загрузки конфигурационных файлов (МЗКФ) (1) и взаимодействует с целевыми модулями алгоблоков АУ (3) и ВОУ (4). Модули загрузки конфигурационных файлов и ГИО реализованы на языке G LabVIEW. Исполняемые модули АУ и ВОУ генерируются из описания на языке Reflex и в виде DLL интегрируются в LabVIEW, что в отличие от подхода, использованного при разработке виртуальных лабораторных стендов, обеспечивает возможность итерационно развивать не только АУ, но и ВОУ.

Управление верификацией ведется оператором через ГИО (рис. 4). ГИО представлен расположенной внизу окна неизменяемой панелью управления верификацией (3) и шестью вкладками: генерации входных сообщений для АУ и ВОУ, переменных АУ (рис. 5), переменных ВОУ, отладочной информации АУ (рис. 6), отладочной информации ВОУ и вкладки «Помощь» с руководством по работе с комплексом. Панель управления верификацией содержит элементы управления режимом отладки (пошаговый / непрерывный), запуска и остановки алгоблоков.

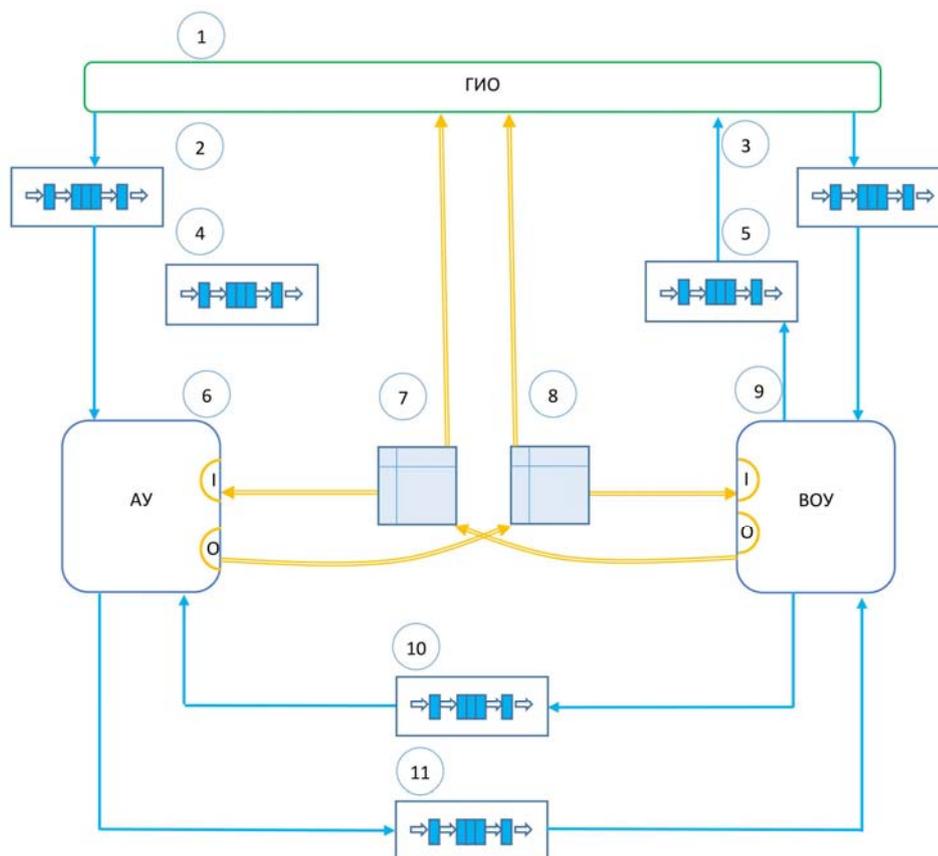


Рис. 2. Архитектура автоматизированного комплекса верификации: 1 – ГИО; 2 – очередь сообщений от ГИО; 3 – очередь сообщений для ВОУ; 4 – очередь сообщений от АУ; 5 – очередь сообщений от ВОУ; 6 – АУ; 7 – буфер входных дискретных сигналов для АУ; 8 – буфер выходных дискретных сигналов для АУ; 9 – ВОУ; 10 – очередь сообщений для имитации аналоговых данных от датчиков / АЦП для АУ; 11 – очередь сообщений для имитации выходных аналоговых сигналов (ЦАП) от АУ

### Используемые условные обозначения

	Алгоблок
	Цифровые порты алгоритмического блока (I – входные, O – выходные)
	Буфер значений дискретных цифровых сигналов
	Очередь сообщений
	Цифровые данные
	Сообщения

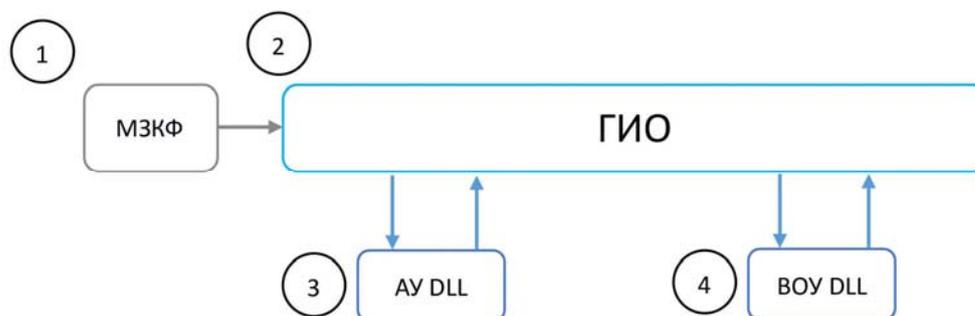


Рис. 3. Архитектура программного комплекса автоматизированной верификации алгоритмов управления

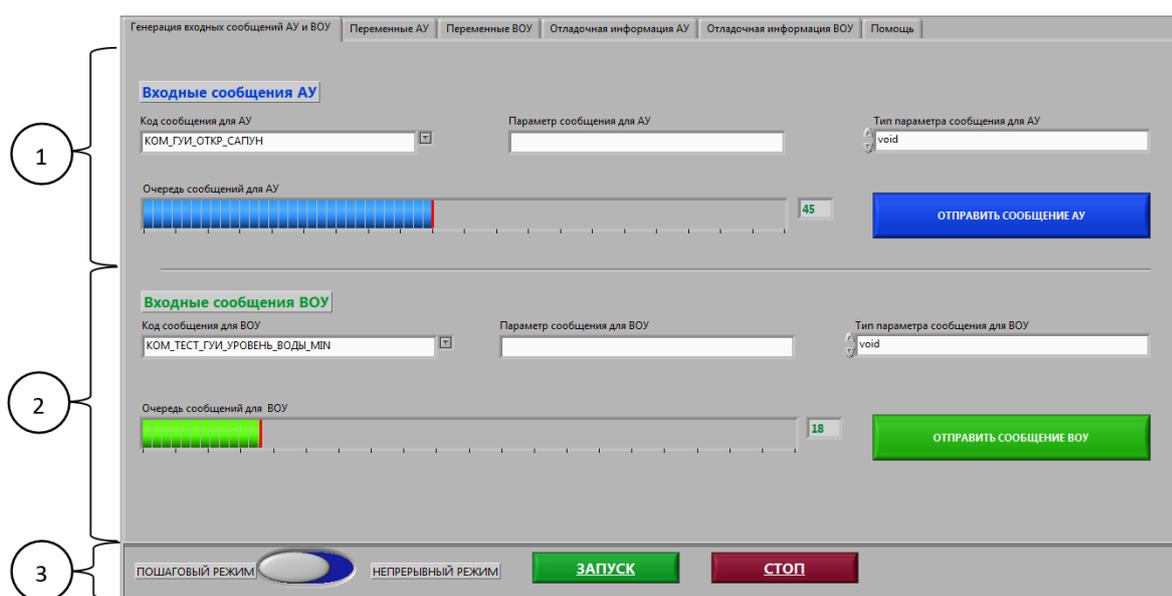


Рис. 4. Графический интерфейс оператора, вкладка «Генерация входных сообщений для АУ и БОУ»: 1 – панель генерации входных сообщений для АУ; 2 – панель генерации входных сообщений для БОУ; 3 – область управления верификацией

Вкладка генерации входных сообщений АУ и БОУ (см. рис. 4) разделена на две области: панель управления АУ (1) и панель управления БОУ (2) для ввода входных сообщений для АУ и БОУ соответственно.

Вкладка «Переменные АУ» (рис. 5) разделена на три области:

- панель значений входных портов АУ (1) с указанием имени входного порта, источника, который генерирует значения порта (БОУ / РУЧН), и значения битов порта, отображаемых цветом;
- панель значений входных и выходных глобальных переменных АУ (2) с указанием типа, имени переменной и ее текущего значения;
- панель режима управления входным портом АУ (3), позволяющая оператору устанавливать значения битов выделенного порта вручную.

Вкладка «Переменные БОУ» выглядит аналогично.

На вкладке «Отладочная информация АУ» (см. рис. 6) отображаются выходные сообщения АУ (1) и текущие состояния процессов АУ (2).

Вкладка «Отладочная информация БОУ» выглядит аналогично.

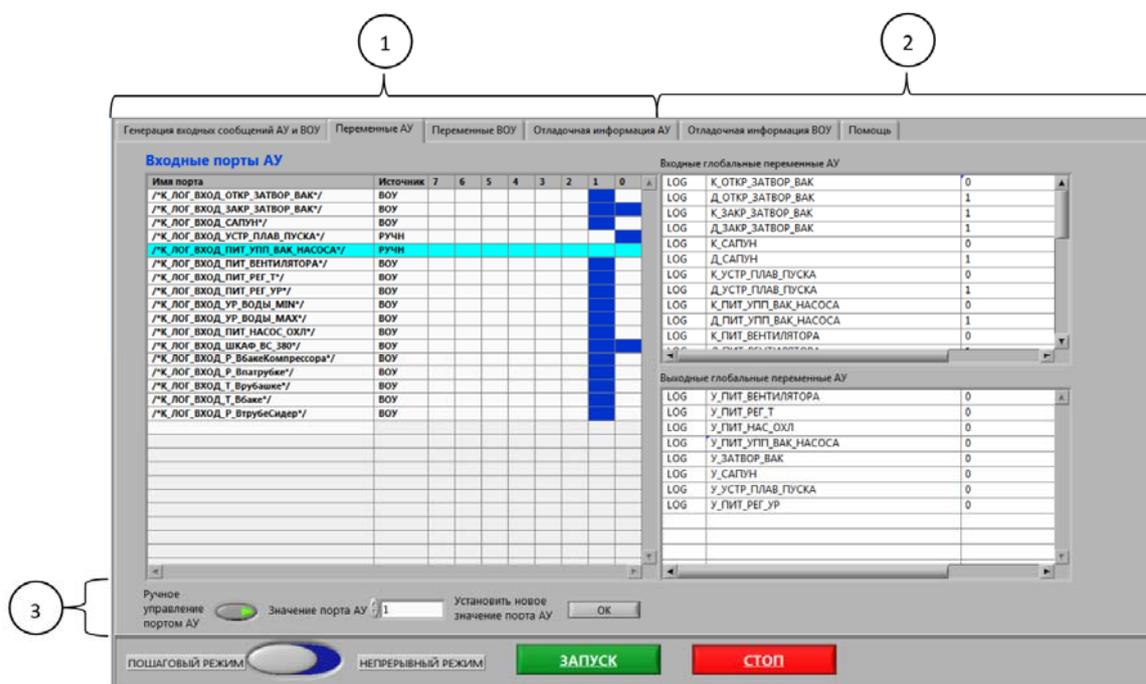


Рис. 5. Графический интерфейс оператора, вкладка «Переменные АУ»:

1 – панель значений входных портов АУ; 2 – панель значений входных и выходных глобальных переменных АУ; 3 – панель режима управления входным портом АУ

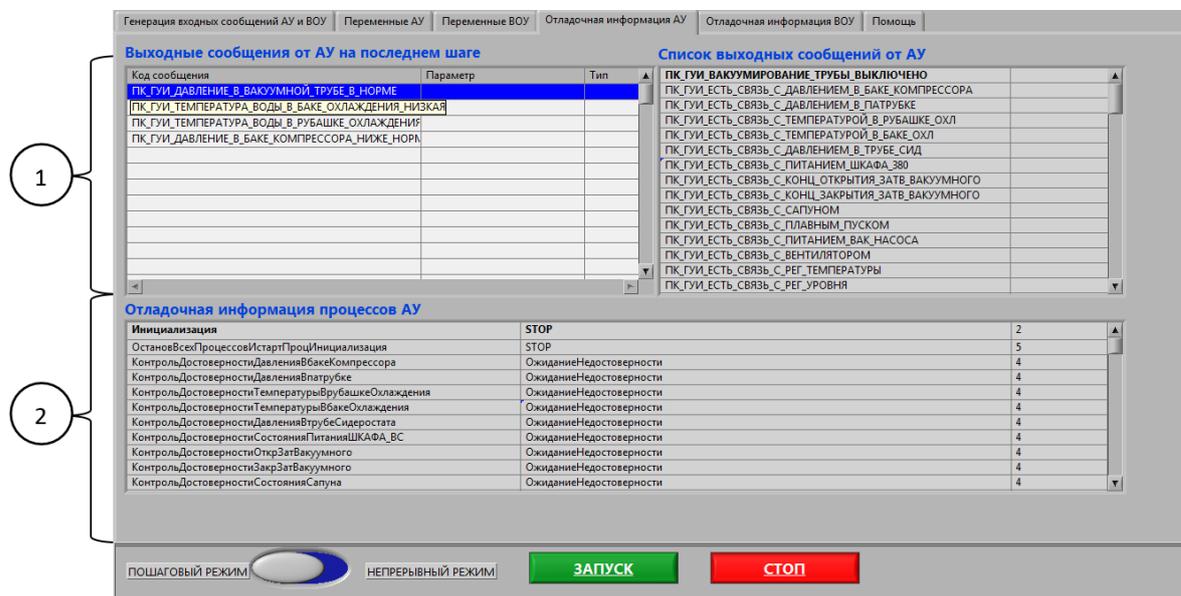


Рис. 6. Графический интерфейс оператора, вкладка «Отладочная информация АУ»:

1 – панель выходных сообщений АУ; 2 – панель текущих состояний процессов АУ

При запуске комплекса МЗКФ на основании конфигурационных файлов, создаваемых транслятором языка Reflex, настраивает вкладки ГИО (имена переменных, портов, процессов, входных / выходных сообщений) и передает управление ГИО. ГИО выделяет память под очереди сообщений, буферы входных / выходных переменных, буфер состояния процессов,

и затем передает указатели на выделенные области памяти алгоблокам. По началу верификации (кнопка «Запуск» на панели управления верификацией) ГИО циклически активизирует алгоблоки в определенной последовательности: сначала ВОУ, затем АУ. Создание тестовых ситуаций производится оператором через вкладки «Генерация входных сообщений АУ и ВОУ», «Переменные АУ», «Переменные ВОУ». Контроль корректности алгоритма управления ведется визуально. Разработка алгоритма управления соответствует общей схеме итерационной разработки (см. рис. 1).

### Практическая апробация подхода

Решение было практически апробировано на задаче автоматизации Большого солнечного вакуумного телескопа (пос. Листвянка, Иркутская обл.) [5; 24]. В проекте был создан и верифицирован алгоритм управления системой вакуумирования. Верифицировалась работа алгоритма при создании вакуума в трубе телескопа, развакуумировании перед техническими работами, реакция алгоритма на изменение температуры окружающей среды, уровня воды в системе охлаждения, на разгерметизацию трубы телескопа, отказ отсечных клапанов, вакуумных заслонок, насосов, вытяжных вентиляторов, датчиков и исполнительных органов системы климат-контроля и т. д.

Верификация была проведена на территории разработчика и обеспечила значительное сокращение общей трудоемкости работ, в частности, продолжительности пусконаладки на целевом объекте была сокращена более чем в два раза.

Выявленные недостатки подхода:

- визуальный контроль за реакцией алгоритма и сложность анализа отображаемой информации не исключает вероятности пропуска ошибки при верификации;
- в силу того что на каждой итерации должна быть заново проверена реакция алгоритма в соответствии со списком тестовых ситуаций, возникает большое количество рутинных действий.

### Заключение

В работе была предложена и рассмотрена автоматизированная схема итерационной разработки алгоритмов управления технологическим объектом. Итерационная схема разработки предполагает верификацию алгоритма управления, специфицированного на процесс-ориентированном языке Reflex, на виртуальном объекте управления до начала приемо-сдаточных испытаний на целевом объекте. Эффективность автоматизированной схемы верификации была подтверждена в проекте автоматизации Большого солнечного вакуумного телескопа при создании ПО подсистемы вакуумирования.

### Список литературы

1. *Закревский А. Д.* Параллельные алгоритмы логического управления. М.: Эдиториал УРСС, 2003. 200 с.
2. *Wagner F., Schmuki R., Wagner T., Wolstenholme P.* Modeling software with finite state machines: a practical approach // Auerbach Publications. 2006. 390 p.
3. *Harel D.* Statecharts: A visual formalism for complex systems // Science of computer programming. 1987. Vol. 8.3. P. 231–274.
4. *Зюбин В. Е.* Программирование ПЛК: языки МЭК 61131-3 и возможные альтернативы // Промышленные АСУ и контроллеры. 2005. № 11. С. 31–35.
5. *Лях Т. В., Зюбин В. Е., Сизов М. М.* Опыт применения языка Reflex при автоматизации Большого солнечного вакуумного телескопа // Промышленные АСУ и контроллеры. 2016. № 7. С. 37–43.
6. *Зюбин В. Е.* К пятилетию стандарта IEC 1131-3. Итоги и прогнозы // Приборы и системы. Управление, контроль, диагностика. 1999. № 1. С. 64–71.
7. *Samek M., Montgomery P.* State Oriented Programming // Embedded Systems Programming. 2000. August. P. 22–43.

8. *Шальто А. А., Туккель Н. И.* SWITCH технология – автоматный подход к созданию программного обеспечения «реактивных» систем // Программирование. 2001. № 5. С. 45–62.
9. *Розов А. С., Зюбин В. Е., Нефедов Д. В.* Программирование встраиваемых микроконтроллерных систем на основе гиперпроцессов // Вестн. НГУ. Серия: Информационные технологии. 2017. Т. 15, № 4. С. 64–73.
10. *Зюбин В. Е.* «Си с процессами» – язык программирования логических контроллеров // Мехатроника, автоматизация, управление. 2006. № 12. С. 31–35.
11. *Зюбин В. Е.* Итерационная разработка управляющих алгоритмов на основе имитационного моделирования объекта управления // Автоматизация в промышленности. 2010. № 11. С. 43–48.
12. *Розов А. С., Лях Т. В., Краснов Д. В., Санжиев Е. С.* Практическая апробация языка IndustrialC на примере автоматизации установки термовакуумного напыления // Вестн. НГУ. Серия: Информационные технологии. 2017. Т. 15, № 3. С. 90–99.
13. *Гаранина Н. О., Зюбин В. Е., Лях Т. В.* Онтологический подход к организации шаблонов требований в рамках системы поддержки формальной верификации программных систем. // Системная информатика. 2017. № 9. С. 111–132.
14. *Shilov N. V., Garanina N. O.* Combined Logics of Knowledge, Time, and Actions for Reasoning about Multi-agent Systems. Knowledge Processing and Data Analysis // Lecture Notes in Computer Science. 2011. Iss. 6581. P. 48–58.
15. *Шелехов В. И.* Верификация и синтез программ сложения на базе правил корректности операторов // Моделирование и анализ информационных систем. 2010. Т. 17, № 4. С. 101–110.
16. *Clarke E. M., Gao S.* Model checking hybrid systems // International Symposium on Leveraging Applications of Formal Methods, Verification and Validation. Lecture Notes in Computer Science. 2014. Iss. 8803. P. 385–386.
17. *Clarke E. M., Grumberg O., Peled D.* Model checking. MIT Press, 1999. 314 p.
18. *Clarke E. M., Emerson A., Sistla P.* Automatic verification of finite-state concurrent systems using temporal logic specifications // ACM Transactions on Programming Languages and Systems. Vol. 8. Iss. 2. 1986. P. 244–263.
19. *Isermann R., Schaffnit J., Sinsel S.* Hardware-in-the-loop simulation for the design and testing of engine-control systems // Control Engineering Practice. 1999. Vol. 7. P. 643–653.
20. *Xiao B., Starke M., King D., Irmingier Ph., Herron A., Ollis B., Xue Y.* Implementation of system level control and communications in a Hardware-in-the-Loop microgrid testbed // IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT). 2016. P. 1–5.
21. *Зюбин В. Е.* Использование виртуальных объектов для обучения программированию информационно-управляющих систем // Информационные технологии. 2009. № 6. С. 79–82.
22. *Лях Т. В., Зюбин В. Е.* Автоматическая верификация алгоритмов управления сложными технологическими объектами на программных имитаторах // Математическое и компьютерное моделирование: Сб. материалов V Междунар. науч. конф., посвящ. памяти Р. Л. Долганова. Омск, 2017. С. 128–130.
23. *Зюбин В. Е., Калугин А. А.* Виртуальные лабораторные стенды: обучение программированию задач промышленной автоматизации // Промышленные АСУ и контроллеры. 2009. № 2. С. 39–44.
24. *Ковадло П. Г., Лубков А. А., Бевзов А. Н., Будников К. И., Власов С. В., Зотов А. А., Колобов Д. Ю., Курочкин А. В., Котов В. Н., Лылов С. А., Лях Т. В., Максимов А. С., Перебейнос С. В., Петухов А. Д., Пещеров В. С., Попов Ю. А., Русских И. В., Томин В. Е.* Система автоматизации Большого солнечного вакуумного телескопа // Автометрия. 2016. Т. 52, вып. 2. С. 97–106.

T. V. Liakh<sup>1</sup>, V. E. Zyubin<sup>1</sup>, N. O. Garanina<sup>2</sup>

<sup>1</sup> Institute of Automation and Electrometry SB RAS  
1 Academician Koptyug Ave., Novosibirsk, 630090, Russian Federation

<sup>2</sup> A. P. Ershov Institute of Informatics Systems SB RAS  
6 Academician Koptyug Ave., Novosibirsk, 630090, Russian Federation

antsys\_nsu@mail.ru, zyubin@iae.nsk.su, garanina@iis.nsk.su

## AUTOMATIC VERIFICATION OF CONTROL ALGORITHMS FOR COMPLEX TECHNOLOGICAL OBJECTS ON SOFTWARE SIMULATORS

The article is devoted to the analysis of process-oriented control algorithms, on the correspondence to input specifications. Article presents the general scheme of verification of the control algorithm, its implementation in an automated form and the results of practical approbation of the approach in the project for the automation of the Large Solar Vacuum Telescope.

*Keywords:* automation, industrial control algorithms, verification, process-oriented programming languages.

### References

1. Zakrevsky A. D. Parallel algorithms of logical control. Moscow, Editorial URSS, 2003. 200 p. (in Russ.)
2. Wagner F., Schmuki R., Wagner T., Wolstenholme P. Modeling software with finite state machines: a practical approach. *Auerbach Publications*, 2006. 390 p.
3. Harel D. Statecharts: A visual formalism for complex systems. *Science of computer programming*, 1987, 8.3, p. 231–274.
4. Zyubin V. E. Programming of PLC: languages IEC 61131-3 and possible alternatives. *Industrial control systems and controllers*, 2005, no. 11, p. 31–35. (in Russ.)
5. Lyakh T. V., Zyubin V. Y., Sizov M. M. Experience of Reflex Language Application at Automation of the Large Solar Vacuum Telescope. *Industrial Automated Control Systems and Controllers*, 2016, no. 7, p. 37–43. (in Russ.)
6. Zyubin V. E. To the fifth anniversary of the IEC 1131-3 standard. Results and forecasts. *Devices and systems. Management, control, diagnostics*, 1999, no. 1, p. 64–71. (in Russ.)
7. Samek M., Montgomery P. State Oriented Programming. *Embedded Systems Programming*, August 2000, p. 22–43.
8. Shalyto A. A., Tukkell N. I. SWITCH technology – automatic approach to software development of “reactive” systems. *Programming*, 2001, no. 5, p. 45–62. (in Russ.)
9. Rozov A. S., Zyubin V. E., Nefedov D. V. Programming of Embedded Microcontroller Systems Based on Hyperprocesses. *Vestnik NSU. Series: Information Technologies*, 2017, vol. 15, no. 4, p. 64–73. (in Russ.)
10. Zyubin V. E. “C with processes” – programming language of logic controllers. *Mechatronics, automation, control*, 2006, no. 12, p. 31–35. (in Russ.)
11. Zyubin V. E. Iterative development of control algorithms based on simulation of control object. *Automation in industry*, 2010, no. 11, p. 43–48. (in Russ.)
12. Rozov A. S., Lyakh T. V., Krasnov D. V., Sanzhiev E. S. Practical approbation of the IndustrialC language on the example of automation of the thermovacuum sputtering installation. *Vestnik NSU. Series: Information Technologies*, 2017, vol. 15, no. 3, p. 90–99. (in Russ.)
13. Garanina N. O., Zyubin V. E., Lyakh T. V. Ontological approach to the organization of requirement templates within the system of support of formal verification of software systems. *System Informatics*, 2017, no. 9, p. 111–132. (in Russ.)
14. Shilov N. V., Garanina N. O. Combined Logics of Knowledge, Time, and Actions for Reasoning about Multi-agent Systems. Knowledge Processing and Data Analysis. *Lecture Notes in Computer Science*, 2011, iss. 6581, p. 48–58.

15. Shelekhov V. I. Verification and synthesis of addition programs based on the correctness rules of operators. *Modeling and analysis of information systems*, 2010, vol. 17, no. 4, p. 101–110. (in Russ.)
16. Clarke E. M., Gao S. Model checking hybrid systems. *International Symposium on Leveraging Applications of Formal Methods, Verification and Validation. Lecture Notes in Computer Science*, 2014, iss. 8803, p. 385–386.
17. Clarke E. M., Grumberg O., Peled D. Model checking. MIT Press, 1999, 314 p.
18. Clarke E. M., Emerson A., Sistla P. Automatic verification of the finite-state concurrent systems using temporal logic specifications. *ACM Transactions on Programming Languages and Systems*, 1986, vol. 8, iss. 2, p. 244–263.
19. Isermann R., Schaffnit J., Sinsel S. Hardware-in-the-loop simulation for the design and testing of engine-control systems. *Control Engineering Practice*, 1999, vol. 7, p. 643–653.
20. Xiao B., Starke M., King D., Irminger Ph., Herron A., Ollis B., Xue Y. Implementation of the system-level control and communications in a Hardware-in-the-Loop microgrid testbed. *IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 2016, p. 1–5.
21. Zyubin V. E. Use of virtual objects for teaching programming of information-control systems. *Information technologies*, 2009, no. 6, p. 79–82. (in Russ.)
22. Liakh T. V., Zyubin V. E. Automatic verification of control algorithms for complex technological objects on software simulators. *Proc. of the V International Scientific Conference Mathematical and computer modeling dedicated to the memory of R. L. Dolganov*. Omsk, 2017, p. 128–130. (in Russ.)
23. Zyubin V. E., Kalugin A. A. Virtual laboratory stands: teaching programming of problems of industrial automation. *Industrial Automated Control Systems and Controllers*, 2009, no. 2, p. 39–44. (in Russ.)
24. Kovadlo P. G., Lubkov A. A., Bezvov A. N., Budnikov K. I., Vlasov S. V., Zotov A. A., Kolobov D. U., Kurochkin A. V., Kotov V. N., Lylov S. A., Lyakh T. V., Maksimov A. S., Perebeynos S. V., Petukhov A. D., Peshcherov V. S., Popov U. A., Russkih I. V., Tomin V. E. Automation system of the Large Solar Vacuum Telescope. *Optoelectronics, Instrumentation and Data Processing*, 2016, vol. 52, no. 2, p. 97–106. (in Russ.)

Received 23.09.2018

*For citation:*

Liakh T. V., Zyubin V. E. Garanina N. O. Automatic Verification of Control Algorithms for Complex Technological Objects on Software Simulators. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 4, p. 85–94. (in Russ.) DOI 10.25205/1818-7900-2018-16-4-85-94

**Е. Д. Махина<sup>1</sup>, Д. Е. Пальчунов<sup>1,2</sup>**

<sup>1</sup> *Новосибирский государственный университет  
ул. Пирогова, 2, Новосибирск, 630090, Россия*

<sup>2</sup> *Институт математики им. С. Л. Соболева СО РАН  
пр. Академика Коптюга, 4, Новосибирск, 630090, Россия*

*ekatermakhina96@gmail.com, palch@math.nsc.ru*

## **ПРОГРАММНАЯ СИСТЕМА ДЛЯ ОПРЕДЕЛЕНИЯ РЕЧЕВЫХ ДЕЙСТВИЙ В ТЕКСТАХ ЕСТЕСТВЕННОГО ЯЗЫКА \***

Статья посвящена разработке методов извлечения речевых действий из текстов естественного языка. В настоящее время одной из наиболее важных проблем является проблема создания человеко-машинных интерфейсов. В данной работе рассматривается задача построения подобных интерфейсов путем автоматизированного извлечения речевых действий из текстов естественного языка. Использование речевых действий позволяет задавать контекст диалога и уточнять смысл употребляемых понятий. По результатам исследования была разработана программная система, которая способна определять типы речевых действий, представленных в текстах естественного языка. Одно из возможных применений результатов данного исследования заключается в использовании разработанной программной системы при создании чат-ботов.

*Ключевые слова:* речевое действие, иллокутивная сила, теоретико-модельные методы, алгебра Линденбаума – Тарского, фрагменты атомарных диаграмм, онтология, естественно-языковые интерфейсы.

### **Введение**

В настоящее время проектирование, построение и пополнение семантических и онтологических моделей различных предметных областей являются актуальными задачами, решаемыми как при автоматизации бизнес-процессов предприятий, так и при разработке программных продуктов для конечных пользователей. Одной из наиболее острых проблем является проблема предоставления интерфейсов на естественном языке для коммуникации человека и компьютера. В частности, такие проблемы ставятся и решаются при создании чат-ботов.

Компьютер не всегда способен «понять», какую цель хотел достичь человек той или иной фразой, исходя только из синтаксиса текста диалога с пользователем. Интерфейсы на естественном языке призваны помочь определять то, чего хочет человек, посредством не только синтаксического, но и семантического анализа текстов естественного языка.

Существуют разные подходы к разработке интерфейсов на естественном языке [1–5]. В рамках данной работы мы используем теоретико-модельные методы [6–11] и теорию речевых действий для создания естественно-языковых интерфейсов.

---

\* Исследование выполнено при частичной финансовой поддержке Президиума СО РАН (проект «Инженерия интенциональных онтологий в дедуктивных и информационных системах» Комплексной программы ФНИ СО РАН П.1).

Целью настоящего исследования является разработка и программная реализация алгоритма определения типов речевых действий, представленных в текстах естественного языка. Достижение данной цели необходимо для решения более общей задачи – извлечения речевых действий из текстов естественного языка и их формального (теоретико-модельного) представления. Извлечённые речевые действия представляются в виде элементов куба алгебры Линденбаума – Тарского теории рассматриваемой предметной области [12; 13].

Одной из важных областей применения результатов данной работы является их использование при разработке чат-ботов нового поколения, которые по своим возможностям будут существенно превосходить существующие. Рассмотрим, например, чат-бот «Алиса» – виртуальный голосовой помощник, созданный компанией Яндекс. Алиса не создает контекста общения с собеседником, и по этой причине не всегда корректно отвечает на вопросы человека. Мы же хотим строить и далее использовать модель диалога между пользователем и чат-ботом.

На основе знаний о диалоге, представленных в модели, будут формироваться дальнейшие реплики чат-бота. Модель формально представляется в виде фрагмента атомарной диаграммы алгебраической системы, а речевые действия – в виде троек формул: <сообщение, побуждение, заявление>. Каждая такая тройка является элементом куба алгебры Линденбаума – Тарского.

Необходимо отметить, что целью данной работы является создание не полностью автоматической, а автоматизированной системы распознавания речевых действий, так как существуют проблемы, которые на настоящем этапе не представляется возможным решать полностью автоматически. К таким проблемам можно отнести полное описание контекста для определения пропозициональной составляющей и иллокутивной силы речевого действия (в частности, обработку пресуппозиций, содержащихся в тексте естественного языка [14; 15]).

### **Теория речевых действий**

Разработка теории речевых действий была начата Дж. Л. Остином [16; 17] и детально разработана Дж. Р. Серлем и Д. Вандервекенем [18–28]. Нами была предложена определенная ревизия классической теории речевых действий [12; 13].

В теории речевых действий диалог между собеседниками рассматривается как совокупность речевых действий. В речевом действии выделяют две основных составляющих: высказывание (пропозициональная составляющая) и иллокутивная сила. Мы расширяем этот подход и рассматриваем текст естественного языка как последовательность речевых действий. Методы определения речевых действий в текстах естественного языка в рамках классической теории речевых действий исследовались в [29].

Высказывание (пропозициональная составляющая речевого действия) отражает общее содержание предложения. Мы формально описываем пропозициональные составляющие речевых действий при помощи фрагментов атомарных диаграмм (т. е. конъюнкций атомарных предложений логики предикатов первого порядка).

В классической теории речевых действий иллокутивная сила раскладывается на 7 компонентов:

- иллокутивная цель;
- мера стремления к иллокутивной цели;
- способ достижения иллокутивной цели;
- ограничения на пропозицию;
- предварительные условия;
- условие искренности;
- сила выражения условия искренности.

В данной работе мы рассматриваем две составляющие иллокутивной силы: иллокутивную цель и интенсивность выражения иллокутивной цели (т. е. меру стремления к иллокутивной цели).

Классическая теория речевых действий предлагает следующую классификацию речевых действий:

- 1) репрезентативы – речевые действия, целью которых является сообщение об общем положении дел;
- 2) директивы – речевые действия, побуждающие совершать или не совершать какие-либо действия;
- 3) комиссивы – речевые действия, направленные на возложение ответственности на говорящего за то, чтобы сделать что-то в будущем;
- 4) экспрессивы – речевые действия, служащие для выражения эмоционального состояния говорящего;
- 5) декларативы – речевые действия, результатом которых является осуществление главной цели высказывания, т. е. после их озвучивания положение дел в реальном мире соответствует сказанному.

Стоит заметить, что данная классификация не представляет собой разбиение на непересекающиеся подмножества, т. е. одно речевое действие может быть представителем нескольких классов.

Разработанная нами ранее классификация речевых действий [12; 13] выделяет три типа речевых действий (более точно – три типа компонентов речевых действий): сообщение, заявление, побуждение. Каждое речевое действие может иметь один, два или три компонента. Указанные типы речевых действий сопоставимы со следующими аспектами сознания собеседника: знания (прошлое, настоящее или будущее), представление о текущей ситуации (настоящее) и намерения (будущее). Например, заявление может менять контекст диалога, в частности изменять интерпретацию, смысл употребляемых понятий.

Таким образом, с одной стороны, речевое действие представляется как пара  $\langle F, \varphi \rangle$  (иногда пишут  $F\langle\varphi\rangle$ ), где  $F$  – иллокутивная сила, а  $\varphi$  – пропозициональная составляющая, а с другой стороны, как тройка предложений логики предикатов первого порядка  $(\varphi_1, \varphi_2, \varphi_3)$ , где  $\varphi_1$  – сообщение,  $\varphi_2$  – заявление, а  $\varphi_3$  – побуждение. При этом формула  $\varphi_i$  может быть тождественно истинной. Это означает, что  $i$ -й компонент речевого действия отсутствует.

Заметим, что классический подход к теории речевых действий обладает рядом существенных недостатков [12; 13].

1. Переход от синтаксиса к семантике, определяемый только через иллокутивные глаголы.

Однако следующие примеры являются разными речевыми действиями с точки зрения классической теории речевых действий, но обозначают они одно и то же желание говорящего, поэтому с точки зрения семантики не имеет смысла разделять их:

- a. Закрой окно.
- b. Я прошу тебя закрыть окно.
- c. Я заявляю свое требование закрыть окно.
- d. Я сообщаю тебе свою просьбу – закрой окно.

2. Наличие в каждом речевом действии не более одной иллокутивной цели. Из-за этого классификация иллокутивных сил становится очень громоздкой.

3. Отсутствие точного определения классов речевых действий. Например, отсутствует четкий критерий, позволяющий ответить на вопрос, что представляет собой заявление. Подобного рода определения крайне важны, так как в настоящей работе для программной реализации необходимы четкие правила для определения типа речевого действия.

4. Изолированность речевого действия от диалога. Речевое действие рассматривается отдельно, как бы «в вакууме».

5. «Юридическая парадигма»: на самом деле меняет мир не слово, а его фиксация в рамках определенного контекста – протокола, регламента и т. д.

Основными преимуществами предлагаемого подхода к классификации речевых действий [12; 13] по сравнению с подходом Остина и Серля является следующее.

1. Речевое действие является композицией компонентов трех типов: сообщение, заявление и побуждение.

2. Каждый тип речевого действия имеет четкое определение.

3. Речевое действие рассматривается в контексте диалога. Это означает, что мы имеем представление о текущей ситуации (например, разговаривают ли два друга либо работода-

тель и подчиненный), и в зависимости от этого мы можем делать те или иные выводы о принадлежности речевого действия к тем или иным классам.

4. Меняется не мир, а собеседник: его знания, представления о текущей ситуации, намерения.

### Определение типа речевого действия

Алгоритм автоматизированного извлечения речевых действий из текстов естественного языка можно представить в виде последовательности шагов:

- 1) распознавание речевого действия в предложении;
- 2) извлечение его пропозициональной части;
- 3) извлечение иллокутивной цели и интенсивности речевого действия.

Одной из задач данной работы является определение типа речевого действия, представленного в предложении естественного языка. Для наглядности рассмотрим следующие примеры:

- 1) Я объявляю войну (заявление)
- 2) Делай уроки! (побуждение)
- 3) В пятницу будет контрольная работа (сообщение)
- 4) В пятницу будет контрольная работа, все должны на нее прийти (сообщение + заявление)
- 5) Можно я оставлю у себя твою книгу еще на неделю, пожалуйста? (побуждение + заявление)
- 6) Пообещай, что больше не будешь так делать (побуждение)
- 7) Какое красивое небо! (заявление)

Разберем подробно пример предложения, содержащего все компоненты речевого действия: «Завтра в 9 часов утра вы должны прийти на сдачу долгов». Во-первых, в данном предложении присутствует сообщение, семантически выраженное в следующем: «Завтра в 9 часов утра состоится сдача долгов». Во-вторых, заявление является частью речевого действия, так как отдается приказ быть всем завтра на сдаче долгов. В-третьих, говорящий (предположим, что это преподаватель) побуждает всех прийти, т. е. он заинтересован в том, чтобы завтра все присутствовали на сдаче долгов, а иначе самому преподавателю не имеет смысла приходить. Перечисленные составляющие представляют 3 разных пропозиции, выделенные из основного предложения.

Для определения того или иного типа речевого действия можно задать определенные вопросы. Если ответы на эти вопросы могут быть получены из предложения, то речевое действие относится к тому или иному типу; в противном случае речевое действие не содержит искомого компонента.

Ниже перечислены вопросы, являющиеся содержательными критериями проверки наличия в речевом действии каждой из трех составляющих. Другими словами, на какие вопросы необходимо ответить, чтобы убедиться, что речевое действие имеет тот или иной компонент. Предполагается, что целью является определение типа речевого действия фраз субъекта *a* в диалоге с субъектом *b*.

1. Ставил ли целью говорящий *a* передать собеседнику *b* какие-то знания (об уже совершившихся, текущих или будущих событиях)?
2. Было ли целью говорящего установить, что имела место определенная речевая ситуация: вопрос, просьба, приказ, мольба и т. д.?
3. Хочет ли говорящий *a*, чтобы у собеседника *b* после разговора было намерение что-то сделать (при этом не имеет значения, было ли такое намерение у *b* до разговора; важно, чтобы после диалога это намерение появилось)?

Данные вопросы можно задавать пользователю в интерактивном режиме в том случае, если нельзя однозначно определить тип речевого действия. Также в случае невозможности однозначной классификации речевого действия можно выдавать пользователю несколько вариантов, отличающихся в зависимости от контекста.

Помимо этого, мы можем задавать вопросы, помогающие выяснить содержание каждой конкретной составляющей речевого действия:

1.  $\varphi_1(a, b) = \text{“что } b \text{ узнал нового?”}$
2.  $\varphi_2(a, b) = \text{“что произошло во время разговора } a \text{ и } b \text{ (какая речевая ситуация имела место)?”}$
3.  $\varphi_3(a, b) = \text{“что намеревается сделать } b\text{?”}$ .

В ходе работы был составлен следующий список правил определения типа речевого действия.

#### **1. Побуждение:**

- a. повелительное наклонение глагола
- b. условное наклонение глагола;
- c. будущее время глагола + 2-е или 3-е лицо;
- d. глагол «давай»;
- e. слово «пожалуйста»;
- f. слова «следует / должен / обязан / необходимо / нужно / надо» и т. д.;
- g. можешь ли ты / не мог бы ты / почему бы тебе не и т. д.

#### **2. Сообщение:**

- a. настоящее время глагола;
- b. прошедшее время глагола;
- c. 1-е или 3-е лицо.

#### **3. Заявление:**

- a. «я хотел бы / мне хотелось бы»;
- b. настоящее время глагола.

#### **4. Заявление + побуждение:**

- a. было бы хорошо / чудесно / здорово и т. д., если бы...;
- b. любой изолированный вопрос.

Видно, что некоторые правила совпадают для разных типов речевых действий. Это означает, что в исходном предложении речевое действие представляет собой комбинацию типов, например сообщение и побуждение, а возможно, и все три типа одновременно.

Предложенный список правил не является финальной версией алгоритма, используемого в программе, т. е. данный список является расширяемым и пополняемым в процессе использования программной системы для разных текстов русского языка.

### **Этапы анализа предложения**

Алгоритм разработанной программной системы состоит из следующих шагов. На вход программной системе подается предложение на русском языке. В данном исходном предложении осуществляется морфологический анализ слов. После этого предложение проверяется на наличие каждого из типов речевого действия (сообщение, заявление, побуждение). На выходе программа выдает найденные типы компонентов речевого действия, представленного в данном предложении.

Опишем этапы работы программной системы более подробно. Итоговой идентификации типа речевого действия в предложении предшествует этап морфологического анализа предложения. Несмотря на то что данный этап не играет определяющей роли в идентификации типа речевого действия, он очень важен для предобработки предложения, так как именно на этом этапе происходит первоначальный выбор того или иного типа речевого действия.

Для морфологического анализа предложения естественного языка используется программа MyStem, предоставленная компанией Яндекс для открытого использования. Программа делает следующее: на вход подается предложение русского языка, а на выходе мы получаем

всю информацию о каждом слове – часть речи, время и наклонение для глаголов, падеж для существительных и т. п. К системе MyStem можно подключить свой словарь; также она позволяет строить гипотетические разборы для слов, не входящих в словарь.

Рассмотрим пример работы программы: для входного файла `sapgir.txt`, содержащего текст “В мурели шлепают пельсиски. В стакелках светится мычай.”, вывод будет следующим:

```
$ mystem -nig sapgir.txt
V{в=PR=|в=S, сокр=(им, ед|им, мн|род, ед|род, мн|дат, ед|дат, мн|вин, ед|вин, мн|твор, ед|твор, мн|п
р, ед|пр, мн)}
мурелки{мурелка?=S, жен, од=(им, мн|род, ед)|мурелка?=S, жен, неод=(им, мн|род, ед|вин, мн)}
шлепают{шлепать=V, несов, пе=непрош, мн, изъяв, 3-л}
пельсиски{пельсиск?=S, имя, муж, од=им, мн|пельсиска?=S, жен, неод=(им, мн|род, ед|вин, мн)}
V{в=PR=|в=S, сокр=(им, ед|им, мн|род, ед|род, мн|дат, ед|дат, мн|вин, ед|вин, мн|твор, ед|твор, мн|п
р, ед|пр, мн)}
стакелках{стакелка?=S, жен, неод=пр, мн|стакелка?=S, жен, од=пр, мн|стакелок?=S, муж, неод=пр, мн|
стакелка?=S, имя, жен, од=пр, мн|стакелк?=S, имя, муж, од=пр, мн|стакелк?=S, муж, неод=пр, мн}
светится{светиться=V, несов, нп=непрош, ед, изъяв, 3-л}
мычай{мычай?=S, муж, неод=(им, ед|вин, ед)}
```

Описанная программа встроена в разработанную программную систему SpeechActs и используется для предварительной обработки предложений.

После проведения морфологического разбора предложения мы переходим к его семантическому анализу. Этот этап играет большую роль в определении типа (комбинации типов) речевого действия, так как первоначально мы выбираем тип в соответствии с определенными морфологическими формами слов, а затем корректируем выбранный тип посредством семантического анализа. Так, к примеру, предложение “Какая прекрасная погода сегодня!” на этапе морфологического анализа мы определим как сообщение, а после семантического анализа поймем, что в действительности это заявление, так как данным предложением говорящий хотел выразить свое отношение к погоде, а в другом контексте это может быть побуждением отдохнуть на природе.

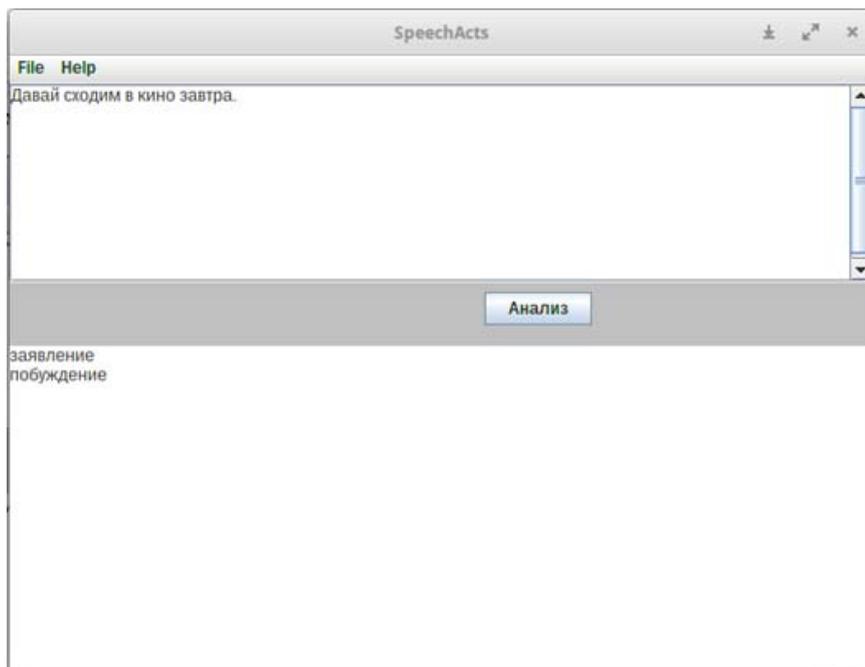
На этапе семантического анализа мы ищем употребленные в предложении глаголы в словаре иллокутивных глаголов, встроенном в программу. Нахождение глагола в одном из словарей определяет принадлежность речевого действия тому или иному типу.

Таким образом, на данном этапе мы подтверждаем выбор типа речевого действия, осуществленный на предыдущем этапе (этапе морфологического анализа), либо добавляем к уже выбранному типу еще один тип речевого действия, определенный в ходе семантического анализа.

## Интерфейс программной системы

Алгоритм работы программной системы заключается в следующем: на вход подается предложение на русском языке. В исходном предложении осуществляется морфологический анализ каждого слова в предложении, после чего предложение проверяется на наличие каждого из типов речевого действия (сообщение, заявление, побуждение). На выходе программа выдает найденные типы речевых действий в предложении из входных данных.

Ниже приведен скриншот пользовательского интерфейса первой версии программной системы.



## Архитектура программной системы

Разработанная программная система состоит из следующих компонентов.

### 1. Controller

Контроллер отвечает за инициализацию всех частей приложения. Он запускает остальные компоненты, а также обеспечивает передачу данных от графического интерфейса к модели и обратно.

### 2. LogicText

Модуль, запускающий программную систему LogicText, разработанную О. Г. Махасоевой и Д. Е. Пальчуновым [8; 9]. Программа делает следующее: каждое слово в предложении приводит к нормальной форме, выбирает из них субъект и объект; само действие является предикатом. В программу встроен словарь с некоторыми известными предикатами, где с каждым глаголом сопоставлены вопросы (кто, что, где, когда и т. д.). Алгоритм работы программной системы LogicText основан на применении теории «Смысл-Текст» И. А. Мельчука [30].

### 3. MyStem

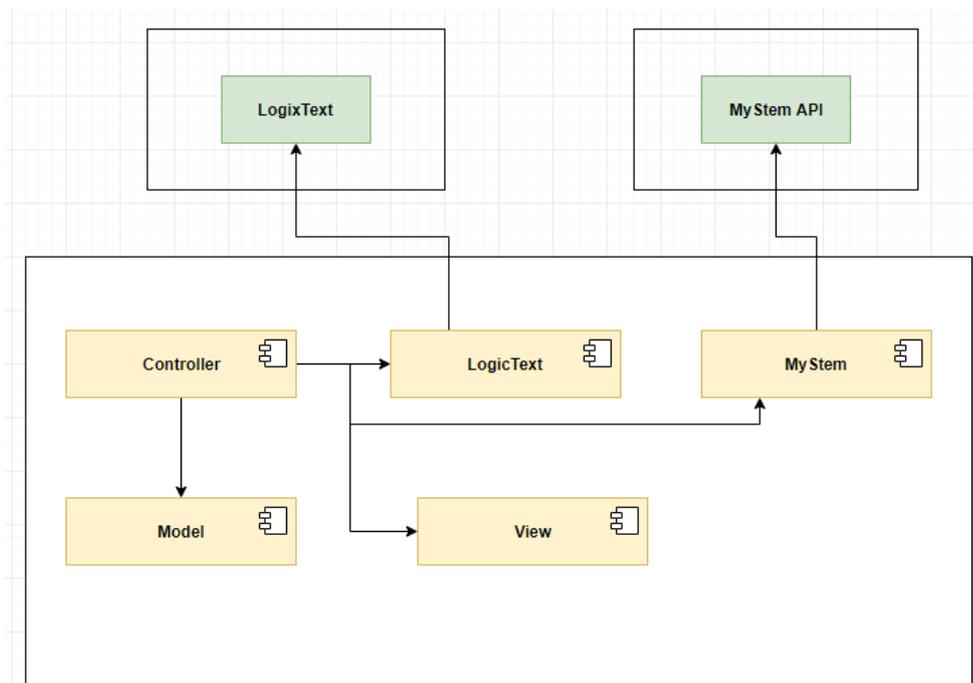
Модуль, который обрабатывает каждое слово в предложении и возвращает полную характеристику каждого из слов.

### 4. Model

Модель представляет собой главный компонент, в котором происходит обработка предложения, его анализ и определение типа речевого действия.

### 5. View

Данный модуль отвечает за отображение введенного пользователем предложения, а также за визуализацию ответа программной системы.



Речевое действие представлено в программной системе в виде композиции трёх компонентов: (сообщение, заявление, побуждение). Если компонент определённого типа присутствует в речевом действии, то приводится соответствующее правило, которое позволило определить этот тип.

## Заключение

В данной работе были проведены исследования речевых действий, представленных в текстах естественного языка. В результате сделана классификация иллокутивных глаголов, соответствующих различным типам речевых действий, а также разработан алгоритм по определению типа (или комбинации типов) речевого действия.

На основе разработанного алгоритма была реализована программная система SpeechActs, которая позволяет определять тип речевого действия по предложению на естественном языке. Программная система также позволяет выводить информацию о том, по какому признаку (правилу) был определен тип речевого действия. Программная система SpeechActs интегрирована с программной системой LogicText, на основе которой она производит формальное представление информации, содержащейся в предложении естественного языка.

Разработанная программная система может быть применена для построения модели диалога, а также для построения модели пользователя при его общении с чат-ботом. Извлечение речевых действий дает возможность формального представления контекста диалога. В частности, речевые действия могут определять и переопределять смысл (определения) используемых понятий в контексте данного диалога.

Для дальнейшего совершенствования разработанной программной системы необходимо добавлять новые правила для более точного определения типа речевого действия. Помимо этого, планируется использовать методы определения интенсивности речевого действия, а также предполагается разработать методы построения фрагментов атомарных диаграмм для полного извлечения речевых действий из текстов естественного языка.

## Список литературы

1. Елисеева О. Е. Естественнo-языковой интерфейс интеллектуальных систем: Учеб. пособие. Минск: Изд-во БГУИР, 2009. С. 84–85.
2. Dezhao S., Schilder F., Smiley C., Brew C., Zielund T., Bretz H., Martin R., Dale C., Pomerville S., Duprey J., Miller T., Harrison J. TR Discover: a natural language interface for querying and analyzing interlinked datasets // Proc. 14<sup>th</sup> Intern. conf. on the Semantic Web: ISWC 2015. Springer Intern. Publ., 2015. P. 21–37.
3. Крайванова В. А. Модель естественнo-языкового интерфейса для систем управления сложными техническими объектами и оценка эффективности алгоритмов на ее основе // Управление большими системами. М.: Изд-во ИПУ РАН, 2009. Вып. 26. С. 158–178.
4. Deshpande A. K., Devale P. R. Natural language query processing using probabilistic context free grammar // Intern. Journ. of Advances in Engineering & Technology. 2012. Vol. 3. No. 2. P. 568–573.
5. Li F., Jagadish H. V. NaLIR: An interactive natural language interface for querying relational databases // Proc. 2014 ACM SIGMOD Intern. conf. on Management of data, ACM. 2014. P. 709–712.
6. Пальчунов Д. Е. Решение задачи поиска информации на основе онтологий // Бизнес-информатика. 2008. № 1 (3). С. 3–13.
7. Dmitry E. Palchunov. Virtual catalog: the ontology-based technology for information retrieval // Knowledge Processing and Data Analysis. LNAI 6581. Berlin; Heidelberg: Springer-Verlag, 2011. P. 164–183.
8. Махасоева О. Г., Пальчунов Д. Е. Автоматизированные методы построения атомарной диаграммы модели по тексту естественного языка // Вестн. НГУ. Серия: Информационные технологии. 2014. Т. 12, № 2. С. 64–73.
9. Махасоева О. Г., Пальчунов Д. Е. Программная система построения атомарной диаграммы модели по тексту естественного языка. Св-во о гос. рег. программы для ЭВМ № 2014619198 от 10.09.2014.
10. Деревянко Д. В., Пальчунов Д. Е. Формальные методы разработки вопросно-ответной системы на естественном языке // Вестн. НГУ. Серия: Информационные технологии. 2014. Т. 12, № 3. С. 34–47.
11. Корсун И. А., Пальчунов Д. Е. Теоретико-модельные методы извлечения знаний о смысле понятий из текстов естественного языка // Вестн. НГУ. Серия: Информационные технологии. 2016. Т. 14, № 3. С. 34–48.
12. Пальчунов Д. Е. О логическом анализе естественного языка // Теория вычислений и языка спецификаций (вычислительные системы). 1995. Вып. 152. С. 61–75.
13. Dmitri E. Pal'chunov. Algebraische Beschreibung der Bedeutung von Aeusserungen der natuerlichen Sprache // Zelger, Josef; Maier, Martin (Hrsg.). GABEK. Verarbeitung und Darstellung von Wissen. Innsbruck; Wien: STUDIEN Verlag, 1999. S. 310–326.
14. Пальчунов Д. Е. Поиск и извлечение знаний: порождение новых знаний на основе анализа текстов естественного языка // Философия науки. 2009. № 4. С. 70–90.
15. Пальчунов Д. Е., Целищев В. В. Проблема извлечения знаний в системе взаимодействия человека и компьютера (онтологии и пресуппозиции) // Философия науки. 2012. № 4 (55). С. 20–35.
16. Остин Дж. Слово как действие // Новое в зарубежной лингвистике. М., 1986. Вып. 17: Теория речевых актов. С. 22–129.
17. Austin J. L. How to do things with words. Oxford Uni. Press, 1975.
18. Серль Дж. Р. Классификация иллокутивных актов // Новое в зарубежной лингвистике. М., 1986. Вып. 17: Теория речевых актов. С. 170–194.
19. Searle J. R. Indirect speech acts. 1975.
20. Searle J. R. A classification of illocutionary acts // Language in society. 1976. Т. 5, № 1. С. 1–23.
21. Searle J. R., Vanderveken D. Foundations of illocutionary logic. CUP Archive, 1985.
22. Searle J. R. Speech acts: An essay in the philosophy of language. CUP, 1969. Vol. 626.
23. Searle J. R. Expression and meaning: Studies in the theory of speech acts. CUP, 1985.

24. *Searle J. R. Austin on Locutionary and Illocutionary Acts // Philosophical Review. 1968. Vol. 77. No. 4. P. 405–424.*
25. *Vanderveken D. Meaning and speech acts: Volume 1, principles of language use. CUP, 1990. Vol. 1.*
26. *Friggieri J. et al. Actions and Speech Actions in the Philosophy of JL Austin. 1993.*
27. *Серль Дж. Р., Вандервекен Д. Основные понятия исчисления речевых актов // Новое в зарубежной лингвистике. М., 1986. Вып. 18: Логический анализ естественного языка. С. 242–264.*
28. *Vanderveken D. Fondements die la logique des attitudes // Manuscripto. 2006. Vol. 29. No. 2.*
29. *Бехтерева С. О. Анализ текстов на естественном языке на основании Теории речевых действий. Выпускная квалификационная работа бакалавра. Новосибирск, 2012.*
30. *Смыслы, тексты и другие захватывающие сюжеты: Сб. ст. в честь 80-летия И. А. Мельчука / Под ред. Ю. Д. Апресяна, И. М. Богуславского, Л. Ваннера, Л. Л. Иомдина, Я. Миличевич, М.-К. Л'Омм, А. Польгера. М.: Языки славянской культуры, 2012. 704 с.*

*Материал поступил в редколлегию 07.08.2018*

**E. D. Makhina<sup>1</sup>, D. E. Palchunov<sup>1,2</sup>**

<sup>1</sup>*Novosibirsk State University  
1 Pirogov Str., Novosibirsk, 630090, Russian Federation*

<sup>2</sup>*S. L. Sobolev Institute of Mathematics SB RAS  
4 Academician Koptuyug Ave., Novosibirsk, 630090, Russian Federation*

*ekatermakhina96@gmail.com, palch@math.nsc.ru*

## **SOFTWARE SYSTEM FOR THE DETERMINATION OF SPEECH ACTS IN NATURAL LANGUAGE TEXTS**

The paper is devoted to the development of methods of speech act extraction from natural language texts. Currently, one of the most significant problems is the problem of creating human-computer interfaces. The article discusses the task of development such interfaces by automated extraction of speech acts from natural language texts. The use of speech acts allows us to set the context of the dialogue and clarify the meaning of the concepts used. According to the results of the study, we develop a software system that is able to determine the types of speech acts represented in the natural language texts. One of the possible applications of the results of this study is to use the developed software system when creating chat bots.

*Keywords:* speech act, illocutionary force, model-theoretical methods, Lindenbaum – Tarski algebra, atomic diagram fragment, ontology, natural language interface.

### **References**

1. *Eliseeva O. E. Estestvenno-yazykovoy interfeys intellektualnykh sistem [Natural Language Interface of Intellectual Systems]. Study guide. Minsk, BSUIR Publ., 2009, p. 84–85. (in Russ.)*
2. *Dezhao S., Schilder F., Smiley C., Brew C., Zielund T., Bretz H., Martin R., Dale C., Pomerville S., Duprey J., Miller T., and Harrison J. TR Discover: a natural language interface for querying and analyzing interlinked datasets. *Proc. 14<sup>th</sup> Intern. conf. on the Semantic Web: ISWC 2015*. Springer Intern. Publ., 2015, p. 21–37.*

3. Krayvanova V. A. Model of natural-language interface for complex technical objects control systems and efficiency evaluation of model-based algorithms. *Upravlenie bolshimi sistemami [Large-scale Systems Control]*. Moscow, Institute of Control Sciences Publ., 2009, vol. 26, p. 158–178 (in Russ.)
4. Deshpande A. K., Devale P. R. Natural language query processing using probabilistic context free grammar. *Intern. Journ. of Advances in Engineering & Technology*, 2012, vol. 3, no. 2, p. 568–573.
5. Li F., Jagadish H. V. NaLIR: An interactive natural language interface for querying relational databases. *Proc. 2014 ACM SIGMOD Intern. conf. on Management of data, ACM*, 2014, p. 709–712.
6. Palchunov D. E. The solution of the problem of information retrieval based on ontologies. *Bisnes-informatika*, 2008, no. 1, p. 3–13. (in Russ.)
7. Dmitry E. Palchunov. Virtual catalog: the ontology-based technology for information retrieval. *Knowledge Processing and Data Analysis. LNAI 6581*. Berlin, Heidelberg, Springer Verlag, 2011, p. 164–183.
8. Makhasoeva O. G., Palchunov D. E. Semi-automatic methods of a construction of the atomic diagrams from natural language texts. *Vestnik NSU. Series: Information Technologies*, 2014, vol. 12, no. 2, p. 64–73. ISSN 1818-7900 (in Russ.)
9. Makhasoeva O. G., Palchunov D. E. Program system for the construction of the atomic diagram of a model from natural language texts. Certificate of the State Registration of the computer program. No. 2014619198, registered 10.09.2014. (in Russ.)
10. Derevyanko D. V., Palchunov D. E. Formal methods of development of the question-answering system on natural language. *Vestnik NSU. Series: Information Technologies*, 2014, vol. 12, no. 3, p. 34–47. (in Russ.)
11. Korsun I. A., Palchunov D. E. Model-Theoretic Methods of Extraction of Knowledge on the Meaning of Concepts from the Natural Language Texts. *Vestnik NSU. Series: Information Technologies*, 2016, vol. 14, no. 3, p. 34–48. (in Russ.)
12. Palchunov D. E. On the logical analysis of natural language. *Theory of calculations and specification languages*. Novosibirsk, 1995, vol. 152: Computing systems, p. 61–75. (in Russ.)
13. Dmitri E. Pal'chunov. Algebraische Beschreibung der Bedeutung von Aeusserungen der natuerlichen Sprache. In: Zelger, Josef; Maier, Martin (Hrsg.). GABEK. Verarbeitung und Darstellung von Wissen. Innsbruck, Wien, STUDIEN Verlag, 1999, S. 310–326.
14. Palchunov D. E. Knowledge retrieval and elicitation: generation of new knowledge based on analysis of natural language texts. *Filosofiya nauki*, 2009, no. 4 (43), p. 70–90. (in Russ.)
15. Palchunov D. E., Tselishchev V. V. The problem of knowledge retrieval in interaction between a man and a computer: ontologies and presuppositions. *Filosofiya nauki*, 2012, no. 4 (55), p. 20–35. (in Russ.)
16. Austin J. L. Word as action. *New in foreign linguistics*. Moscow, 1986, iss. 17: The theory of speech acts, p. 22–129.
17. Austin J. L. How to do things with words. Oxford Uni. Press, 1975.
18. Searle J. R. Classification of illocutionary acts. *New in foreign linguistics*. Moscow, 1986, iss. 17: The theory of speech acts, p. 170–194.
19. Searle J. R. Indirect speech acts. 1975.
20. Searle J. R. A classification of illocutionary acts. *Language in society*, 1976, vol. 5, no. 1, p. 1–23.
21. Searle J. R., Vanderveken D. Foundations of illocutionary logic. CUP Archive, 1985.
22. Searle J. R. Speech acts: An essay in the philosophy of language. CUP, 1969, vol. 626.
23. Searle J. R. Expression and meaning: Studies in the theory of speech acts. CUP, 1985.
24. Searle J. R. Austin on Locutionary and Illocutionary Acts. *Philosophical Review*, 1968, vol. 77, no. 4, p. 405–424.
25. Vanderveken D. Meaning and speech acts. CUP, 1990, vol. 1: Principles of language use.
26. Friggieri J. et al. Actions and Speech Actions in the Philosophy of JL Austin. 1993.
27. Searle J. R., Vanderveken D. Basic concepts of calculus of speech acts. *New in foreign linguistics*. Moscow, 1986, iss. 18: Logical analysis of natural language, p. 242–264.
28. Vanderveken D. Fondements die la logique des attitudes. Manuscrito, 2006, vol. 29, no. 2.

29. Bekhtereva S. O. Natural language analysis based on the theory of speech acts. Bachelor final qualification work. Novosibirsk, 2012. (in Russ.)

30. Meanings, Texts, and Other Exciting Things: a Festschrift to Commemorate the 80<sup>th</sup> Anniversary of Professor Igor Alexandrovich Melchuk. Eds. Yu. Apresyan, I. Boguslavsky, M.-C. L'Homme, L. Iomdin, J. Miliüeviü, A. Polguère, L. Wanner. Moscow, Yazyki slavyanskoy kul'tury, 2012, 704 p. (in Russ.)

*Received 07.08.2018*

*For citation:*

Makhina E. D., Palchunov D. E. Software System for the Determination of Speech Acts in Natural Language Texts. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 4, p. 95–106. (in Russ.) DOI 10.25205/1818-7900-2018-16-4-95-106

УДК 004.934

DOI 10.25205/1818-7900-2018-16-4-107-114

**Ю. И. Молородов<sup>1,2</sup>, Д. А. Ходорченко<sup>2</sup>**

<sup>1</sup> *Институт вычислительных технологий СО РАН  
пр. Академика Лаврентьева, 6, Новосибирск, 630090, Россия*

<sup>2</sup> *Новосибирский государственный университет  
ул. Пирогова, 1, Новосибирск, 630090, Россия*

*yumo@ict.sbras.ru, daria.khodorchenko@gmail.com*

## **РАЗРАБОТКА ИНФОРМАЦИОННОЙ СИСТЕМЫ ПО КЛЕЩЕВОЙ ОПАСНОСТИ НА ОСНОВЕ ОНТОЛОГИИ ПРЕДМЕТНОЙ ОБЛАСТИ \***

Предложен подход к разработке состава и структуры Интернет-ресурса на основе онтологии предметной области. Описана предметная область, связанная с распространением клещей и переносимыми ими заболеваниями, представлена ее семантическая структура, выделены основные разделы информационного наполнения и показаны способы навигации по ресурсу.

*Ключевые слова:* иксодовый клещ, инфекционные агенты, онтологический подход, информационная система.

### **Введение**

Развитие научной деятельности в современном обществе приводит к росту роли компьютерных технологий. Возрос поток информации, что обусловило необходимость использования новых способов ее хранения, представления, формализации, систематизации и автоматической обработки. Появились способы создания баз знаний, позволяющих использовать их для различных практических целей.

Появления технологий Semantic Web<sup>1</sup> привело к появлению систем, способных без участия человека извлекать нужную информацию из текста. Гипертекстовые страницы Semantic Web имеют дополнительную разметку, которая содержит сведения о семантике элементов страницы. Важным компонентом Semantic Web является понятие онтологии, описывающее смысл семантической разметки.

Как правило, под онтологией понимают систему понятий некоторой предметной области, которая представляется как набор сущностей, соединенных различными отношениями [1].

Создание основанного на развитой онтологической структуре интеллектуального научно-го Интернет-ресурса (ИНИР) или портала по определенным тематикам позволит обеспечить эффективный доступ к информации и ее последующую обработку. Онтология как основа ресурса позволяет, помимо структуризации данных, производить первоначальную верификацию данных, исходя из заданных в ней правил [2; 3]. Это помогает улучшить качество вносимой информации и в ряде случаев исключить противоречивые факты.

Социально значимой информацией являются данные об инфекционных заболеваниях, переносимых иксодовыми клещами, вызванных их разнообразием и сложной структурой при-

---

\* Исследования выполнены при частичной поддержке гранта РФФИ № 18-07-01457, Интеграционного проекта СО РАН № АААА-А18-118022190008-8 (№ 0316-2018-0002) и темы госзадания № АААА-А17-117120670141-7 (№ 0316-2018-0009).

<sup>1</sup> Semantic Web Wikipedia, the free encyclopedia. URL: [https://en.wikipedia.org/wiki/Semantic\\_Web](https://en.wikipedia.org/wiki/Semantic_Web)

родных очагов этих заболеваний. Интеграцию сведений в виде информационного ресурса, или информационной системы (ИС), ориентированной на анализ клещевой опасности на основе данных полевых и лабораторных работ, полученных разными группами исследователей, обеспечит использование подхода, основанного на онтологии. Этот ресурс позволит не только осветить проблему заражения трудноизлечимыми заболеваниями, передаваемыми через укус клеща, но и предоставить площадку для сбора и обмена информацией по данной проблеме в научных кругах. Различное территориальное расположение специалистов и отсутствие единой информационной базы являются серьезной помехой в научной и практической деятельности. Становится актуальной организация эффективного доступа не только к публикациям, описывающим методы и подходы к исследованию генетического разнообразия инфекционных агентов, переносимых клещами разного типа, но и к фактическому материалу, связанному с клещами, территорией и средой обитания насекомых. Целью работы является описание подходов к разработке информационного ресурса, основанного на онтологическом подходе, ориентированном на эпидемиологическую опасность, на примере Новосибирской области.

### Проблема клещевой опасности

Ареал обитания клещей очень широк. Встречаются виды, обитающие даже в Арктике и Антарктике. Они являются основными переносчиками множества вирусных, бактериальных и протозойных (паразитических простейших) возбудителей заболеваний человека и животных: клещевого энцефалита, боррелиоза, эрлихиоза и др. Вирус клещевого энцефалита (ВКЭ) и боррелии – наиболее социально опасный среди перечисленных возбудителей природно-очаговых инфекций человека [4]<sup>2</sup>.

На территории России встречается около 60 видов иксодовых клещей, основными переносчиками *инфекционных заболеваний* (клещевой энцефалит и клещевой боррелиоз) являются таежный клещ *Ixodes persulcatus*, обитающий в азиатской части страны, и европейский лесной клещ *Ixodes ricinus*, обитающий в европейской части страны. Недавно обнаружена новая разновидность вируса Кемерово [5].

В зависимости от вида клещи могут содержать различные *инфекционные агенты* – вирусы, бактерии и пр. Часто их называют патогенами. Многие переносимые клещами заболевания имеют сходные ранние симптомы, что затрудняет точную диагностику и лечение<sup>3</sup>. Иногда клещи могут переносить несколько возбудителей заболеваний, а это еще сильнее осложняет постановку диагноза и оказание медицинской помощи. Только полная информация о клещах и переносимых ими инфекциях может помочь врачу правильно диагностировать заболевание и назначить лечение.

Для обеспечения эффективных мер профилактики вирусных заболеваний необходим пространственный и временной анализ распространения клещей, в том числе инфицированных теми или иными возбудителями. Отсутствие эффективной технологии раннего выявления известных и новых патогенов и прогнозирования их распространения является одной из важных и острых проблем. И в этой связи одним из наиболее перспективных способов контроля над возбудителями инфекций может стать система непрерывного наблюдения. Важным шагом к созданию подобной системы является пространственный и временной анализ на основе геоинформационных технологий [6]. Геоинформационная система (ГИС) – система сбора, хранения, анализа и графической визуализации пространственных (географических) данных и связанной с ними информации о необходимых объектах. Понятие ГИС также используется в более узком смысле – как инструмент (программный продукт), позволяющий пользователям искать, анализировать и редактировать как цифровую карту местности, так и дополнительную информацию об объектах.

<sup>2</sup> Иксодовые клещи // МЧС России. URL: [http://www.mchs.gov.ru/dop/info/individual/Bitovie\\_CHS/item/7816051/](http://www.mchs.gov.ru/dop/info/individual/Bitovie_CHS/item/7816051/) (дата обращения 05.05.2018).

<sup>3</sup> Природно-очаговые инфекции / ФБУН Центральный НИИ Эпидемиологии Роспотребнадзора. URL: <https://www.cmd-online.ru/vracham/spravochnik-vracha/prirodno-ochagovye-infektsii/> (дата обращения 05.05.2018).

## Онтология клещевой опасности

Основной задачей создаваемой информационной системы является получение, интеграция и предоставление данных и знаний в интересах научных исследований в предметной области. Концептуальной основой систематизации знаний и информации предметной области являются онтологии. Как правило, они исполняют роль модели предметной области [2; 3]. Онтология является ядром, базовым компонентом информационной модели. С ее помощью можно не только описать систему знаний ИНИР, но и создать формальные структуры для представления его контента. Онтология содержит понятия моделируемой области, связывающие их отношения, атрибуты понятий и отношений, ограничения на значения атрибутов, а также аксиомы, определяющие семантику понятий и отношений. Формализм, используемый в технологии построения порталов научных знаний, обеспечивает компактное и непротиворечивое описание понятий проблемной и предметной областей портала и разнообразных семантических связей между ними, а также выстраивание понятий в иерархию «общее – частное» и поддержку наследования свойств по этой иерархии [7].

Важным моментом при разработке онтологии предметной области является построение таксономий. Таксономия – это предметная (тематическая) классификация, которая группирует термины в виде управляемых словарей (тезаурусов) и упорядочивает их в виде иерархических структур. Разработанная онтология предметной области по клещевой опасности включает четыре базовые иерархические структуры: иерархию объектов, иерархию предметов исследования, иерархию методов исследования и иерархию научных результатов [1]. Свойства каждого понятия описываются с помощью атрибутов и ограничений, налагаемых на область их значений. Понятия базовых онтологий связаны между собой ассоциативными отношениями, выбор которых осуществлялся исходя не только из полноты представления проблемной и предметной областей портала, но и из удобства навигации по его информационному пространству и поиска информации (рис. 1).

Одним из этапов создания ИНИР является изучение предметной области, связанной с клещевой опасностью, и формулирование некоторого представления структуры знаний об этой области. При работе со всем многообразием информации, имеющей отношение к клещам и связанной с ними эпидемиологической опасности, были выделены следующие основные объекты исследования: собственно клещи, местность их распространения (биотоп), инфекционные агенты, переносимые клещами, и гены, которые обнаруживают у клещей и у некоторых инфекционных агентов. Для каждого объекта исследования была составлена таксономия, содержащая элементы, актуальные для предметной области, связанной с клещевой опасностью, на примере Новосибирской области [5; 6].

В результате изучения предметной области были выделены основные понятия: клещ, территория и местность распространения, переносимые клещами патогены и вирусные заболевания. В онтологию были также включены персоны, профессионально связанные с этой областью: первооткрыватели, исследователи, специалисты санитарно-эпидемиологической службы и лица, принимающие решения.

Для каждого вида клещей были определены связанные с ним ключевые понятия. Схема связей между понятиями представлена на рис. 1. Для того чтобы формализовать структуру полученной информации о клещах, была построена иерархическая структура онтологии (рис. 2).



Рис. 1. Схема связей между ключевыми понятиями

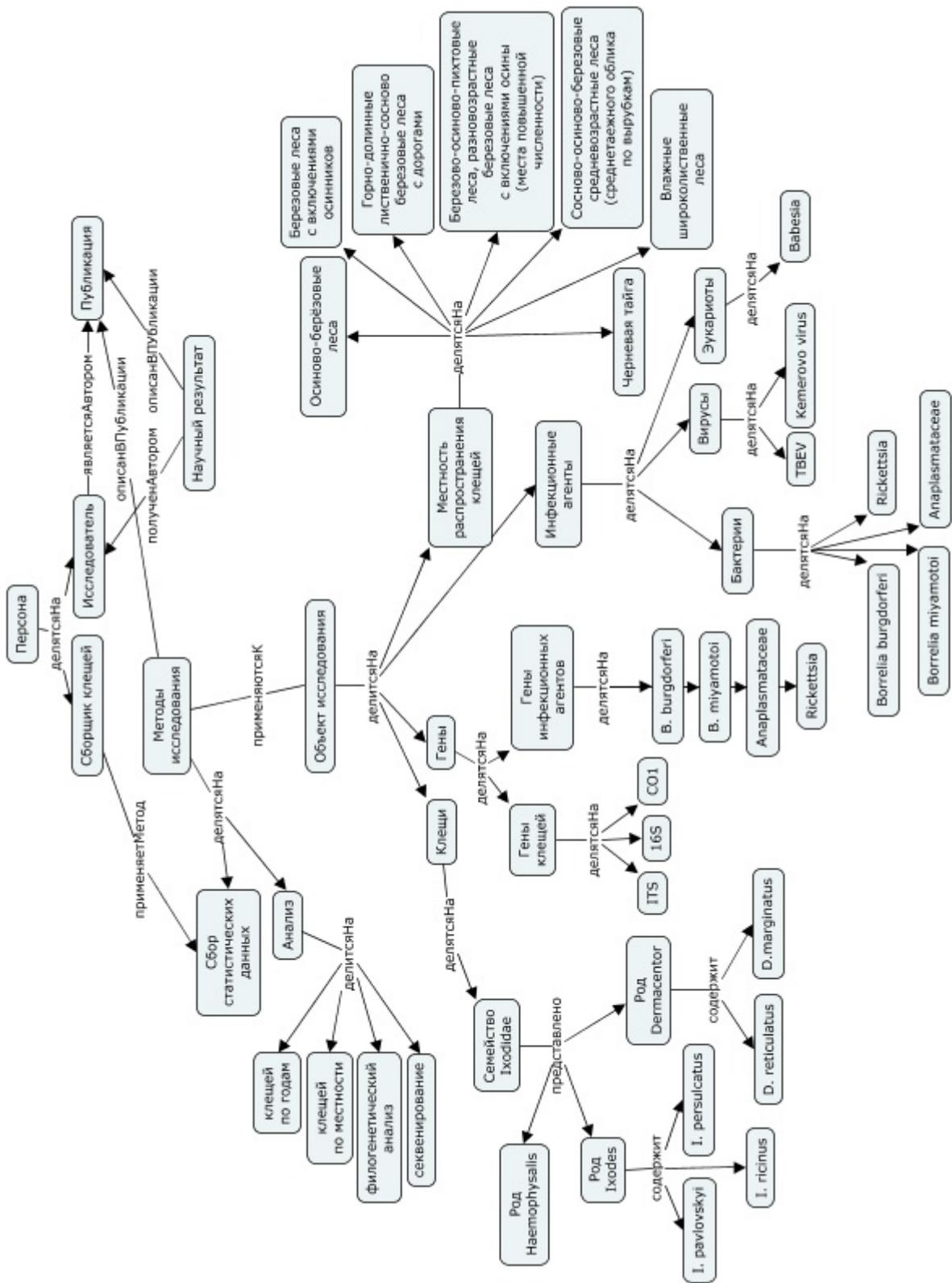


Рис. 2. Онтология предметной области

Для конкретизации информации об исследуемых объектах, были определены методы исследования выбранных объектов. Наиболее широко используемыми методами являются секвенирование биологического материала и анализ генетической информации. Анализ подразделяется на несколько групп: анализ количества клещей по различным территориям, анализ клещей на одной территории по годам и филогенетический анализ генома клещей [6].

Онтология также содержит информацию о персонах, которые собирали исходный фактический материал, а потом использовали указанные методы. Этими персонами могут быть сборщики информации, которые собирали статистические данные о клещах, либо исследователи, которые в результате обработки и анализа данных получали научный результат, который, как правило, формулировался в виде публикаций.

### Архитектура информационной системы

Концептуальным базисом создаваемого интернет-ресурса (портала знаний) является описанная выше онтология. Онтология портала вводит формальные описания понятий предметной области в виде классов объектов и отношений между ними, тем самым задавая структуры для представления реальных объектов и их связей [7]. В соответствии с этим данные на портале представлены в виде семантической сети, т. е. как множество разнотипных взаимосвязанных информационных объектов [3]. Содержательный доступ к систематизированным знаниям и информационным ресурсам обеспечивается с помощью информационной системы, предоставляющей развитые средства навигации и поиска. Архитектуру ИС определяют ее компоненты, их функции и взаимодействие.

Система разработана на основе шаблона проектирования MVC<sup>4</sup>. Моделью является хранилище данных, представлением – пользовательский интерфейс, контроллером – интерпретатор действий пользователя (рис. 3). Использование данного подхода позволяет проводить разработку, модификацию или замену каждого компонента независимо друг от друга.

При работе с ИС можно выделить две важные части – получение данных и изменение данных. В каждой из них можно выделить следующие компоненты: *получение* – навигационная и поисковая системы, *изменение* – редакторы данных и онтологии.

Пользовательский интерфейс реализован с помощью технологии JavaServer Pages<sup>5</sup>. Функциями представления являются навигация по ресурсу, отображение контента и результатов поиска. Информация об объекте онтологии представляется в виде HTML-страницы, где показываются свойства и их значения, а также связи с другими объектами в виде гиперссылки. Отображение данных зависит от прав доступа пользователя. Так, незарегистрированному пользователю доступны только навигация и поиск, а эксперту знаний – весь функционал ресурса. На рис. 4 представлен интерфейс информационного ресурса.

Особое внимание при разработке информационной модели ресурса было уделено связям между экземплярами классов онтологии, поскольку данные, указанные в атрибутах и описаниях, являются основным источником знаний для пользователей. С этой точки зрения важно поддерживать не только навигацию через иерархию наследования, определяемую отношениями is-A, subclassOf и т. п., но и другими видами отношений [8].

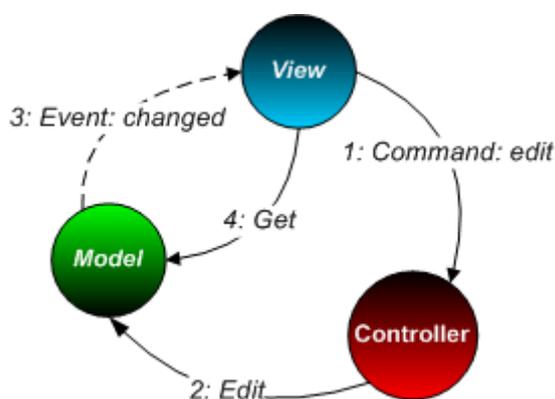


Рис. 3. Шаблон проектирования MVC

<sup>4</sup> Model – view – controller. [Электронный ресурс] Wikipedia, the free encyclopedia. URL: <https://en.wikipedia.org/wiki/Model-view-controller/>.

<sup>5</sup> JavaServer Pages Wikipedia, the free encyclopedia. URL: [https://ru.wikipedia.org/wiki/JavaServer\\_Pages/](https://ru.wikipedia.org/wiki/JavaServer_Pages/).

[Вход](#) | [Редактор онтологии](#) | [Редактор данных](#)

**Свойства объекта**

Ixodes persulcatus	
Название	Таежный клещ

**Связи объекта**

Переносит	
<b>Инфекционные агенты</b>	
<a href="#">Borrelia burgdorferi</a>	
<a href="#">Tick-borne encephalitis virus</a>	
Распространена	
<b>Территория</b>	
<a href="#">Алтайский край</a>	
<a href="#">Амурская область</a>	
<a href="#">Архангельская область</a>	
<a href="#">Вологодская область</a>	
<a href="#">Забайкальский край</a>	
<a href="#">Ивановская область</a>	
<a href="#">Иркутская область</a>	

Рис. 4. Интерфейс информационного ресурса

Онтология как базовый компонент информационной системы должна быть приведена к виду, пригодному для машинной обработки. Базовый компонент был разработан на языке веб-онтологий OWL<sup>6</sup> в редакторе онтологий Protégé<sup>7</sup>. Для единообразного хранения онтологии и данных было принято решение использовать RDF-хранилище. Был выбран Jena Fuseki Server как бесплатное мультиплатформенное средство, поддерживающее запрос данных через SPARQL Query Language, изменение данных через SPARQL Update, а также логический вывод для проверки согласованности базы знаний<sup>8</sup>.

Контроллер связывает между собой модель и представление, интерпретируя действия пользователя для отображения или изменения данных. Контроллер реализован через Java Servlet интерфейс. Так как модель RDF служит только для описания данных, но не их обработки, то необходимо использовать сторонние средства. Был использован язык запросов SPARQL<sup>9</sup>. Функциональные компоненты ресурса были разделены по типу SPARQL-запросов на две группы. Навигационная и поисковая системы только обращаются к базе данных, а редактор данных и онтологии еще и модифицирует её.

## Заключение

Построена концептуальная модель информационной системы. На ее основе определено представление сущностей и отношений (связей между сущностями), и обеспечивается поддержка архитектуры универсальной информационной системы, связанной с конкретной областью научных знаний, ориентированных на эпидемиологическую обстановку конкретного ареала. Концептуальная модель включает в себя основные сущности: методы и объекты исследований, научный результат, инфекционные агенты, иксодовые клещи, персоны и публикации. Важной составляющей концептуальной модели являются публикации и данные,

<sup>6</sup> Semantic Web Wikipedia, the free encyclopedia. URL: [https://en.wikipedia.org/wiki/Semantic\\_Web](https://en.wikipedia.org/wiki/Semantic_Web)

<sup>7</sup> URL: <http://protege.stanford.edu/>.

<sup>8</sup> URL: [http://jena.apache.org/documentation/serving\\_data/](http://jena.apache.org/documentation/serving_data/).

<sup>9</sup> URL: <https://en.wikipedia.org/wiki/SPARQL/>.

включая факты – особый вид документа. В свою очередь, факты понимаются как характеристика сущностей, описываемых в онтологии информационной системы, представляемой как единичное значение данных. По результатам изучения предметной области, ориентированной на эпидемиологическую проблему, вызванную инфицированными иксодовыми клещами, была выполнена структуризация информации. На основе структуризации составлена онтология предметной этой области, которая стала базовым компонентом при построении информационной модели ресурса.

### Список литературы

1. Gruber T. R. Towards Principles for the Design of Ontologies Used for Knowledge Sharing // International Workshop on Formal Ontology, March. Padova, Italy, 1993.
2. Загорулько Ю. А., Загорулько Г. Б., Шестаков В. К., Кононенко И. С. Концепция и архитектура тематического интеллектуального научного интернет-ресурса // Тр. XV Всерос. науч. конф. RCDL'2013. Ярославль: ЯрГУ, 2013. С. 57–62.
3. Загорулько Ю. А. Технология разработки интеллектуальных научных Интернет-ресурсов, ориентированная на экспертов предметной области // Инфраструктура научных информационных ресурсов и систем: Сб. избр. науч. ст. Труды Четвертого Всероссийского симпозиума / Под ред. Е. В. Кудашева, В. А. Серебрякова. М.: ВЦ РАН, 2014. Т. 1. С. 69–86.
4. Колонин Г. В. Мировое распространение иксодовых клещей. М.: Наука, 1978. 70 с.
5. Tkachev S., Panov V., Dobler G., Tikunova N. First detection of Kemerovo virus in *Ixodes pavlovskyi* and *Ixodes persulcatus* ticks collected in Novosibirsk region, Russia // Ticks tick-borne dis. 2014. No. 5. P. 494–496.
6. Molorodov Yu. I., Chernenko V. V., Fedotov A. M. Web-based repository for spatial monitoring viral tick-borne pathogens // Eurasian Journal of Mathematical and Computer Applications. 2016. Vol. 4, № 4. P. 37–43.
7. Загорулько Ю. А., Загорулько Г. Б. Онтологии и их практическое применение в системах, основанных на знаниях // Всероссийская конференция с международным участием «Знания – Онтологии – Теории» (ЗОНТ-2011) / Ин-т математики им. С. Л. Соболева СО РАН. Новосибирск, 2011. Т. 1.
8. Грегер С. Э. Проектирование и реализация онтологии навигационной системы сайта // Объектные системы. 2012. № 1 (6).

Материал поступил в редколлегию 17.06.2018

**Yu. I. Molorodov<sup>1,2</sup>, D. A. Khodorchenko<sup>2</sup>**

<sup>1</sup> Institute of Computational Technologies SB RAS  
6 Academician Lavrentiev Ave., Novosibirsk, 630090, Russian Federation

<sup>2</sup> Novosibirsk State University  
1 Pirogov St., Novosibirsk, 630090, Russian Federation

yumo@ict.sbras.ru, daria.khodorchenko@gmail.com

### DEVELOPMENT OF THE INFORMATION SYSTEM FOR TICK-HAZARD ON THE BASIS OF THE OBJECTIVE ONTOLOGY

The paper proposes an approach to the development of the composition and structure of an Internet resource, based on the domain ontology. The subject area associated with the spread of ticks and the diseases they transmit is described, its semantic structure is presented, the main sections of the content are highlighted, and ways to navigate the resource are shown.

*Keywords:* ixodes tick, pathogens, ontology, information system, portal of knowledge.

## References

1. Gruber T. R. Towards Principles for the Design of Ontologies Used for Knowledge Sharing. *International Workshop on Formal Ontology*. March, Padova, Italy, 1993.
2. Zagorulko Yu. A., Zagorulko G. B., Shestakov V. K., Kononenko I. S. Concept and architecture of thematic intelligent scientific Internet resource. *Proc. of the XV All-Russian Scientific Conference RCDL'2013*. Yaroslavl, YarSU Publ., 2013, p. 57–62. (in Russ.)
3. Zagorulko Yu. A. Technology of development of intellectual scientific Internet resources, focused on experts of the subject domain. *Infrastructure of scientific information resources and systems. Collection of selected scientific articles. Proceedings of the Fourth All-Russian Symposium* Eds. E. V. Kudasheva, V. A. Serebryakov. Moscow, Computing Center of the Russian Academy of Sciences, 2014, vol. 1, p. 69–86. (in Russ.)
4. Kolonin G. V. World distribution of ixodic ticks. Moscow, Science, 1978, 70 p. (in Russ.)
5. Tkachev S., Panov V., Dobler G., Tikunova N. First detection of Kemerovo virus in *Ixodes pavlovskiyi* and *Ixodes persulcatus* ticks collected in Novosibirsk region, Russia. *Ticks tick-borne dis*, 2014, no. 5, p. 494–496.
6. Molorodov Yu. I., Chernenko V. V., Fedotov A. M. Web-based repository for spatial monitoring viral tick-borne pathogens. *Eurasian Journal of Mathematical and Computer Applications*, 2016, vol. 4, № 4, p. 37–43.
7. Zagorulko Yu. A., Zagorulko G. B. Ontologies and their practical application in systems based on knowledge. *All-Russian Conference with international participation "Knowledge-Ontologies-Theories" (ZONT-2011)*. Novosibirsk, S. Sobolev Institute of Mathematics SB RAS, 2011, vol. 1. (in Russ.)
8. Greger S. E. Design and implementation of the ontology of the navigation system of the site. *Object systems*, 2012, № 1 (6). (in Russ.)

Received 17.06.2018

### *Acknowledgements:*

The work is supported by RFBR (grant 18-07-01457), projects AAAA-A18-118022190008-8 (0316-2018-0002) and AAAA-A17-117120670141-7 (0316-2018-0009).

### *For citation:*

Molorodov Yu. I., Khodorchenko D. A. Development of the Information System for Tick-Hazard on the Basis of the Objective Ontology. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 4, p. 107–114. (in Russ.) DOI 10.25205/1818-7900-2018-16-4-107-114

**П. В. Мызников**

*Новосибирский государственный университет  
ул. Пирогова, 1, Новосибирск, 630009, Россия*

*miznikov72@gmail.com*

## **РАЗРАБОТКА ПРЕЦЕДЕНТНО-ОРИЕНТИРОВАННОГО ПОДХОДА ОБРАТНОГО РЕИНЖИНИРИНГА WEB-ИНТЕРФЕЙСОВ**

Статья посвящена применению рассуждений на основе прецедентов (case-based reasoning, CBR) в области web-разработки. Учитывая собственный практический опыт в данной сфере, автор предлагает автоматизировать составление html/css-макета путём агрегирования прецедентов кода из предыдущих опытов, что на методологическом уровне крайне близко к подходу CBR, а именно: предлагается методика конструирования систем, генерирующих html-код из растрового изображения и основанных на подходе рассуждений на основе прецедентов. В частности, в статье изложена оригинальная теория составления структуры изображения и описан алгоритм получения такой структуры. Кроме того, рассматривается модификация подхода case-based reasoning, которая позволяет получить требуемый результат решения поставленной задачи. Также описываются результаты прохождения экспериментов разработанной системы в определённых условиях.

*Ключевые слова:* рассуждения на основе прецедентов, компьютерное зрение, оптическое распознавание символов, обратный инжиниринг, web-разработка.

### **Введение**

Web-технологии являются одним из самых развитых направлений современных информационных технологий. Они используются не только для разработки коммерческих сайтов компаний, но и как один из важных узлов любой ИТ-инфраструктуры. Соответственно можно говорить о том, что практически вся отрасль информационных технологий в том или ином виде использует web-разработку.

Вместе с тем автоматизация самой web-разработки является недостаточно полной. Существуют средства автоматизации развёртывания сред окружения, механизмы конструирования каркасов приложений и некоторые другие направления. Однако почти не охваченным остаётся одна из ключевых задач: создание HTML/CSS-макетов. С одной стороны, эта часть считается одной из наиболее рутинных; с другой – ряд особенностей этого процесса не позволяет применить классические средства автоматизации. К таким особенностям можно отнести: наличие не формализуемых требований к макету, вариативность корпоративных стандартов написания кода, требование кроссбраузерности и кроссплатформенности, что предполагает постоянную адаптацию решения к новым средам.

Автоматизация создания HTML/CSS-макетов не только ускорит процесс разработки web-приложений, но и сделает его более гибким, позволяя масштабировать его результаты, проверять гипотезы, помогать тестированию. Это обуславливает целесообразность описываемой разработки.

Ввиду сказанного объектом исследования является обратный реинжиниринг в области web-технологий, а предметом – автоматизация обратного реинжиниринга интерфейсов web-

приложения. Цель работы – создать технологию построения систем генерации html-кода на основе растрового изображения.

Методическую основу теоретических оснований составляет подход *case-based reasoning* (CBR, рассуждения на основе прецедентов). Кратко описывая его суть, можно сказать, что это способ решения проблем путём адаптации решений аналогичных проблем в прошлом к текущей ситуации. Такой подход выбран в связи с поставленной в работе гипотезой о том, что формализация изображений (чем, по сути, является создание HTML-макета) посредством трансдуктивных выводов, или рассуждений по аналогии, порождает результат, наиболее близкий к тому, что производит человек. К тому же такая схема позволяет решить описанные выше проблемы автоматизации.

Следует отметить, что поставленная выше гипотеза обуславливает использование подхода, который отличает данное исследование от смежных (см. раздел «Обзор литературы»), что можно рассматривать как элемент научной новизны. Кроме того, к новым результатам можно отнести разработку алгоритма извлечения структуры элементов изображения.

### Обзор литературы

В последние годы в зарубежной практике появляется повышенный интерес к поставленной проблеме.

Полученный в работе [1] результат (проект DeepCoder) даёт основания полагать, что современный уровень развития искусственного интеллекта позволяет автоматизировать создание артефактов работы ИТ-специалистов, которые ранее считались неподдающимися автоматизации. Так, DeepCoder – это система генерирования программного кода на основании входных и выходных данных требуемого алгоритма. Хотя такая проблематика отличается от рассматриваемой в текущей работе, разработки, полученные в ходе этого проекта, послужили основой для многих других исследований [2; 3], родственных рассматриваемой задаче. К таким разработкам можно отнести, в частности, использование моделей машинного обучения, основанных на рекуррентных нейронных сетях, и специальные техники поиска блоков кода.

Разработку достаточно близкой задачи осуществлял Т. Нгуен [4]: обратный инжиниринг интерфейсов мобильного приложения. Напрямую перенести этот опыт на инжиниринг web-интерфейсов не получится. Однако сильная сторона исследования – техники оптического распознавания символов (OCR) – может быть повторно использована для распознавания элементов изображения страницы Интернет-сайта.

Самым последним успешным случаем решения описываемой задачи можно назвать работу организации Uizard Technologies [5] под названием «pix2code». Авторы отмечают схожесть данной задачи с задачей генерирования текста на естественном языке, описывающего изображение. Соответственно, и метод они выбирают схожий: в основе их решения лежит каскад рекуррентных нейронных сетей типа LSTM (long-short term memory, долгая краткосрочная память). При несомненных достоинствах этого подхода и успешных практических результатах следует выделить недостатки.

Наиболее значимым видится, как утверждают сами авторы, отсутствие использования экспертной эвристики. Принимая преимущества этой особенности в некоторых условиях, необходимо отметить, что в общем случае это ведёт к ограниченности подхода, так как многие требования к реализации макетов нельзя извлечь исключительно из изображения: необходим контекст, который существует в знаниях экспертов.

Другой важный недостаток – формирование обучающей выборки. Нейронные сети, несомненно, мощный инструмент машинного обучения, однако он требует на вход огромное количество данных (порядка миллионов записей), чтобы быть по-настоящему эффективным. Безусловно, учитывая колоссальные размеры Интернета, такую выборку собрать технически несложно. Но возникает вопрос о репрезентативности данных: далеко не всякий web-сайт положительно скажется на обучении нейронных сетей. В связи с этим появляется дополнительная задача определения качества HTML-кода, что видится существенным увеличением сложности реализации.

В работах [6; 7] рассуждения на основе прецедентов и методы онтологического моделирования применялись для диагностирования заболеваний позвоночника. Знания извлекались и обрабатывались при помощи онтологической модели данной предметной области [8]. В качестве прецедентов рассматривались истории болезни пациентов. Множество прецедентов формально было представлено в виде прецедентной модели, являющейся частным случаем булевозначной модели [9]. Порождение оценочных знаний о болезнях пациентов производилось при помощи нечётких моделей, полученных фазификацией булевозначных прецедентных моделей [9; 10].

Опыт и результаты этих исследований, учёт их сильных и слабых сторон будут использованы в текущей работе.

### Постановка задачи

Опишем подробнее, что требуется от системы. Алгоритм должен распознать структуру изображения, выделить как можно больше характеристик объектов в этой структуре (цвет, шрифт, отступы...) и сгенерировать код, реализующий построение распознанной структуры.

При этом не требуется:

- распознать **все** атрибуты,
- воссоздать картинку пиксель-в-пиксель.

В идеале достаточно, чтобы алгоритм безошибочно сгенерировал около 85 % кода так, чтобы HTML-специалист любой квалификации мог без труда добавить оставшиеся 15 %, при этом чтобы не нужно было принципиально менять структуру кода, а также думать о решении стандартных задач. Поставим этот показатель в качестве ориентира при разработке прототипа.

Во введении и обзоре литературы были упомянуты условия, которые следует учесть при решении задачи. Необходимо разработать ряд требований к системе таких, чтобы можно было получить положительные результаты при работе в этих условиях.

- **Условие 1. Наличие не формализуемых требований к макету.** Изображение само по себе не несёт всю полноту информации о том, каким должен быть конечный HTML-макет. Соответственно, требуется, чтобы в разрабатываемую систему можно было заносить дополнительное знание о требованиях.

- **Условие 2. Вариативность корпоративных стандартов написания кода.** Существует несколько методологий HTML-вёрстки (адаптивная – не адаптивная, bem-методология, bootstrap и т. д.). Более того, как правило, в каждой команде разработчиков существуют внутренние стандарты. Соответственно, от системы требуется умение создавать код в разных стилях и стандартах.

- **Условие 3. Требование кроссбраузерности и кроссплатформенности.** HTML-код должен не только быть валидным, но и одинаково восприниматься всеми современными браузерами и операционными системами.

Кроме того, необходимо ввести ограничения к входным данным на этапе разработки прототипа. Изображение сайта не должно содержать:

- градиентный тип заливки,
- элементы анимации,
- «выпадающие элементы» (элементы, возникающие на странице в результате определённых событий).

Данные ограничения вводятся на этапе создания прототипа в целях лучшего решения базовых проблем. В дальнейшем предполагается создание алгоритма, работающего и без этих условий.

### Архитектура

В основе архитектуры лежит попытка воссоздать процесс составления кода человеком. Базовое утверждение состоит в том, что человек, формализуя визуальную информацию, преимущественно применяет трансдуктивные рассуждения: он извлекает из изображения структуры разного уровня абстракции и описывает эти структуры по аналогии со своим или чужим прошлым опытом.

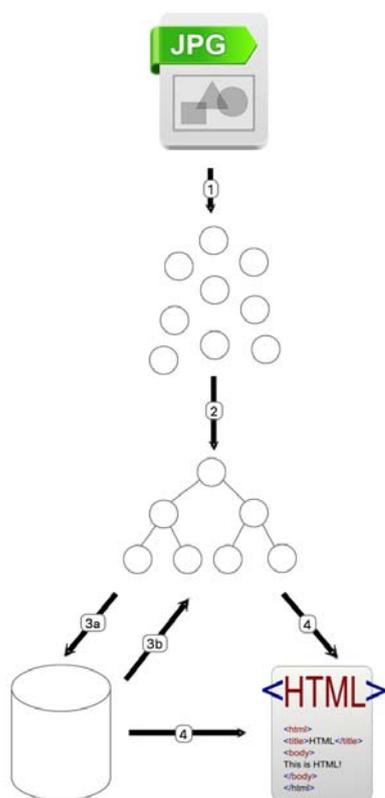


Рис. 1. Схема генерации HTML-кода из изображения

В частности, такая ситуация наблюдается при написании html-кода. Встречая незнакомый паттерн расположения блоков, специалист находит в справочниках и вообще в литературе, как реализовать это на языке html. Далее, с опытом, обращений к справочникам становится всё меньше, но принцип остаётся тем же: сталкиваясь с очередным шаблоном, он достаёт из памяти подходящий пример и адаптирует к текущей ситуации.

Данный процесс реализуется следующим образом (рис. 1):

1) из изображения извлекаются все элементы (текст, картинки, блоки и т. п.);

2) извлечённые элементы собираются в древовидную структуру согласно специальному алгоритму;

3) совершается префиксный обход дерева, на каждой итерации которого происходит обращение к базе прецедентов с целью поиска подходящего описания архитектурного паттерна, лежащего в обрабатываемом узле, на языках HTML и CSS;

4) полученные на предыдущем шаге артефакты сохраняются в виде требуемых файлов.

Шаги 1 и 4 являются инженерными задачами, которые уже имеют решения в общедоступной практике. Шаги 2 и 3 – наукоёмкая часть архитектуры. В разделах «Извлечение структуры изображения» и «Построение СВР-системы» приводится разработка теории и алгоритмов, которые были использованы для реализации этих шагов.

## Извлечение структуры изображения

Цель данного этапа – получить так называемую **структуру** растрового изображения, т. е. математический объект заданного формата, описывающий взаимное расположение частей изображения. Поставленная задача является задачей построения дерева. В данной статье предлагается жадный тип алгоритма решения.

На вход алгоритму подаётся множество узлов, каждый из которых содержит ровно один элемент изображения, упорядоченное по координатам левой верхней точки описанных прямоугольников.

Общая идея заключается в последовательной обработке элементов данного множества и попытке на каждой итерации объединить обрабатываемый элемент со следующим в определённый узел (рис. 2). При объединении элементов в узлы мощность множества будет сокращаться. Алгоритм завершит работу, когда во множестве останется ровно один элемент – именно этот элемент, представляющий собой дерево всех элементов, является решением задачи.

Проведём оценку сложности алгоритма по времени и по памяти.

Заметим, что алгоритм не использует дополнительную память, т. е. не зависит от количества элементов в множестве. Соответственно, оценка по памяти равна  $O(1)$ .

С оценкой по времени ситуация сложнее. В зависимости от расположения элементов, алгоритм может обладать совершенно разной сложностью.

В лучшем случае изображение состоит из ряда элементов, расположенных в линию, горизонтальную или вертикальную. Тогда алгоритму достаточно пройти один раз по множеству и объединить их все в один определённый узел. Оценка сложности по времени в этом случае

линейная –  $O(n)$ . К сожалению, такой случай почти не встречается на практике (только, быть может, в очень узких задачах).

Чтобы оценить худший случай, проанализируем, как собирается дерево. При каждом проходе элементов (см. рис. 2, цикл на строке 2) собираются узлы, которые в конечном дереве будут лежать на одном уровне. При этом каждый такой проход в худшем случае оценивается как  $O(n)$ . Соответственно, оценка сложности всего алгоритма равна  $O(n * h)$ , где  $h$  – высота дерева. Наихудшая высота дерева равна количеству элементов: на каждом проходе цикла получается максимум один узел. Получаем  $O(n^2)$  как оценку времени выполнения алгоритма в худшем случае. Хорошая новость состоит в том, что таких примеров на практике хоть и больше, чем наилучших случаев, но всё же не так много.

**Require:** *nodes* – множество элементов, упорядоченных по абсциссе и ординате

```

1: while |nodes| > 1 do
2:   for orientation ← [HOR,VERT] do
3:     i ← 0
4:     while i < |nodes| do
5:       suitableNode ← findSuitableNode(nodesi, orientation)
6:       if suitableNode is not null then
7:         if nodei is composite then
8:           nodesi.addChild(suitableNode)
9:         else
10:          newNode ← Node()
11:          newNode.addChild(nodesi)
12:          newNode.addChild(suitableNode)
13:          newNode.orientation = orientation
14:          nodesi ← newNode
15:        delete suitableNode
16:        i ← i + 1

```

**Ensure:** |nodes| == 1 и nodes<sub>1</sub> – корневой узел дерева, содержащий все элементы из исходного множества

Рис. 2. Алгоритм построения иерархии элементов изображения

В среднем ожидается, что на каждой итерации цикла (см. рис. 2, строка 2) количество элементов будет логарифмически сокращаться, поэтому среднюю сложность можно оценить как  $O(n \log n)$ . Для более точных данных об истинном распределении наилучших, наихудших и прочих случаев, а также справедливости приведённых в данном параграфе формул необходимо провести дополнительные эксперименты и исследования.

### Рассуждения на основе прецедентов

Рассуждения на основе прецедентов (CBR, case based reasoning) – одно из успешных направлений развития искусственного интеллекта. Будучи изобретённым в конце 70-х гг. XX века, case-based reasoning использовался в разработке многих интеллектуальных систем.

Работы Шэнка и Абельсона 1977 г. [11] считаются истоками рассуждений на основе прецедентов. Роджер Шенк предложил следующую концепцию: наше знание о мире организовано в виде пакетов памяти, хранящих эпизоды жизни. Такие пакеты (MOPs – memory organizations packets) и их элементы не изолированы, а пересекаются с нашими ожиданиями развития событий (сценариями). В свою очередь, MOPs образуют иерархию, где более общие

пакеты объединяют более специфичные. Если МОР содержит ситуацию, в которой некоторая проблема была успешно решена, и человек находит себя в подобной ситуации, то он стремится использовать предыдущий опыт, чтобы найти решение. Таким образом, вместо следования общему набору правил происходит повторное применение схемы решения в новых, но схожих условиях.

В 1990-х гг. было написано большое количество научных работ, посвящённых этому направлению [12–17]. Тем не менее в 2000-х гг. популярность CBR стала падать вследствие вытеснения другими подходами. Несмотря на это, рассуждения на основе прецедентов остаются достаточно эффективной методологией, которую до сих пор используют в самых разных предметных областях.

Кратко описывая суть рассуждений на основе прецедентов, можно сказать, что это способ решения проблем путём адаптации решения схожих проблем в прошлом к текущей ситуации. Процесс нахождения такого решения называется CBR-циклом и состоит из 4 фаз [14].

5) **Извлечение.** Из хранилища прецедентов извлекается прецедент с проблемой, наиболее похожей на ту проблему, которую нужно решить.

6) **Адаптация.** Решение извлечённого прецедента адаптируется так, чтобы решить исходную проблему.

7) **Оценка.** Полученное решение оценивается: если оно не удовлетворяет заданным критериям, то либо извлекаются дополнительные прецеденты и снова происходит попытка адаптации, либо процедура завершается с неудачным результатом.

8) **Сохранение.** Прецедент с полученным решением сохраняется в хранилище прецедентов.

Более подробно остановимся на представлении прецедентов. В CBR прецедент – это тройка элементов:

- проблема – состояние мира / предметной области, при котором возникает прецедент,
- решение – приемлемый в данном контексте ответ для проблемы,
- результат – состояние мира / предметной области после применения решения.

### Построение CBR-системы генерации кода

Как было сказано в предыдущем разделе, чтобы построить CBR-систему, необходимо определить следующие вещи: формат прецедента, меру сходства, процедуру адаптации, словарь, а также реализацию 4 шагов CBR-цикла.

В разделе «Архитектура» сказано, что генерация кода производится путём префиксного обхода дерева – структуры изображения. На каждом этапе этого обхода запускается CBR-цикл. Требуется определить, что является проблемой, решением и результатом в контексте решаемой задачи. Иными словами, описать структуру прецедента.

- **Проблема:** набор свойств обрабатываемого узла и свойств его прямых потомков.
- **Решение:** шаблоны html- и css-кода, реализующие вёрстку описываемой структуры.
- **Результат:** html/css-код, сгенерированный из решения-шаблона применительно к конкретному случаю.

Так, при обработке каждого узла дерева формируется <<проблема>>: описание текущего элемента структуры изображения вместе с описанием структуры изображения на один уровень детализации глубже. В хранилище лежат прецеденты, проблемы которых представляют собой типовые случаи расположения элементов, а решения – типовые варианты их вёрстки. Задача CBR-цикла – используя хранимые решения, получить html/css-код для реализации текущей структуры. Опишем этот цикл подробнее. Для условных обозначений назовём обрабатываемый прецедент общепринятым термином *новый прецедент*. Дальнейшие термины будем вводить по мере их появления.

**Извлечение.** *Новый прецедент* содержит проблему, но не имеет решения и результата. Задача данного этапа – извлечь из хранилища прецеденты, максимально близкие по проблеме к *новому прецеденту*. В общем случае имеем задачу  $n$ -классовой классификации, где  $n$  – количество прецедентов в хранилище. На этапе построения концепции выберем в качестве решения метод  $k$  ближайших соседей с евклидовой мерой расстояния. Для этого векторизуем и нормализуем проблемы. Это означает, что, во-первых, все категориальные признаки при-

ведём к числовому виду; во-вторых, заменим абсолютные значения относительными по формуле  $x_i = \frac{x_i}{x_{\max} - x_{\min}}$ . Далее, вычисляем расстояния между текущим вектором и векторами

прецедентов в хранилище  $d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$  и выбираем прецедент, расстояние до которого оказалось минимальным. В терминах CBR такой прецедент называется *извлечённым*.

**Адаптация.** Шаг адаптации реализуется алгоритмом, принимающим на вход проблему *нового* прецедента и решение *извлечённого*, а отдающим на выход результат. Прецедент, составленный из проблемы *нового* прецедента, решения *извлечённого* прецедента и результата, полученного на шаге адаптации, называется *решённым*, а само решение называется *предлагаемым*.

В текущем исследовании предлагается деривационный тип адаптации, т. е. регенерация старого решения в новых условиях. Так как решение прецедента – это шаблон кода, то результат получается путём применения выбранного языка шаблонизации. Трудность заключается в том, что результаты разных прецедентов в ходе обработки структуры могут конфликтовать друг с другом. Особенно это касается конфликтов в css-коде: ситуации, когда различные результаты описывают по-разному один и тот же css-класс. Поэтому на этапе адаптации важным является процесс разрешения конфликтов.

**Оценка.** В классическом подходе CBR этап оценки служит валидацией полученного результата. Прецедент, который положительно прошёл данный этап, называется *протестированным*, а его решение – *подтверждённым*. Что касается текущей работы, задача оценки качества html/css-кода заслуживает отдельного исследования. На этапе проверки концепции данный шаг опускается, а проверка полученного решения осуществляется для всего документа целиком, а не на каждой итерации CBR-цикла, как предлагается в методологии.

**Сохранение.** Последний шаг цикла призван сохранить подтверждённое решение в базе, для того чтобы использовать его в будущем для решения схожих проблем. В текущей работе этот шаг несколько отличается от классического подхода. В отличие от использования одного хранилища прецедентов разрабатываемая CBR-система поддерживает «глобальное» хранилище и несколько «локальных» хранилищ (см. таблицу). Первый тип – это база прецедентов, используемая по умолчанию для всех изображений. Второй – база прецедентов, используемая только в рамках обработки одного документа. Можно сказать, что «локальное» хранилище – это контекст документа: оно содержит результаты решённых проблем в ходе его обработки.

#### Отличия локального и глобального хранилищ

Локальное хранилище	Глобальное хранилище
Хранится в оперативной памяти	Хранится во внешней памяти
Уничтожается после обработки документа	Существует независимо от обработки документа
Заполняется в процессе обработки документа	Заполняется специальной процедурой до обработки документа
Прецеденты имеют приоритет при извлечении	

Смысл состоит в том, что на этапе сохранения решённый прецедент всегда записывается только в локальное хранилище. При этом на этапе извлечения поиск происходит во всех базах, но преимущество есть у прецедентов в локальном хранилище.

Другими словами, глобальное хранилище служит источником прецедентов, имеющих решения генерации кода типичных примеров расположения элементов. Далее локальное хранилище используется для точной и быстрой адаптации этих решений в контексте конкретного документа.

## Результаты и дальнейшая работа

Компьютерная система, реализующая описанный выше подход, была протестирована на множестве изображений существующих сайтов. Такой вид тестирования позволил оценить, насколько близко к человеческому решению получает результат построенный алгоритм. В качестве примера рассмотрим одну из страниц сайта Новосибирского Государственного Университета.

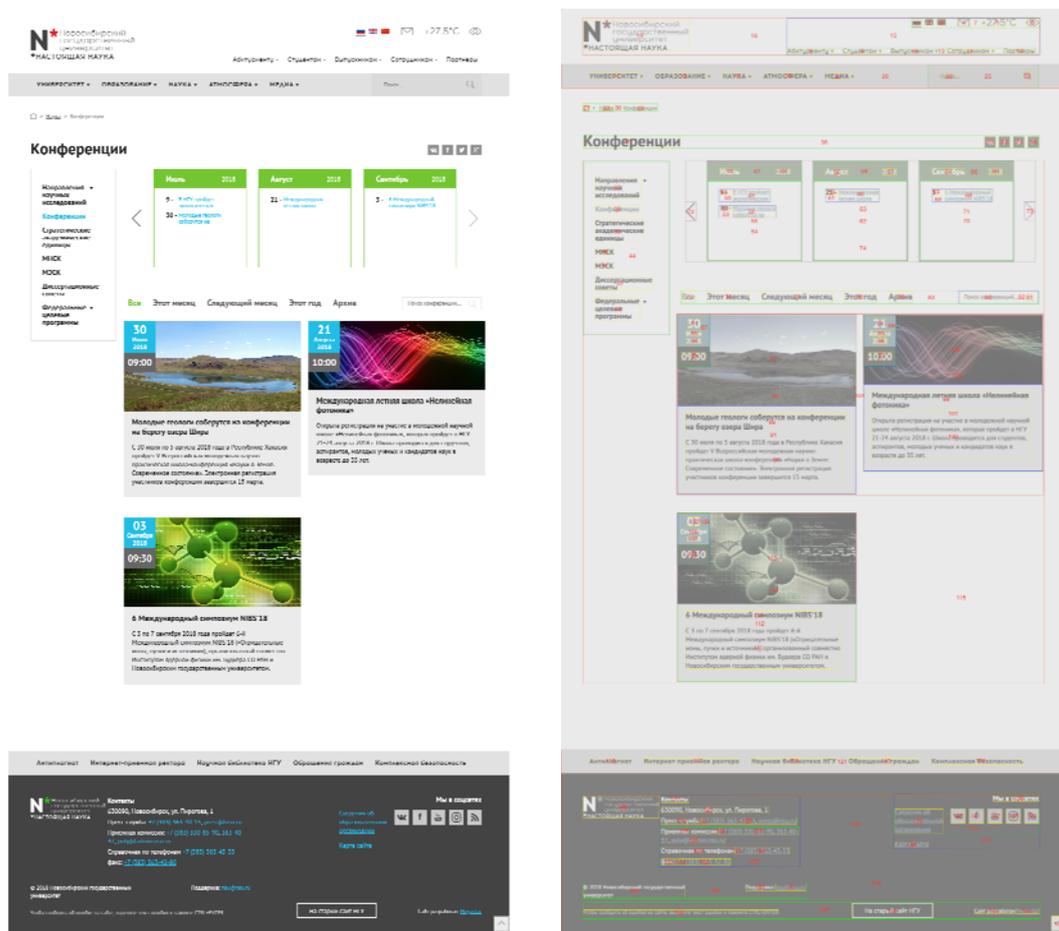


Рис. 3. Оригинал обрабатываемого изображения (слева) и выделенные элементы и узлы (справа). Узлы обведены цветными прямоугольниками и пронумерованы

В результате выполнения первого этапа обработки изображения были выделены элементы и сгруппированы в узлы (рис. 3). Далее, описанный выше алгоритм сформировал из полученных узлов древовидную структуру (рис. 4). Полученный результат имеет достаточно хорошее качество (рис. 5). Несмотря на некоторые ошибки (не был выбран подходящий шрифт, не подгружены некоторые изображения, небольшие ошибки в размерах), система справилась с самой главной задачей: воссоздала структуру элементов и сгенерировала легко поддерживаемый человеком код.

Из данного примера видно, что прототип системы удовлетворяет поставленным требованиям. В общем, учитывая эксперименты на других изображениях, можно утверждать, что средняя точность работы системы составляет 87 %. Также стоит заметить, что увеличение структурной сложности (добавление колонок, «резиновой» ширины) ожидаемо снижает результаты. Эти выводы полезны для определения следующих направлений развития.

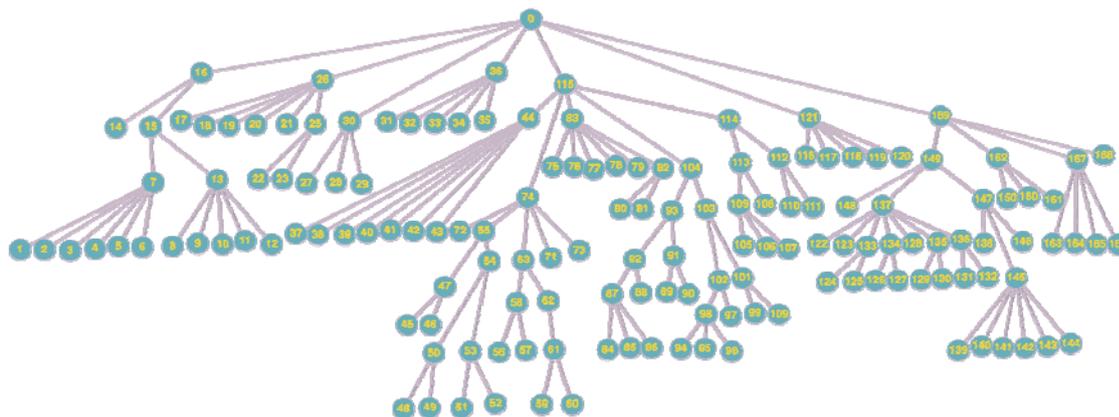


Рис. 4. Структура элементов обрабатываемого изображения. Числа внутри узлов соответствуют индексам меток, расставленных на рис. 3

Помимо количественных требований, в постановке задачи были выдвинуты три условия, в которых система должна эффективно функционировать: 1) наличие не формализуемых требований к макету, 2) вариативность корпоративных стандартов написания кода, 3) требование к кроссбраузерности и кроссплатформенности. Выполнение всех этих условий учитывается в системе за счёт вариативности заполнения хранилища прецедентов. Действительно, помимо автоматического получения множества прецедентов, есть возможность создавать их с помощью экспертов. В этом случае в решениях-шаблонах будет находиться экспертный опыт и дополнительное знание. А производя замену хранилищ, каждое из которых специфицировано для разных стилей и стандартов, можно получать заданный результат в разном исполнении.

Следующими шагами в развитии разработанной концепции являются:

- 1) разработка процедуры автоматического заполнения хранилища прецедентов;
- 2) сбор тренировочного набора реальных примеров изображений и проведение экспериментов на этом наборе;
- 3) разработка наилучшей меры близости прецедентов;
- 4) разработка этапа оценки в CBR-цикле.

Таким образом, получен прототип интеллектуальной системы, эффективность которой доказана экспериментально. Предлагаемый подход является перспективным решением многих практических проблем в web-разработке. Важно и то, что разработанная теоретическая модель распознавания структур изображений может быть обобщена и использована в других вопросах компьютерного зрения.

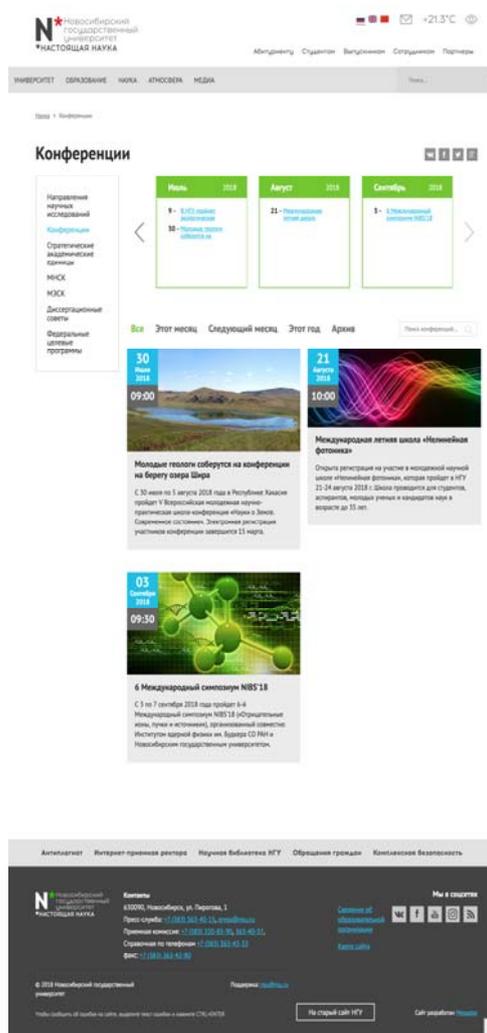


Рис. 5. Изображение web-страницы, сгенерированной системой

**Список литературы**

1. DeepCoder: Learning to Write Programs / M. Balog [et al.]. 2016. Nov. arXiv: 1611.01989. URL: <http://arxiv.org/abs/1611.01989>.
2. TerpreT: A Probabilistic Programming Language for Program Induction / A. L. Gaunt [et al.]. 2016. Aug. arXiv: 1608.04428. URL: <http://arxiv.org/abs/1608.04428>.
3. Latent Predictor Networks for Code Generation / W. Ling [et al.]. 2016. Mar. arXiv: 1603.06744. URL: <http://arxiv.org/abs/1603.06744>.
4. *Nguyen T. A., Csallner C.* Reverse Engineering Mobile Application User Interfaces with REMAUI (T) // 2015 30<sup>th</sup> IEEE/ACM International Conference on Automated Software Engineering (ASE). Nov. 2015. P. 248–259. DOI 10.1109/ASE.2015.32.
5. *Beltramelli T.* pix2code: Generating Code from a Graphical User Interface Screenshot. 2017. arXiv: 1705.07962. URL: <http://arxiv.org/abs/1705.07962>.
6. *Palchunov D., Yakhyaeva G., Yasinskaya O.* Software system for the diagnosis of the spine diseases using case-based reasoning // Proceedings of the International Conference on Biomedical Engineering and Computational Technologies (SIBIRCON / SibMedInfo – 2015). 28–30 October 2015. P. 205–210.
7. *Пальчунов Д. Е., Яхьяева Г. Э., Ясинская О. В.* Программная система, осуществляющая case-based reasoning для диагностирования заболеваний позвоночника // Сибирский научный медицинский журнал. 2016. Т. 36, № 1. С. 97–104.
8. *Пальчунов Д. Е., Яхьяева Г. Э., Ясинская О. В.* Применение теоретико-модельных методов и онтологического моделирования для автоматизации диагностирования заболеваний // Вестник НГУ. Серия: Информационные технологии. 2015. Т. 13, вып. 3. С. 42–51.
9. *Пальчунов Д. Е., Яхьяева Г. Э.* Нечеткие алгебраические системы // Вестник НГУ. Серия: Математика, механика, информатика. 2010. Т. 10, № 3. С. 75–93.
10. *Пальчунов Д. Е., Яхьяева Г. Э.* Нечёткие логики и теория нечётких моделей // Алгебра и логика. 2015. Т. 54, № 1. С. 109–118.
11. *Schank R., Abelson R.* Scripts, Plans, Goals and Understanding. New Jersey, 1977. 248 p.
12. *Slade S., Henry P.* Case-Based Reasoning: A Research Paradigm // AI Magazine. 1991. Vol. 12, no. 1. P. 41–55. ISSN 0738-4602. DOI: <http://dx.doi.org/10.1609/aimag.v12i1.883>.
13. *Golding A. R., Rey M., Rosenbloom P. S.* Improving Rule-Based Systems through Case-Based Reasoning // Knowledge Creation Diffusion Utilization. 1991. April 1998.
14. *Kolodner J. L.* An introduction to case-based reasoning // Artificial Intelligence Review. 1992. Vol. 6, no. 1. P. 3–34. ISSN 0269-2821. DOI 10.1007/BF00155578.
15. *Watson I. A. N., Marir F.* Case-based reasoning: A review // The Knowledge Engineering Review. 1994. Vol. 9, no. 4. P. 327–354.
16. *Aamodt A., Plaza E.* CBR: foundational issues, methodological variations and system approaches // AI Communications. 1994. Vol. 7, no. 1. P. 39–59.
17. *Ram A., Santamaria J.* Continuous case-based reasoning // Artificial Intelligence. 1997. Vol. 90, no. 1–2. P. 25–77. ISSN 00043702. DOI 10.1016/S00043702(96)000379.

**P. V. Myznikov**

*Novosibirsk State University  
1 Pirogov St., Novosibirsk, 630090, Russian Federation*

*miznikov72@gmail.com*

## **DEVELOPMENT OF THE CASE-BASED APPROACH OF WEB INTERFACES REVERSE REENGINEERING**

The article is devoted to implementation of case-based reasoning in web development. Given personal practical experience in this field the author proposes to automate html/css markup building with aggregation of code samples from previous cases, which is very similar to case-based reasoning approach on a methodological level. Namely, the article describes a technique of construction of systems generating html code from a bitmap based on CBR. In particular, an original theory of building an image structure is presented in the article. Besides, the article considers modification of case-based reasoning approach that allows to get the result required. Also the results of practical experiments are described.

*Keywords:* case based reasoning, computer vision, optical character recognition, reverse engineering, web development.

### **References**

1. DeepCoder: Learning to Write Programs. M. Balog [et al.]. 2016, Nov. arXiv: 1611.01989. URL: <http://arxiv.org/abs/1611.01989>.
2. TerpreT: A Probabilistic Programming Language for Program Induction. A. L. Gaunt [et al.]. 2016, Aug. arXiv: 1608.04428. URL: <http://arxiv.org/abs/1608.04428>.
3. Latent Predictor Networks for Code Generation. W. Ling [et al.]. 2016, Mar. arXiv: 1603.06744. URL: <http://arxiv.org/abs/1603.06744>.
4. Nguyen T. A., Csallner C. Reverse Engineering Mobile Application User Interfaces with REMAUI (T). *2015 30<sup>th</sup> IEEE/ACM International Conference on Automated Software Engineering (ASE)*, Nov. 2015, p. 248–259. DOI 10.1109/ASE.2015.32.
5. Beltramelli T. pix2code: Generating Code from a Graphical User Interface Screenshot. 2017. arXiv: 1705.07962. URL: <http://arxiv.org/abs/1705.07962>.
6. Palchunov D., Yakhyaeva G., Yasinskaya O. Software system for the diagnosis of the spine diseases using case-based reasoning. *Proceedings of the International Conference on Biomedical Engineering and Computational Technologies (SIBIRCON / SibMedInfo-2015)*. 28–30 October 2015, p. 205–210.
7. Palchunov D. E., Yakhyaeva G. E., Yasinskaya O. V. Software system for diagnosing spinal diseases using case-based reasoning. *Siberian Scientific Medical Journal*, 2016, vol. 36, issue 1, p. 97–104. (in Russ.)
8. Palchunov D. E., Yakhyaeva G. E., Yasinskaya O. V. Application of model-theoretic methods and ontological modeling to automate the diagnosis of diseases. *Vestnik NSU. Series: Information Technologies*, 2015, vol. 13, no. 3, p. 42–51. (in Russ.)
9. Palchunov D. E., Yakhyaeva G. E. Fuzzy algebraic systems. *Vestnik NSU. Series: Mathematics, Mechanics, Informatics*, 2010, vol. 10, no. 3, p. 75–93. (in Russ.)
10. Palchunov D. E., Yakhyaeva G. E. Fuzzy logics and fuzzy model theory. *Algebra and Logic*, 2015, vol. 54, no. 1, p. 74–80. (in Russ.)
11. Schank R., Abelson R. *Scripts, Plans, Goals and Understanding*. New Jersey, 1977. 248 p.
12. Slade S., Henry P. Case-Based Reasoning: A Research Paradigm. *AI Magazine*, 1991, vol. 12, no. 1, p. 41–55. ISSN 0738-4602. DOI: <http://dx.doi.org/10.1609/aimag.v12i1.883>.
13. Golding A. R., Rey M., Rosenbloom P. S. Improving Rule-Based Systems through Case-Based Reasoning. *Knowledge Creation Diffusion Utilization*, 1991, April 1998.

14. Kolodner J. L. An introduction to case-based reasoning. *Artificial Intelligence Review*, 1992, vol. 6, no. 1, p. 3–34. ISSN 0269-2821. DOI 10.1007/BF00155578.
15. Watson I. A. N., Marir F. Case-based reasoning: A review. *The Knowledge Engineering Review*, 1994, vol. 9, no. 4, p. 327–354.
16. Aamodt A., Plaza E. CBR: foundational issues, methodological variations and system approaches. *AI Communications*, 1994, vol. 7, no. 1, p. 39–59.
17. Ram A., Santamaria J. Continuous case-based reasoning. *Artificial Intelligence*, 1997, vol. 90, no. 1–2, p. 25–77. ISSN 00043702. DOI 10.1016/S00043702(96)000379.

*Received 10.08.2018*

*For citation:*

Myznikov P. V. Development of the Case-Based Approach of Web Interfaces Reverse Reengineering. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 4, p. 115–126. (in Russ.) DOI 10.25205/1818-7900-2018-16-4-115-126

УДК 519.6, 550.832  
DOI 10.25205/1818-7900-2018-16-4-127-139

**О. В. Нечаев<sup>1</sup>, В. Н. Глинских<sup>1-3</sup>**

<sup>1</sup> *Институт нефтегазовой геологии и геофизики им. А. А. Трофимука СО РАН  
пр. Академика Коптюга, 3, Новосибирск, 630090, Россия*

<sup>2</sup> *Новосибирский государственный университет  
ул. Пирогова, 1, Новосибирск, 630090, Россия*

<sup>3</sup> *Новосибирский государственный технический университет  
пр. К. Маркса, 20, Новосибирск, 630090, Россия*

*oleg.v.nechaev@gmail.com, GlinskikhVN@ipgg.sbras.ru*

## **ТРЕХМЕРНОЕ МОДЕЛИРОВАНИЕ И ИНВЕРСИЯ ДАННЫХ КОМПЛЕКСА МЕТОДОВ ЭЛЕКТРОКАРОТАЖА В МОДЕЛЯХ СРЕД С НАКЛОНОМ ГЛАВНЫХ ОСЕЙ ТЕНЗОРА ЭЛЕКТРИЧЕСКОЙ АНИЗОТРОПИИ \***

Впервые формулируются прямая и обратная задачи электрического каротажа об определении компонент тензора электрической анизотропии горных пород по комплексу измерений методами бокового каротажного зондирования и бокового каротажа в нефтегазовых скважинах. Рассматриваются особенности конечно элементной аппроксимации прямой задачи, использующей базисные функции большого порядка. Исследуется возможность восстановления горизонтальной и вертикальной составляющих удельного электрического сопротивления, а также угла наклона главных осей тензора электрической анизотропии.

*Ключевые слова:* метод конечных элементов, многоуровневый метод, прямая и обратная задачи, электрический каротаж, тензор удельного электрического сопротивления.

### **Введение**

Начиная от поиска перспективных объектов, определения их запасов и заканчивая контролем за разработкой месторождений активно используются современные геофизические методы, позволяющие получить исчерпывающую информацию о нефтяных и газовых резервуарах. При изучении геологической среды большую роль играет реконструкция удельного электрического сопротивления (УЭС) горных пород по данным электрокаротажных зондирований [1–4]. Оценка содержания углеводородов в резервуарах выполняется по значениям УЭС горных пород на основе измерений в скважине методами электрического каротажа.

Существуют геологические среды, интерпретация данных измерений в которых требует специализированного математического описания и соответствующих вычислительных алгоритмов. К таким относятся слоистые осадочные породы, представленные переслаиванием тонких прослоев разного вещественного состава и электрофизических свойств, например сильно электропроводящих глинистых отложений и слабо электропроводящих нефтесодержащих песчаников. Такая среда является электрически анизотропной, поскольку значения

---

\* Исследование выполнено при финансовой поддержке РФФИ и Правительства Новосибирской области в рамках научного проекта № 17-45-540530.

*Нечаев О. В., Глинских В. Н.* Трехмерное моделирование и инверсия данных комплекса методов электрокаротажа в моделях сред с наклоном главных осей тензора электрической анизотропии // Вестн. НГУ. Серия: Информационные технологии. 2018. Т. 16, № 4. С. 127–139.

УЭС в плоскости напластования и в вертикальном направлении существенно отличаются. Если не учитывать электрическую анизотропию при интерпретации данных электрического каротажа, это будет приводить к существенным ошибкам при определении оценки нефтесодержания. Эффект электрической анизотропии слабо изучен при исследовании трещиноватых карбонатных коллекторов, в которых наличие электрической анизотропии обусловлено системой субвертикальных трещин. Таким типом анизотропии обладают палеозойские сложнопостроенные коллекторы. Поэтому решение задачи об определении электрической анизотропии горных пород по данным электрического каротажа является чрезвычайно актуальной и имеет большое прикладное значение в промышленной геофизике. Требуется создание новых эффективных алгоритмов и реализация программных средств, адаптированных под традиционные комплексы методов каротажа скважин.

Начиная с 1990-х гг. задачу определения параметров макроанизотропного коллектора решают, применяя индукционные зонды, где генерация токов и измерения выполняется в трехкомпонентных ортогональных катушках [5–11]. Необходимость использования таких сложных систем, а также применения соответствующих ресурсоемких вычислительных схем и процедур интерпретации заметно снижает эффективность их применения. В настоящее время ведутся разработки принципиально новых каротажных систем, предназначенных для определения характеристик макроанизотропных коллекторов [12–15]. В представленной работе изучение электрической анизотропии геологических объектов базируется на интерпретации данных широко используемых методов электрического каротажа, исходно не предназначенных для изучения макроанизотропных сред.

В рамках настоящей работы изучение эффекта электрической анизотропии выполняется по комплексу электрических каротажных методов – бокового каротажного зондирования (БКЗ) и бокового каротажа (БК). Необходимость их комплексирования обусловлена различной чувствительностью к компонентам электрической анизотропии. Особенности этих методов таковы, что сигналы зондов БКЗ имеют достаточную чувствительность к вертикальной компоненте анизотропии, а БК – к радиальной. Таким образом, по комплексу данных БКЗ и БК становится возможным определение элементов диагонального тензора электрической анизотропии, а также оценки наклона его главных осей.

### Прямая задача

Геофизическую модель рассматриваемой среды будем описывать при помощи функции  $\rho(w, x)$ , задающей зависимость УЭС от пространственных координат  $x$  и вектора параметров  $w$ . Последний будет содержать значения параметров геоэлектрической модели. В данной работе будут рассматриваться простые модели, состоящие из однородного пласта, скважины и корпуса зонда. Диаметр и УЭС скважины, а также геометрические характеристики корпуса каротажного зонда будем полагать фиксированными.

В дальнейших обозначениях как один, так и оба параметра функции УЭС могут опускаться. УЭС будет полагаться анизотропным – имеющим горизонтальную и вертикальную составляющие. Тензор УЭС будет характеризовать угол наклона локальной системы координат относительно оси скважины, в которой он является диагональным. Необходимо отметить, что наклон локальной системы координат описывается тремя углами. Но в рассматриваемом случае каротажные зонды являются симметричными относительно вертикальной оси координат, а также диагональный тензор имеет только две независимые компоненты – горизонтальную ( $x$  и  $y$  компоненты) и вертикальную ( $z$  компонента) составляющие. Таким образом, тензор УЭС будет иметь только три независимых параметра и выглядеть следующим образом:

$$\rho = \begin{bmatrix} \rho_h & 0 & 0 \\ 0 & \rho_h \cos(\alpha) \cos(\alpha) + \rho_v \sin(\alpha) \sin(\alpha) & \rho_v \cos(\alpha) \sin(\alpha) - \rho_h \sin(\alpha) \cos(\alpha) \\ 0 & \rho_v \cos(\alpha) \sin(\alpha) - \rho_h \sin(\alpha) \cos(\alpha) & \rho_v \cos(\alpha) \cos(\alpha) + \rho_h \sin(\alpha) \sin(\alpha) \end{bmatrix},$$

где  $\rho_h$  и  $\rho_v$  – горизонтальная и вертикальная компоненты тензора УЭС соответственно,  $\alpha$  – угол наклона оси зонда относительно локальной системы координат, в которой тензор УЭС является диагональным.

В настоящей работе рассматриваются повсеместно используемые на практике при изучении геологических разрезов, вскрытых бурением, два метода электрокаротажа – БКЗ и БК.

Результатом измерения зондов БКЗ является кажущееся УЭС, которое вычисляется по формуле  $\rho_k = k \frac{\varphi_M - \varphi_N}{I_0^A}$ , где  $\rho_k$  – кажущееся сопротивление,  $k$  – коэффициент зонда,

$\varphi_M$  – потенциал, измеренный на электроде  $M$ ,  $\varphi_N$  – потенциал, измеренный на электроде  $N$ ,  $I_0^A$  – сила тока, протекающего через токовый электрод  $A$ . В БКЗ используются трехэлектродные зонды, которые имеют следующее обозначение: А2.0М0.5N. Здесь расстояние между электродами  $A$  и  $M$  составляет 2,0 м, а между  $M$  и  $N$  – 0,5 м.

Результатом измерения зонда БК также является кажущееся УЭС, вычисленное по той же формуле. Но в этом случае  $\varphi_M$  – потенциал, измеренный на токовом электроде, а  $\varphi_N$  – потенциал поверхности Земли или потенциал, измеренный на достаточном удалении от токового электрода, соответственно. В отличие от зондов БКЗ токовый электрод зонда БК окружен экранирующими электродами. Их главной задачей является фокусировка тока, достигаемая устранением части токов, протекающих вдоль зонда в окрестности токового электрода. Значение тока, протекающего через токовый электрод, можно получить при помощи следующей формулы:

$$I_0^A = \int_S J dS = \int_S \sigma \vec{E} dS = - \int_S \sigma \text{grad} \varphi^A dS,$$

где  $J$  – плотность тока, протекающего через поверхность электрода,  $\sigma$  – удельная электрическая проводимость,  $\vec{E}$  – напряженность электрического поля,  $\varphi^A$  – электрический потенциал на электроде,  $S$  – поверхность электрода.

Величину коэффициента зонда выбирают таким образом, чтобы кажущееся УЭС, измеренное зондом, совпадало с УЭС однородной среды.

Таким образом, результатом измерений являются каротажные диаграммы значений кажущегося УЭС в зависимости от глубины (расстояния) по скважине, описываемые вектором  $\rho^R$ .

Под прямой задачей моделирования процесса каротажа при помощи зондов БКЗ и БК будем понимать определение измеренного кажущегося УЭС по заданной функции распределения УЭС с фиксированным вектором модельных параметров. Как следует из определения кажущегося УЭС, для нахождения его значения необходимо знать значение электрического потенциала на измерительных и токовых электродах.

Далее при решении прямых задач БКЗ и БК будут учитываться не только особенности геоэлектрического строения околоскважинного пространства, но конструктивные особенности зондов (диаметр корпуса, геометрические размеры электродов и их взаимное расположение). Распределение электрического потенциала  $\varphi$  в области моделирования описывается следующей краевой задачей:

$$-\text{div} \sigma \text{grad} \varphi = 0, \quad (1)$$

$$\varphi|_{\Gamma^0} = 0, \quad (2)$$

$$\sigma \frac{\partial \varphi}{\partial \vec{n}} \Big|_{\Gamma^1} = 0, \quad (3)$$

$$\varphi|_{\Gamma^A} = \varphi_A, \quad (4)$$

$$\varphi|_{\Gamma^S} = \varphi_S, \quad (5)$$

где  $\varphi$  – потенциал напряженности  $\vec{E}$  электрического поля,  $\vec{E} = -\text{grad} \varphi$ ,  $\sigma$  – удельная электрическая проводимость  $\sigma = \rho^{-1}$ ,  $\Gamma^0$  – внешняя граница области, на которой электрический потенциал считается равным нулю (расстояние от токового электрода до внешней границы выбирается таким образом, чтобы оно существенно не влияло на значение электрического потенциала на измерительных электродах),  $\Gamma^1$  – диэлектрическая поверхность зонда,  $\Gamma^A$  – поверхность токового электрода,  $\Gamma^S$  – поверхность экранирующих электродов зонда БК (в случае зонда БКЗ соответствующее краевое условие опускается),  $\varphi_A$  – напряжение на токовом электроде,  $\varphi_S$  – напряжение на экранирующих электродах.

Представим искомый электрический потенциал в виде суммы неизвестной функции  $\varphi_0$ , равной нулю на поверхностях токового и экранирующих электродов и границе  $\Gamma^0$ , и некоторой функции  $\varphi_\Gamma$ . Данная функция  $\varphi_\Gamma$  должна удовлетворять краевым условиям исходной задачи и обладать достаточной гладкостью:

$$\begin{aligned} \varphi &= \varphi_0 + \varphi_\Gamma, & \varphi_\Gamma|_{\Gamma^0} &= 0, & \varphi_\Gamma|_{\Gamma^A} &= \varphi_A, \\ \varphi_\Gamma|_{\Gamma^S} &= \varphi_S, & \varphi_0|_{\Gamma^0 \cup \Gamma^A \cup \Gamma^S} &= 0, \\ \sigma \frac{\partial \varphi_0}{\partial \vec{n}} \Big|_{\Gamma^1} &= 0, \end{aligned} \quad (6)$$

Конкретный вид функции  $\varphi_\Gamma$  выбирается во время дискретизации задачи.

Подставим выражение (6) в задачу (1)–(5) и сформулируем новую краевую задачу относительно неизвестной функции  $\varphi_0$ :

$$-\text{div} \sigma \text{grad} \varphi_0 = f, \quad (7)$$

$$\sigma \frac{\partial \varphi_0}{\partial \vec{n}} \Big|_{\Gamma^1} = 0, \quad (8)$$

$$\varphi_0|_{\Gamma} = 0, \quad (9)$$

где  $f = \text{div} \sigma \text{grad} \varphi_\Gamma$ , а  $\Gamma = \Gamma^0 \cup \Gamma^A \cup \Gamma^S$ .

Для решения краевой задачи (7)–(9) будем использовать метод конечных элементов [16]. Пусть  $\Omega$  – трехмерная, возможно, неоднородная по физическим свойствам область с липшицнепрерывной границей.

Введем следующие функциональные пространства:

$$H^1(\Omega) = \left\{ \psi \in L^2(\Omega) \mid \text{grad} \psi \in L^2(\Omega) \right\},$$

$$H_0^1(\Omega) = \left\{ \psi \in H^1(\Omega) \mid \psi|_{\Gamma} = 0 \right\},$$

где  $L^2(\Omega)$  – пространство Лебега, для элементов которого, определим следующее скалярное произведение:

$$(u, v) = \int_{\Omega} u v d\Omega.$$

Для краевой задачи (7)–(9) сформулируем следующую вариационную постановку.

Найти  $\varphi_0 \in H_0^1(\Omega)$  такое, что  $\forall v \in H_0^1(\Omega)$  выполняется

$$\int_{\Omega} \text{grad} \varphi_0 \cdot (\sigma \text{grad} v d\Omega) = \int_{\Omega} f v d\Omega. \quad (10)$$

Краевые условия (8) и (9) учитываются в данной постановке естественным образом [16].

Приближенное решение  $\varphi_0^h$  вариационной постановки (10) будем искать в виде разложения по некоторому множеству базисных функций, образующему конечномерное подпространство  $H^h(\Omega)$  пространства  $H_0^1(\Omega)$  [16]:

$$U_0^h = \sum_{i=1}^N x_i \psi_i, \quad (11)$$

где  $\psi_i$  –  $i$ -я базисная функция,  $x_i$  – вес  $i$ -й базисной функции в разложении.

Дискретная вариационная постановка тогда будет иметь следующей вид.

Найти  $\varphi_0^h \in H^h(\Omega)$  такое, что  $\forall v^h \in H^h(\Omega)$  выполняется

$$\int_{\Omega} \text{grad} \sum_{i=1}^N x_i \psi_i \cdot (\sigma \text{grad} v^h) d\Omega = \int_{\Omega} f v^h d\Omega. \quad (12)$$

Поскольку данная постановка должна выполняться для всех функций из подпространства  $H^h(\Omega)$ , а любую такую функцию можно представить при помощи разложения по базису этого подпространства, нам достаточно потребовать, чтобы равенство (12) выполнялось только для всех базисных функций  $\psi_i$ . В результате искомое приближенное решение должно удовлетворять  $N$  следующим равенствам:

$$\int_{\Omega} \text{grad} \sum_{i=1}^N x_i \psi_i \cdot (\sigma \text{grad} \psi_j) d\Omega = \int_{\Omega} f \psi_j d\Omega, \quad j = 1, 2, \dots, N.$$

В результате приближенное решение  $\varphi_0^h$  полностью определяется вектором весов  $x = (x_1, x_2, \dots, x_N)$ , найти который можно, решив систему линейных алгебраических уравнений (СЛАУ):

$$Ax = b, \quad (13)$$

где элементы матрицы и вектора правой части можно вычислить следующим образом:

$$[A]_{i,j} = \int_{\Omega} \text{grad} \psi_i \cdot (\sigma \text{grad} \psi_j) d\Omega,$$

$$[b]_i = \int_{\Omega} f \psi_i d\Omega.$$

В качестве базисных функций  $\psi_i$  будем использовать полиномы высоких порядков, определенных на тетраэдральной сетке, построенной в расчетной области [17]. Далее под порядком базисных функций будем понимать максимальный порядок полиномов, использованных при определении базиса. Для улучшения спектральных свойств матриц, получаемых после дискретизации исходной задачи, можно ортогонализировать базисные функции. Полная ортогонализация привела бы к резкому увеличению количества ненулевых элементов матрицы. В [17] предлагается проводить частичную ортогонализацию, т. е. разбить базисные функции на множество групп, а затем выполнить ортогонализацию только внутри каждой группы. Базисные функции высоких порядков могут быть ассоциированы с узлом, ребром, гранью или с самим тетраэдром. Это зависит от того, как определяется степень свободы конкретной базисной функции. Поскольку одно ребро, грань или элемент для базисов высоких порядков ассоциированы с несколькими функциями, будем использовать это свойство в качестве разделителя на группы. Определение групп ортогонализации подобным образом не приводит к увеличению количества ненулевых элементов матрицы, а также к изменению ее портрета. В [17] для ортогонализации используется стандартное скалярное произведение.

В этой работе базисные функции внутри одной группы ортогонализуются относительно билинейной формы, которая используется для построения вариационной постановки. Таким образом, процесс ортогонализации будет эквивалентен использованию некоторой блочно-диагональной матрицы.

Матрица А СЛАУ (13) является положительно определенной и может быть легко приведена к симметричному виду [16]. Исходя из этого для нахождения решения (13) воспользуемся методом сопряженных градиентов. Поскольку решаемая задача является трехмерной, то время решения СЛАУ (13) будет существенным. Это особенно важно при решении обратной задачи, когда основные временные затраты будут приходиться на нахождение решений множества прямых задач. Для уменьшения требуемых вычислительных ресурсов воспользуемся многоуровневым алгоритмом [18]. Это полный аналог многосеточного метода [16], но в отличие от него для нахождения решения прямой задачи используется не последовательность вложенных друг в друга сеток, а последовательность вложенных функциональных подпространств. Для формирования данной последовательности вложенных подпространств воспользуемся свойством иерархичности используемых для построения дискретной вариационной постановки базисных функций [16]. А именно, множество базисных функций порядка  $p - 1$  является подмножеством множества базисных функций порядка  $p$ . Таким образом, операции интерполяции и проектирования между двумя пространствами базисных функций разного порядка вводятся естественным образом.

Также существенным является вопрос выбора порядка базисных функций, используемых для построения дискретной задачи. С одной стороны, повышение порядка базисных функций ведет к увеличению порядка аппроксимации получаемого приближенного решения и, следовательно, к возможности использования для моделирования более грубых сеток – обладающих существенно меньшим количеством узлов. С другой стороны, при увеличении порядка используемых базисных функций также увеличивается количество степеней свободы, связанных с различными геометрическими примитивами расчетной сетки. Например, для базиса первого порядка с каждым узлом сетки связана одна степень свободы, и размерность результирующей СЛАУ будет равна количеству узлов расчетной сетки. В то же время для базиса второго порядка, помимо степеней свободы, связанных с узлами сетки, возникают степени свободы, связанные с ребрами сетки. Таким образом, размерность СЛАУ будет равна количеству узлов сетки плюс количество ребер сетки. При этом необходимо отметить, что переход к базису второго порядка на одной и той же сетке не только увеличивает размерность СЛАУ, но также увеличивает и заполненность матрицы не нулевыми элементами, что также приводит к дополнительным вычислительным затратам.

Поэтому встает вопрос оптимального выбора порядка базисных функций, обеспечивающий получение приближенного решения с заданной точностью за минимальное вычислительное время.

В табл. 1 представлена зависимость времени вычисления приближенного решения прямой задачи БКЗ (модель «зонд – скважина – пласт»). Для каждого выбранного порядка базисных функций подбиралась такая тетраэдральная сетка, что относительная погрешность нахождения кажущегося УЭС была не более двух процентов. В этой таблице используются следующие обозначения:  $P$  – порядок полиномов, используемых для построения базисных функций;  $T_{CG}$  – время (в секундах), необходимое для уменьшения относительной нормы невязки СЛАУ до  $10^{-8}$  при помощи метода сопряженных градиентов;  $T_{ML}$  – время (в секундах), необходимое для уменьшения относительной нормы невязки СЛАУ до  $10^{-8}$  при помощи многоуровневого метода. Метод сопряженных градиентов использовался с блочно-диагональным предобуславливанием, или, иными словами, использовалась частичная ортогонализация базисных функций, описанная ранее.

Как следует из табл. 1, использование базисных функций высоких порядков существенно уменьшает вычислительные затраты при решении прямой задачи БКЗ. В свою очередь, использование базисных функций больших порядков позволяет применять для решения СЛАУ многоуровневый алгоритм (при использовании базиса первого порядка он вырождается в простой метод сопряженных градиентов), что также повышает вычислительную эффективность разработанных методов моделирования каротажа. Два этих подхода дают суммарное ускорения в 10 раз.

Таблица 1  
Время вычисления приближенного решения  
прямой задачи БКЗ

$P$	$T_{CG}, c$	$T_{ML}, c$
1	1 567	1 567
2	575	297
3	320	202
4	378	155

Используемые методы решения СЛАУ не исключают использования других способов предобуславливания, что также может дополнительно ускорить нахождение решения. Но это ускорение будет пропорциональным для всех рассмотренных вариантов расчетов, и не сможет нивелировать выигрыш именно от использования базиса высокого порядка и многоуровневого алгоритма.

Необходимо сказать о способе повышения быстродействия вычислений с использованием графических процессоров (GPU) [19; 20]. Так, в работе [21; 22] разработан и программно реализован алгоритм моделирования данных БКЗ на основе метода конечных элементов и высокопроизводительных гетерогенных вычислений на центральном процессоре (CPU) и GPU. Численное решение этой прямой задачи сводится к решению СЛАУ, для которого используется разложение Холецкого с последующим решением двух вспомогательных СЛАУ с треугольными матрицами. Особенностью реализации рассматриваемой задачи является то, что необходимо решать много СЛАУ с одной и той же матрицей, но с разными правыми частями, поэтому вычислительные затраты на решение двух СЛАУ с треугольными матрицами значительно превышают затраты на разложение матрицы основной СЛАУ. В работе показано, что это приводит к низкой эффективности использования только GPU и требует разработки гетерогенного метода CPU-GPU вычислений, который позволит повысить быстродействие.

### Обратная задача

Сформулируем обратную задачу каротажа как задачу поиска вектора параметров геоэлектрической модели  $w$ , минимизирующего функцию  $f(w)$  [23]:

$$f(w) = \sum_{i=1}^n \left[ \frac{\rho_i^R - \rho_i^M(w)}{\rho_i^R} \right]^2,$$

$$w^* = \arg \min_w f(w),$$

где  $n$  – количество измеренных кажущихся УЭС, полученных при каротаже;  $\rho_i^R$  – измеренное кажущееся УЭС;  $\rho_i^M(w)$  – смоделированное кажущееся УЭС, рассчитанное для модели околоскважинного пространства, заданной вектором параметров  $w$ ;  $w^*$  – решение обратной задачи.

Для решения поставленной обратной задачи воспользуемся методом покоординатного спуска [24].

Определим следующий вспомогательный алгоритм минимизации функции вдоль направления  $d$  из точки  $u$  с начальным шагом  $\beta$ :

```

dMin( $f, y, d, \beta$ ):
 $\eta = 0$ 
while true do
if  $f(y + \beta \cdot d) < f(y) - 10^{-6} \cdot \beta^2$  then  $\eta = \beta, \beta = \beta \cdot 2$ 
else return  $\eta$ 
end while

```

Результатом работы данного алгоритма является шаг  $\eta$ , на который необходимо сместиться из точки  $y$  вдоль направления  $d$ , чтобы получить новое приближенное решение  $y + \eta \cdot d$  задачи минимизации. Величина шага  $\eta$  выбирается так, чтобы новое приближение находилось как можно дальше от текущего, а не как обеспечивающая наибольшее уменьшение целевой функции. Тогда  $k$ -я итерация алгоритма покоординатного спуска примет следующий вид:

```

 $y_0 = x_{k-1}$ 
for  $i = 1, 2, \dots, n$  do
 $\beta = dMin(f, y_{i-1}, d_i, \eta_i)$ 
if  $\beta = 0$  then  $d_i = -1 \cdot d_i, \beta = dMin(f, y_{i-1}, d_i, \eta_i)$ 
if  $\beta = 0$  then  $\eta_i = \eta_i \cdot 0,5$ 
 $y_i = y_{i-1} + \beta_i \cdot d_i$ 
end for

```

где  $x_k$  – приближенное решение задачи минимизации на  $k$ -й итерации;  $d_i$  – вектор, параллельный  $i$ -й оси координат в начальный момент времени  $d_i = e_i$ ;  $e_i$  –  $i$ -я ось координат,  $\eta_i$  – шаг поиска нового приближенного решения вдоль  $i$ -й оси координат. Эффективность использования алгоритма минимизации для решения обратных задач рассматривается в [24].

### Результаты решения обратной задачи каротажа

С использованием разработанного алгоритма трехмерной численной инверсии данных методов электрокаротажа проведено тестирование в типичных моделях сред, учитывающих наклон главных осей тензора анизотропии. В приведенных тестах для решения прямой задачи использовались базисные функции 4-го порядка. Анализ точности восстановления горизонтальной и вертикальной компонент УЭС, а также угла наклона оси скважины относительно локальной системы координат, в которой тензор УЭС является диагональным, выполнен в следующих модельных задачах. Первая включает скважину радиусом 0,108 м с УЭС бурового раствора 1 Ом·м и пласт с горизонтальной компонентой УЭС 4 Ом·м, вертикальной – 9 Ом·м, а также различные углы наклона осей тензора УЭС относительно скважины ( $\alpha = 0, 15, 30, 45, 60, 75$  и  $90^\circ$ ). Во второй – горизонтальная и вертикальная компоненты УЭС пласта имеют значения 15 и 75 Ом·м соответственно, и те же углы наклона главных осей тензора УЭС.

Для получения вектора измеренных кажущихся УЭС выполнено моделирование процесса каротажа с использованием четырех зондов БКЗ: А0.4М0.1N, А1.0М0.1N, А2.0М0.5N, А4.0М0.5N, и зонда БК. При решении обратных задач использовались измерения с одной глубины. К полученным синтетическим каротажным данным предварительно были добавлены нормально распределенные случайные величины, пропорциональные начальным данным. Так, в табл. 2 приведены результаты решения обратной задачи для первой и второй моделей с дисперсией шума 0,025. В таблице используются следующие обозначения:  $\alpha$  – угол наклона осей тензора к оси скважины у исходной модели,  $\rho_h$  – восстановленная горизонтальная ком-

понента УЭС,  $\rho_v$  – восстановленная вертикальная компонента УЭС,  $\alpha_i$  – восстановленный угол наклона главных осей тензора к оси скважины.

Таблица 2

## Результаты решения обратной задачи

$\alpha$	$\rho_h$	$\rho_v$	$\alpha_i$
Для модели: пласт с горизонтальной компонентой УЭС 4 Ом·м, вертикальной компонентой УЭС 9 Ом·м, дисперсия шума 0,025			
0	3,95	12,5	0,812
15	4,00	12,8	18,9
30	4,02	12,54	28,5
45	4,46	11,6	37,1
60	4,71	10,9	82,1
75	4,64	11,5	90,7
90	4,06	12,3	94,6
Для модели: пласт с горизонтальной компонентой УЭС 15 Ом·м, вертикальной компонентой УЭС 75 Ом·м, дисперсия шума 0,025			
0	15,01	79,1	0,59
15	13,6	101,5	31,2
30	11,98	110,8	49,2
45	16,0	85,2	41,27
60	19,0	58,5	64,2
75	20,8	60,4	88,8
90	17,6	72,6	89,9

На основании результатов, приведенных в табл. 2, можно сделать следующие выводы: точнее всего удастся восстановить горизонтальную составляющую УЭС пласта, восстановленная вертикальная составляющая имеет несколько большую погрешность по сравнению с горизонтальной, при этом погрешность восстановления угла наклона в некоторых случаях может составлять до  $20^\circ$ . С одной стороны, такая погрешность восстановления угла наклона является значительной, но с другой стороны, она постулирует саму возможность восстановления угла при помощи зондов, изначально не рассчитанных на решение подобных задач. При решении практических задач теперь появляется возможность такой оценки, что дает дополнительную существенную информацию о строении околоскважинного пространства.

### Заключение

На основании проведенных вычислительных экспериментов можно сделать вывод, что совместное использование в конечноэлементной аппроксимации базисных функций больших порядков и многоуровневого алгоритма решения СЛАУ позволят существенно ускорить решение прямых и обратных задач каротажа. Использование комплекса данных БКЗ и БК при решении совместной обратной задачи позволяет восстанавливать значения элементов диагонального тензора удельного электрического сопротивления, а также проводить оценку угла наклона главных осей тензора электрической анизотропии. Относительная погрешность

идентификации угла наклона обусловлена тем, что изначально каротажные методы БКЗ и БК не были предназначены для решения задач, связанных с изучением электрической анизотропии. Однако полученные результаты по определению компонент тензора электрической анизотропии и оценке угла наклона его главных осей (пусть и с некоторой погрешностью) открывают принципиально новые возможности при интерпретации повсеместно применяющихся методов БКЗ и БК для изучения геологических разрезов в Западной и Восточной Сибири. В заключение отметим, что с использованием разработанного алгоритма впервые становится возможным изучение эффекта электрической анизотропии при исследовании трещиноватых карбонатных коллекторов, в которых наличие электрической анизотропии обусловлено системой наклонных трещин. Именно этим типом анизотропии обладают глубокопрогруженные коллекторы в донорском фундаменте, которые являются чрезвычайно перспективными. Информация о наклоне трещин и их преимущественном направлении играет существенную роль при проектировании наклонно-горизонтальных скважин, достоверная оценка которой возможна методами электрокаротажа.

### Список литературы

1. Эпов М. И., Глинских В. Н. Быстрое двумерное моделирование высокочастотного электромагнитного поля для задач каротажа // Геология и геофизика. 2003. Т. 44, № 9. С. 942–952.
2. Глинских В. Н., Никитенко М. Н., Эпов М. И. Моделирование и инверсия данных электромагнитного каротажа в пластах конечной мощности, вскрытых на биополимерных и нефтяных буровых растворах // Геология и геофизика. 2013. Т. 54, № 11. С. 1803–1813.
3. Глинских В. Н., Никитенко М. Н., Эпов М. И. Линеаризованные решения прямых и обратных двумерных задач высокочастотного электромагнитного каротажа в проводящих средах с учетом токов смещения // Геология и геофизика. 2013. Т. 54, № 12. С. 1942–1951.
4. Михайлов И. В., Глинских В. Н., Никитенко М. Н., Суродина И. В. Совместная численная инверсия данных индукционных и гальванических каротажных зондирований в моделях геологических сред с осевой симметрией // Геология и геофизика, 2017. Т. 58, № 6. С. 935–947.
5. Табаровский Л. А., Каганский А. М., Эпов М. И. Электромагнитное поле гармонического источника в анизотропной цилиндрически-слоистой среде // Геология и геофизика. 1976. № 3. С. 94–99.
6. Табаровский Л. А., Эпов М. И. Геометрическая и частотная и фокусировка при изучении анизотропных пластов // Электромагнитные методы исследования скважин. Новосибирск: ИГиГ СО АН СССР, 1982. С. 67–129.
7. Anderson B. I., Barber T. D., Lüling M. G. The response of induction tools to dipping, anisotropic formations // Transactions of the SPWLA 36th Annual Logging Symposium, 1995. Pap. D.
8. Rabinovich M., Bepalov A., Corley B., Merchant G., Wang T., Quint E., Morrison J. Effect of fractures on multi-component and multiarray induction logs // Transactions of the SPWLA, 45<sup>th</sup> Annual Logging Symposium, 2004. Pap. UU.
9. Wang T. A weak-anisotropy approximation to multicomponent induction responses in cross-bedded formations // Geophysics. 2006. Vol. 71. Iss. 4. P. F61–F66.
10. Zhang Zh., Yu L., Krieghauser B. et al. Determination of relative angles and anisotropic resistivity using multicomponent induction logging data // Geophysics. 2004. Vol. 69, № 4. P. 898–909.
11. Zhdanov M. S., Tartaras E., Gribenko A. Fast 3D Imaging from a Single Borehole Using Tensor Induction Logging Data // Petrophysics. 2004. Vol. 45, № 2. P. 167–178.
12. Эпов М. И., Глинских В. Н., Никитенко М. Н. Способ измерения удельной электропроводности и электрической макроанизотропии горных пород. Патент на изобретение RU 2525149. Оpubл. 10.08.2014. Бюлл. № 22.
13. Эпов М. И., Еремин В. Н., Маништейн А. К., Петров А. Н., Глинских В. Н. Устройство для измерения удельной электропроводности и электрической макроанизотропии горных пород. Патент на изобретение RU 2528276. Оpubл. 10.09.2014. Бюлл. № 25.

14. Эпов М. И., Еремин В. Н., Петров А. Н., Глинских В. Н. Электромагнитный зонд для каротажа в нефтегазовых скважинах. Патент на изобретение RU 2583867. Оpubл. 10.05.2016. Бюлл. № 13.
15. Эпов М. И., Никитенко М. Н., Глинских В. Н. Математическое обоснование нового электромагнитного зонда с тороидальными катушками для высокоразрешающего каротажа нефтегазовых скважин // Вестник НГУ. Серия: Информационные технологии. 2018. Т. 16, № 1. С. 113–129.
16. Шайдулов В. В. Многосеточные методы конечных элементов. М.: Наука, 1989. 288 с.
17. Webb J. P. Hierarchical vector basis functions of arbitrary order for triangular and tetrahedral finite elements // IEEE Transactions on antennas and propagation. 1999. Vol. 47, № 8. P. 1244–1253.
18. Aghabarati A., Webb J. P. Multilevel Methods for adaptive Finite Element Analysis of Electromagnetic Scattering // IEEE Transactions on Antennas and Propagation. 2013. Vol. 61, № 11. P. 5597–5606.
19. Глинских В. Н., Эпов М. И., Лабутин И. Б. Моделирование диаграмм электромагнитного каротажа на графических процессорах // Вычислительные технологии. 2008. Т. 13, № 6. С. 50–60.
20. Глинских В. Н., Горбатенко В. А. Инверсия данных электромагнитного каротажа на графических ускорителях // Вычислительные технологии. 2015. Т. 20, № 1. С. 25–37.
21. Глинских В. Н., Дудаев А. Р., Нечаев О. В. Высокопроизводительные гибридные CPU-GPU вычисления в задаче электрического каротажа нефтегазовых скважин // Вычислительные технологии. 2017. Т. 22, № 3. С. 16–31.
22. Glinskikh V. N., Dudaev A. R., Nechaev O. V. High-performance simulation of electrical logging data in petroleum reservoirs using graphics processors // Parallel Computational Technologies (PCT-2017). Communications in Computer and Information Science. 2017. Vol. 753. P. 186–200.
23. Zhdanov M. S. Geophysical inverse theory and regularization problems. Amsterdam: Elsevier Science, 2002.
24. Нечаев О. В., Глинских В. Н. Быстрый прямой метод решения обратной задачи электрического каротажа в нефтегазовых скважинах // Вестник НГУ. Серия: Информационные технологии. 2017. Т. 15, № 4. С. 53–63.

Материал поступил в редколлегию 19.09.2018

**O. V. Nechaev<sup>1</sup>, V. N. Glinskikh<sup>1-3</sup>**

<sup>1</sup> Trofimuk Institute of Petroleum Geology and Geophysics SB RAS  
3 Academician Koptyug Ave., Novosibirsk, 630090, Russian Federation

<sup>2</sup> Novosibirsk State University  
1 Pirogov Str., Novosibirsk, 630090, Russian Federation

<sup>3</sup> Novosibirsk State Technical University  
20 K. Marx Ave., Novosibirsk, 630090, Russian Federation

oleg.v.nechaev@gmail.com, GlinskikhVN@ipgg.sbras.ru

**THREE-DIMENSIONAL SIMULATION  
AND INVERSION OF LATERAL LOGGING SOUNDING  
AND LATERAL LOGGING DATA IN MEDIA WITH TILT OF THE MAIN AXES  
OF THE DIELECTRIC ANISOTROPY TENSOR**

The paper for the first time formulates the direct and inverse problems of electrical logging on the determination of the components of the electrical anisotropy tensor of rocks as measured by the complex of lateral logging sounding and lateral logging methods. The features of the finite element approximation of the direct problem using basic functions of large order are considered. The possi-

bility of reconstructing the horizontal and vertical resistivity components, as well as the angle of tilt of the axes of the electrical anisotropy tensor, is investigated.

*Keywords:* finite element method, multilevel method, forward and inverse problem, electrical logging, electrical resistivity tensor.

## References

1. Epov M. I., Glinskikh V. N. Fast two-dimensional simulation of high-frequency electromagnetic field in induction logging. *Russian Geology and Geophysics*, 2003, vol. 44, no. 9, p. 904–915. (in Russ.)
2. Glinskikh V. N., Nikitenko M. N., Epov M. I. Numerical modeling and inversion of electromagnetic logs in the wells drilled with biopolymer and oil-mud. *Russian Geology and Geophysics*, 2013, vol. 54, no. 11, p. 1409–1416. (in Russ.)
3. Glinskikh V. N., Nikitenko M. N., Epov M. I. Processing high-frequency electromagnetic logs from conducting formations: linearized 2D forward and inverse solutions with regard to eddy currents. *Russian Geology and Geophysics*, 2013, vol. 54, no. 12, p. 1515–1521. (in Russ.)
4. Mikhaylov I. V., Glinskikh V. N., Nikitenko M. N., Surodina I. V. Joint inversion of induction and galvanic logging data in axisymmetric geological models. *Russian Geology and Geophysics*, 2017, vol. 58, no. 6, p. 752–762. (in Russ.)
5. Tabarovskiy L. A., Kagansky A. M., Epov M. I. The electromagnetic field of a harmonic source in an anisotropic cylindrical-layered medium. *Geology and Geophysics*, 1976, vol. 3, p. 94–99. (in Russ.)
6. Tabarovskiy L. A., Epov M. I. Geometric and frequency and focusing in the study of anisotropic seams. *Electromagnetic Methods of Well Research*. Novosibirsk, IGI SB AS USSR, 1982, p. 67–129. (in Russ.)
7. Anderson B. I., Barber T. D., Lüling M. G. The response of induction tools to dipping, anisotropic formations. *Transactions of the SPWLA 36<sup>th</sup> Annual Logging Symposium*, 1995, pap. D.
8. Rabinovich M., Bernalov A., Corley B., Merchant G., Wang T., Quint E., Morrison J. Effect of fractures on multi-component and multiarray induction logs. *Transactions of the SPWLA, 45<sup>th</sup> Annual Logging Symposium*, 2004, pap. UU.
9. Wang T. A weak-anisotropy approximation to multicomponent induction responses in cross-bedded formations. *Geophysics*, 2006, vol. 71, iss. 4, p. F61–F66.
10. Zhang Zh., Yu L., Krieghauser B. et al. Determination of relative angles and anisotropic resistivity using multicomponent induction logging data. *Geophysics*, 2004, vol. 69, no. 4, p. 898–909.
11. Zhdanov M. S., Tartaras E., Gribenko A. Fast 3D Imaging from a Single Borehole Using Tensor Induction Logging Data. *Petrophysics*, 2004, vol. 45, no. 2, p. 167–178.
12. Epov M. I., Glinskikh V. N., Nikitenko M. N. Method for measuring the specific electric conductivity and electrical macroanisotropy of rocks. Invention patent RU 2525149. 10.08.2014. № 22. (in Russ.)
13. Epov M. I., Eremin V. N., Manstein A. K., Petrov A. N., Glinskikh V. N. Device for measuring the specific electrical conductivity and electrical macroanisotropy of rocks. Invention patent RU 2528276. 10.09.2014. № 25. (in Russ.)
14. Epov M. I., Eremin V. N., Petrov A. N., Glinskikh V. N. Electromagnetic tool for logging in oil and gas wells. Invention patent RU 2583867. 10.05.2016. № 13. (in Russ.)
15. Epov M. I., Nikitenko M. N., Glinskikh V. N. Mathematical substantiation of a new electromagnetic tool with toroidal coils for high-resolution logging of oil and gas wells. *Vestnik NSU. Series: Information Technology*, 2018, vol. 16, no. 1, p. 113–129. (in Russ.)
16. Shaydurov V. V. Multigrid methods of finite elements. Moscow, Science, 1989, p. 288. (in Russ.)
17. Webb J. P. Hierarchical vector basis functions of arbitrary order for triangular and tetrahedral finite elements. *IEEE Transactions on antennas and propagation*, 1999, vol. 47, no. 8, p. 1244–1253.

18. Aghabarati A., Webb J. P. Multilevel Methods for adaptive Finite Element Analysis of Electromagnetic Scattering. *IEEE Transactions on Antennas and Propagation*, 2013, vol. 61, no. 11, p. 5597–5606.
19. Glinskikh V. N., Epov M. I., Labutin I. B. Electromagnetic logging data simulation on GPUs. *Computational Technologies*, 2008, vol. 13, no. 6, p. 50–60. (in Russ.)
20. Glinskikh V. N., Gorbatenko V. A. Electromagnetic logging data inversion on GPU. *Computational Technologies*, 2015, vol. 20, no. 1, p. 25–37. (in Russ.)
21. Glinskikh V. N., Dudaev A. R., Nechaev O. V. High-performance CPU-GPU heterogeneous computing in resistivity logging of oil and gas wells. *Computational Technologies*, 2017, vol. 22, no. 3, p. 16–31. (in Russ.)
22. Glinskikh V. N., Dudaev A. R., Nechaev O. V. High-performance simulation of electrical logging data in petroleum reservoirs using graphics processors. *Parallel Computational Technologies (PCT-2017). Communications in Computer and Information Science*, 2017, vol. 753, p. 186–200.
23. Zhdanov M. Geophysical inverse theory and regularization problems. Amsterdam, Elsevier Science, 2002.
24. Nechaev O. V., Glinsky V. N. Fast direct method for solving the inverse problem of electric logging in oil and gas wells. *Vestnik NSU. Series: Information Technology*, 2017, vol. 15, no. 4, p. 53–63. (in Russ.)

*Received 19.09.2018*

*For citation:*

Nechaev O. V., Glinskikh V. N. Three-Dimensional Simulation and Inversion of Lateral Logging Sounding and Lateral Logging Data in Media with Tilt of the Main Axes of the Dielectric Anisotropy Tensor. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 4, p. 127–139. (in Russ.) DOI 10.25205/1818-7900-2018-16-4-127-139

**А. Б. Нугуманова<sup>1</sup>, Е. М. Байбурин<sup>1</sup>, М. Е. Мансурова<sup>2</sup>, В. Б. Барахнин<sup>3,4</sup>**

<sup>1</sup> *Восточно-Казахстанский государственный университет им. С. Аманжолова  
ул. 30-й Гвардейской Дивизии, 34, Усть-Каменогорск, 070002, Казахстан*

<sup>2</sup> *Казахский национальный университет им. аль-Фараби  
пр. аль-Фараби, 71, 050040, Алматы, Казахстан*

<sup>3</sup> *Институт вычислительных технологий СО РАН  
пр. Акад. Лаврентьева, 6, Новосибирск, 630090, Россия*

<sup>4</sup> *Новосибирский государственный университет  
ул. Пирогова, 1, Новосибирск, 630090, Россия*

*{anugumanova, ebaiburin}@vkgu.kz, mansurova.madina@gmail.com, bar@ict.nsc.ru*

## **АВТОМАТИЧЕСКОЕ ИЗВЛЕЧЕНИЕ РЕШЕТОК ПОНЯТИЙ ИЗ МЕДИЦИНСКИХ ТЕКСТОВ НА ОСНОВЕ КОМБИНАЦИИ АНАЛИЗА ФОРМАЛЬНЫХ ПОНЯТИЙ И ТЕХНОЛОГИЙ БУТСТРАППИНГА\***

Рассматривается новый способ извлечения понятий из текстов предметной области на основе комбинации анализа формальных понятий и бутстрап-технологии информационного поиска. Анализ формальных понятий представляет собой мощный аппарат автоматического вывода понятий предметной области, однако он рассчитан на высокое качество входных данных, без пропусков и неточностей. Получение таких наборов данных напрямую из текстов затруднено в силу сильной разреженности текстовых корпусов. Соответственно, представляется перспективным улучшение качества входных данных за счет применения бутстраппинга – технологии, обеспечивающей интеллектуальный поиск фрагментированной информации в сети Интернет. Цель данной работы – показать, что при правильном выборе исходных шаблонов поиска бутстраппинг, основанный на использовании открытых ресурсов Интернета как ценных источников знаний, превращается в эффективный инструмент поддержки концептуального моделирования.

*Ключевые слова:* анализ формальных понятий, бутстраппинг, извлечение информации, поверхностный лингвистический анализ, информационный поиск.

### **Введение**

В данной работе мы предлагаем новый способ автоматического извлечения понятий из текстов медицинской тематики, основанный на заполнении пропусков в сильно разреженных матрицах совместной встречаемости терминов. Наш подход опирается на строгие и изящные формулировки анализа формальных понятий – метода, который, как отмечается в работе [1], использует язык алгебры для описания понятий и их иерархий.

Входные данные для анализа формальных понятий представляются в виде объектно-признаковой таблицы, отражающей распределение признаков по объектам предметной об-

---

\* Работа выполнена при частичной поддержке РФФИ (проект 18-07-01457).

ласти. Таблица является бинарной: если объект, указанный в строке таблицы, обладает признаком, указанном в столбце таблицы, то на пересечении соответствующих строки и столбца ставится 1, иначе 0. Математический аппарат анализа формальных понятий позволяет выделить в этой таблице множества объектов, обладающих одинаковыми наборами признаков. Считается, что каждое такое множество объектов и их признаков образует одно формальное понятие. Объекты этого множества называются объемом понятия (экстенционалом), а признаки этого множества – содержанием понятия (интенционалом).

Особенность нашей работы заключается в том, что в качестве входной объектно-признаковой таблицы мы используем матрицу совместной встречаемости терминов. Термины, образующие строки и столбцы матрицы, мы получаем, извлекая из текстов медицинской тематики конструкции, удовлетворяющие лексико-синтаксическим шаблонам вида «Существительное + Существительное». Один из самых результативных шаблонов этого вида представляет собой пару «Существительное + Существительное в родительном падеже». Так, например, при помощи этого шаблона мы извлекаем такие словосочетания, как «приступ холестистита», «приступ панкреатита», «воспаление нерва», «нарушение кровоснабжения», «обострение гайморита», «осложнение гайморита» и т. д. Термины, стоящие в указанных парах на первом месте, мы записываем в столбцы матрицы (признаки), а термины, стоящие на втором месте, – в строки (объекты). Также возможна работа и с транспонированной матрицей, тогда термины, стоящие на первом месте, мы записываем в строки, а термины, стоящие на втором месте, – в столбцы.

К сожалению, полученную таким способом входную матрицу невозможно напрямую использовать для извлечения формальных понятий, так как для этого она слишком неоднородная и разреженная. Дело в том, что строки этой матрицы отражают весь гетерогенный набор объектов, встречающихся в текстах медицинской тематики: названия заболеваний, лекарств, методов лечения, органов человеческого тела и т. д. Соответственно и столбцы этой матрицы отражают весь спектр разнородных признаков, каждый из которых присущ только одной «своей» группе объектов (например, признаками «обострение» и «осложнение» описываются заболевания, но не органы, а признаками «воспаление» и «отек» описываются органы, но не заболевания). Как следствие, увеличивается и без того сильная разреженность, изначально присущая текстовому корпусу.

Для устранения неоднородности мы подвергаем исходную матрицу кластеризации, в результате чего из нее получают более однородные и менее разреженные матрицы меньшего размера – кластеры, представляющие собой группы объектов со сходными признаками. Эти матрицы мы используем как отправные шаблоны для бутстраппинга – технологии, позволяющей находить недостающую информацию по ее начальным фрагментам. В частности, мы извлекаем из указанных матриц все нулевые пары «объект – признак» и для каждой такой пары формируем поисковый запрос к Интернету как к более репрезентативному корпусу текстов. Таким образом, для каждой нулевой пары «объект – признак» мы проверяем, существует ли устойчивый контекст употребления данного объекта с данным признаком в таком сверхбольшом текстовом корпусе, как Интернет. Если такой контекст существует и является устойчивым, то мы заменяем соответствующий нулевой элемент в матрице на единичный.

Только после восстановления пропусков и увеличения количества значимой информации в матрицах мы используем их для извлечения понятий. Как отмечалось выше, для этого мы применяем аппарат анализа формальных понятий, который позволяет не только автоматически выводить понятия предметной области, но и формировать их иерархии. Цель нашей работы заключается в том, чтобы продемонстрировать на конкретном корпусе текстов, как работает предлагаемый подход и насколько эффективно он решает задачу автоматического извлечения понятий.

В соответствии с поставленной целью дальнейшее изложение работы ведется следующим образом. Сначала мы приводим основные положения анализа формальных понятий. В последующих разделах мы описываем 4 основных этапа предлагаемого подхода.

- Этап 1: построение исходной объектно-признаковой матрицы на основе поверхностного лингвистического анализа.
- Этап 2: разбиение исходной объектно-признаковой матрицы на ряд однородных матриц посредством кластеризации.

- Этап 3: восстановление пропусков в полученных однородных матрицах методом бутстраппинга.
- Этап 4: извлечение из восстановленных матриц понятий предметной области и построение их иерархий на основе анализа формальных понятий.

В заключение мы излагаем основные выводы и приводим план будущей работы.

### Математический аппарат анализа формальных понятий

Входными данными для анализа формальных понятий (АФП) служит информация о распределении признаков среди объектов предметной области [2]. Указанная информация записывается в виде так называемого формального контекста  $K$ , который представляет собой тройку

$$K = \langle G, M, I \rangle,$$

где  $G$  – это множество объектов,  $M$  – множество признаков,  $I$  – соответствие между  $G$  и  $M$ :  $gIm$  означает, что объект  $g \in G$  обладает признаком  $m \in M$ . Оставим пока в стороне вопрос, каким образом из предметной области выбираются объекты и их признаки для формирования формального контекста. Предположим, что формальный контекст уже задан, и покажем, как в заданном контексте выделяются формальные понятия.

Пусть в формальном контексте  $K = \langle G, M, I \rangle$  выбраны произвольные подмножества объектов  $A \subseteq G$  и признаков  $B \subseteq M$ . Операторы Галуа для указанных подмножеств определяются следующим образом:

- $A' = \{m \in M \mid \forall g \in A gIm\}$ , т. е.  $A'$  – это множество признаков, которыми обладают все объекты из  $A$ ;
- $B' = \{g \in G \mid \forall m \in B gIm\}$ , т. е.  $B'$  – это множество признаков, которыми обладают все объекты из  $B$ .

Тогда формальным понятием контекста  $K$  называется пара вида  $(A, B)$ ,  $A \subseteq G$  и  $B \subseteq M$ , такая что  $A' = B$  и  $B' = A$ . Множества  $A$  и  $B$  называются соответственно объемом и содержанием формального понятия  $(A, B)$ .

Между содержанием и объемом понятия существует обратная зависимость: чем больше содержание, тем меньше объем. Другими словами, чем больше признаков содержит данное понятие, тем меньше объектов оно охватывает, и наоборот.

Два понятия  $(A_1, B_1)$  и  $(A_2, B_2)$  называют частично упорядоченными:  $(A_1, B_1) \leq (A_2, B_2)$ , если объем первого понятия входит (вложен) в объем второго:  $A_1 \subseteq A_2$ . Множество всех понятий контекста  $K$ , упорядоченных по вложению их объемов, называется решеткой понятий. Для визуального представления решеток применяются диаграммы Хассе, где сверху показаны наименьшие по объему понятия, а снизу – наибольшие [1; 3].

Поясним основные положения АФП на конкретном примере.

На рис. 1 представлен фрагмент формального контекста, сформированного на основе автоматической обработки корпуса медицинских текстов. Контекст задан обычным для АФП способом – в виде объектно-признаковой таблицы.

A	B	C	D	E	F	G	H	I	J	K	L
	"наличие"	"уровень"	"количеств..."	"содержан..."	"активност..."	"раствор"	"концентр..."	"потребле..."	"накоплен..."	"добавлен..."	"утилизац..."
"амилаза"		X	X	X	X	X	X			X	
"аминотра..."		X	X	X	X	X	X			X	
"глюкоза"	X	X	X	X	X	X	X	X	X	X	X
"креатинин"	X	X	X	X	X	X	X	X	X	X	X
"лактат"	X	X	X	X	X	X	X	X	X	X	X
"мочевина"	X	X	X	X	X	X	X	X	X	X	X
"трансфер..."		X	X	X	X	X	X			X	

Рис. 1. Фрагмент формального контекста

В табл. 1 перечислены все четыре формальных понятия, выведенных из рассматриваемого формального контекста. Первое формальное понятие самое широкое по объему, оно включает в себя все 7 объектов, поскольку все они описываются такими общими признаками, как «уровень», «количество», «содержание», «активность» и «концентрация». Утрируя, можно сказать, что в реальности этому формальному понятию соответствует понятие вещества. Второе формальное понятие включает в себя 3 объекта, которые обладают такими дополнительными признаками, как «раствор» и «добавление». Аналогично можно сказать, что в реальности этому формальному понятию соответствует понятие растворимого вещества. Самым узким по объему формальным понятием является четвертое, оно содержит всего 2 объекта: «глюкоза» и «лактат». Этому формальному понятию соответствует понятие вещества, которое можно употреблять в пищу.

Таблица 1

Формальные понятия, выведенные из формального контекста на рис. 1

№	Понятие	Объем понятия	Содержание понятия
1	$(A_1, B_1)$	$A_1 = \{\text{амилаза, аминотрансфераза, глюкоза, креатинин, лактат, мочеви́на, трансфераза}\}$	$B_1 = \{\text{уровень, количество, содержание, активность, концентрация}\}$
2	$(A_2, B_2)$	$A_2 = \{\text{амилаза, глюкоза, лактат}\}$	$B_2 = B_1 \cup \{\text{раствор, добавление}\}$
3	$(A_3, B_3)$	$A_3 = \{\text{глюкоза, креатинин, лактат, мочеви́на}\}$	$B_3 = B_1 \cup \{\text{утилизация, накопление, наличие}\}$
4	$(A_4, B_4)$	$A_4 = \{\text{глюкоза, лактат}\}$	$B_4 = B_2 \cup B_3 \cup \{\text{потребление}\}$

На рис. 2 изображена диаграмма Хассе, состоящая из двух решеток понятий, построенных на основе рассматриваемого формального контекста. Однако, анализируя полученные иерархии и классы объектов, эксперт предметной области мог бы заметить, что они не корректны. Например, креатинин и мочеви́на несправедливо исключены из понятия «растворимое вещество», – они тоже могут растворяться. В этом и состоит ключевая проблема анализа формальных понятий: если контекст является разреженным и ограниченным, то выходные понятия и их иерархии будут формироваться некорректно.

### Построение входной объектно-признаковой матрицы

Для автоматического построения входной объектно-признаковой матрицы на основе корпуса текстов мы используем поверхностный лингвистический анализ (shallow linguistic analysis). Поверхностный лингвистический анализ – это весьма популярный подход к обработке естественного языка, когда речь идет о создании практических приложений в области анализа текстов, ориентированных на конкретные интересы пользователей [4].

Как следует из его названия, подход не фокусируется на глубоком (многоуровневом и многоаспектном) разборе текстов, его цель – быстрое выделение и анализ только тех текстовых фрагментов, которые содержат релевантную с точки зрения пользователя информацию [5]. Благодаря этому подход демонстрирует

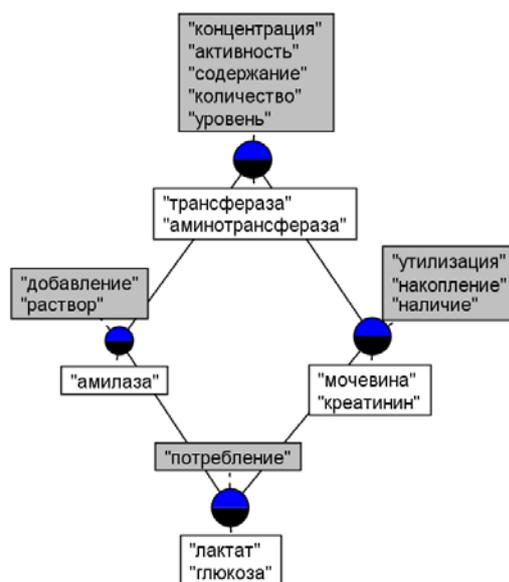


Рис. 2. Диаграмма Хассе

высокую производительность и устойчивость на сколь угодно больших корпусах текстов. Вместе с тем он требует обязательной спецификации того, что считать релевантной информацией, поскольку для выделения такой информации из текста необходима настройка шаблонов, определяющих соответствие между естественно-языковыми высказываниями и знаниями предметной области.

В этом и заключается основной недостаток поверхностных лингвистических методов. Во-первых, как отмечается в [5], подобная «тонкая» настройка методов под конкретную предметную область препятствует их переносу на другие предметные области или задачи. Во-вторых, даже внутри исходной предметной области остается неохваченным большой пласт релевантной информации, так как живая структура естественного языка плохо поддается шаблонизации.

Применительно к нашей задаче релевантной является информация о медицинских объектах и их признаках. Под медицинскими объектами мы понимаем объекты и явления, которые описываются в текстах медицинской тематики: это заболевания, диагнозы, способы лечения, лекарственные препараты, тело человека и т. д.

Новизна нашего подхода заключается в том, что указанные объекты и их признаки мы ищем в текстах корпуса на основе поверхностных лингвистических шаблонов вида «Существительное + Существительное в родительном падеже». Другими словами, мы формируем некоторое начальное множество объектов и их признаков путем извлечения из текстов пар слов, удовлетворяющих заданному шаблону. В качестве признаков мы рассматриваем слова, стоящие в парах на первом месте, а в качестве объектов – слова, стоящие на втором месте.

Однако предлагаемый нами способ поиска объектов и признаков наследует недостатки, присущие поверхностным лингвистическим методам. Во-первых, он плохо адаптируется к изменениям параметров исходной задачи. Во-вторых, если использовать предлагаемый шаблон напрямую, без дополнительного разбора предложения, то он не покрывает всего множества релевантных объектов и их признаков. Например, объекты и их признаки, разделенные определениями, будут потеряны, т. е. конструкции вида «воспаление нерва», «осложнение отита» и т. п. будут извлечены, в то время как конструкции вида «воспаление зрительного нерва», «осложнение гнойного отита» и т. п. будут пропущены. Поэтому нам приходится расширять данный шаблон.

Для практической реализации предлагаемого способа извлечения объектов и их свойств мы используем программный пакет PullEnti – свободно распространяемый набор средств разработки (SDK) для создания приложений по анализу текстов [6; 7]. Мы подключаем указанные средства разработки в наш проект, реализуемый на языке C#, как внешнюю DLL-библиотеку. На рис. 3 представлен пример работы SDK PullEnti по извлечению объектов и их признаков из текстового фрагмента, описывающего симптомы гайморита: «Обычно гайморит – это следствие осложнений после инфекционного заболевания, например, скарлатины, гриппа, простуды. Основные симптомы гайморита – затрудненное дыхание, постоянный насморк, заложенность носа и головная боль». Из данного фрагмента были извлечены такие пары объектов и их признаков, как «симптом гайморита», «заложенность носа», «следствие осложнения», «следствие заболевания».

На рис. 4 представлен еще один пример работы SDK PullEnti. Здесь объекты и их признаки извлекаются из текстового фрагмента, описывающего осложнение гнойной ангины: «Синдром Лемьера – редкое, но серьезное осложнение гнойной ангины, иногда ее называют постангинальным сепсисом». Из данного фрагмента были извлечены пары «синдром Лемьера» и «осложнение ангины». Как мы видим, несмотря на то что в этом фрагменте объекту «ангина» предшествует определение «гнойная», а признаку «осложнение» предшествует определение «серьезное», это не мешает программе правильно установить объект и его признак.

Извлеченные таким образом пары «объект – признак» служат исходным материалом для заполнения объектно-признаковой матрицы: все названия объектов (без повторений) записываются в строках матрицы, а все названия признаков (тоже без повторений) записываются в столбцах. Для каждой извлеченной пары «объект – признак» на пересечении соответствующих ей строки и столбца матрицы ставится 1. Все остальные элементы матрицы, для которых нет соответствий, обнуляются. Ранее на рис. 1 был представлен фрагмент одной из построенных таким образом объектно-признаковых матриц.

Список объектов (можно выбирать текущий)		Текущий объект	
Тип	Краткое описание	Атрибут	Значение
Семантический объект	ГАЙМОРИТ	Класс	Object
Семантический объект	ОСЛОЖНЕНИЕ	Значение	СИМПТОМ
Семантический объект	ИНФЕКЦИЯ	Свойство	ГАЙМОРИТ
Актант предиката	Объект: ИНФЕКЦИОННОЕ		
Семантический объект	ЗАБОЛЕВАНИЕ; ИНФЕКЦИОННОЕ		
Семантический объект	СЛЕДСТВИЕ; ОСЛОЖНЕНИЕ и ЗАБОЛЕВАНИЕ		
Семантический объект	НАПРИМЕР		
Семантический объект	СКАРЛАТИНА		
Семантический объект	ГРИПП		
Семантический объект	ПРОСТУДА		
Семантический объект	СИМПТОМ; ГАЙМОРИТ		
Семантический объект	ЗАТРУДНЕНИЕ		

Рис. 3. Работа программного пакета PullEnti по извлечению объектов и их признаков из текста с описанием симптомов гайморита

Тип	Краткое описание	Атрибут	Значение
Семантический объект	"ЛЕМЬЕРА"	Класс	Object
Семантический объект	СИНДРОМ; "ЛЕМЬЕРА"	Значение	ОСЛОЖНЕНИЕ
Семантический объект	РЕДКОЕ	Свойство	СЕРЬЕЗНОЕ
Семантический объект	ГНОЙ	Свойство	АНГИНА; ГНОЙНАЯ
Актант предиката	Объект: ГНОЙНАЯ		
Семантический объект	АНГИНА; ГНОЙНАЯ		
Семантический объект	СЕРЬЕЗНОЕ ОСЛОЖНЕНИЕ; АНГИНА		
Семантический объект	ПОСТАНГИНАЛЬНЫЙ СЕПСИС		

Рис. 4. Работа программного пакета PullEnti по извлечению объектов и их признаков из текста с описанием осложнения гнойной ангины

### Разбиение (кластеризация) входной объектно-признаковой матрицы

На рис. 5 приведен фрагмент формального контекста, сформированного на основе текстов медицинской направленности при помощи программного пакета Pullenti и предложенного нами шаблона. Анализируя фрагмент, можно заметить, что признаки, стоящие в парах с объектами, обозначающими болезни, не используются в паре с объектами, обозначающими органы, и наоборот. Например, можно сказать «осложнение гайморита» или «осложнение гриппа», но нельзя сказать «осложнение мозга», и наоборот, можно сказать «отек пазухи» или «отек мозга», но нельзя сказать «отек гриппа». Благодаря наличию таких отличительных признаков мы можем выделить во входной объектно-признаковой матрице группы (кластеры) родственных объектов. Если теперь в каждом кластере убрать нулевые признаки и оставить только те, которые присутствуют хотя бы у одного члена кластера, то мы получим на основе этих кластеров новые менее разреженные объектно-признаковые матрицы меньших размеров.

На рис. 6 показаны результаты иерархической кластеризации объектно-признаковой матрицы, построенной на основе рассмотренного выше формального контекста. Первый кластер, как и ожидалось, содержит названия болезней, а второй – названия органов. Внутри второго кластера в отдельный подкластер выделены объекты «гайморова пазуха» и «печень», которые в формальном контексте встречались вместе с признаком «лечение». Лечение – это пример гибридного признака, который употребляется как с болезнями, так и с органами (сравните: «лечение гепатита» и «лечение печени»).

В данной работе в качестве метода кластеризации мы используем аггломеративную кластеризацию по Уорду. При использовании метода Уорда на каждом шаге алгоритма происходит слияние кластеров, которое приводит к минимальному увеличению дисперсии внутри объединенного кластера:

$$\Delta(A, B) = \sum_{i \in A \cup B} \|\bar{x}_i - \bar{m}_{A \cup B}\|^2 - \sum_{i \in A} \|\bar{x}_i - \bar{m}_A\|^2 - \sum_{i \in B} \|\bar{x}_i - \bar{m}_B\|^2,$$

где  $A, B$  – это кластеры до слияния,  $\bar{x}_i$  и  $m$  – объекты кластеров и их центры,  $\Delta(A, B)$  – целевая функция, которая называется ценой слияния.

	профилактика	лечение	воспаление	отек	увеличение	осложнение
аппендицит	1	1	0	0	0	1
аппендикс	0	0	1	0	1	0
гайморит	1	1	0	0	0	1
гайморова пазуха	0	1	1	1	1	0
печень	0	1	1	1	1	0
гепатит	1	1	0	0	0	1
мозг	0	0	1	1	1	0
грипп	1	1	0	0	0	1

Рис. 5. Фрагмент формального контекста

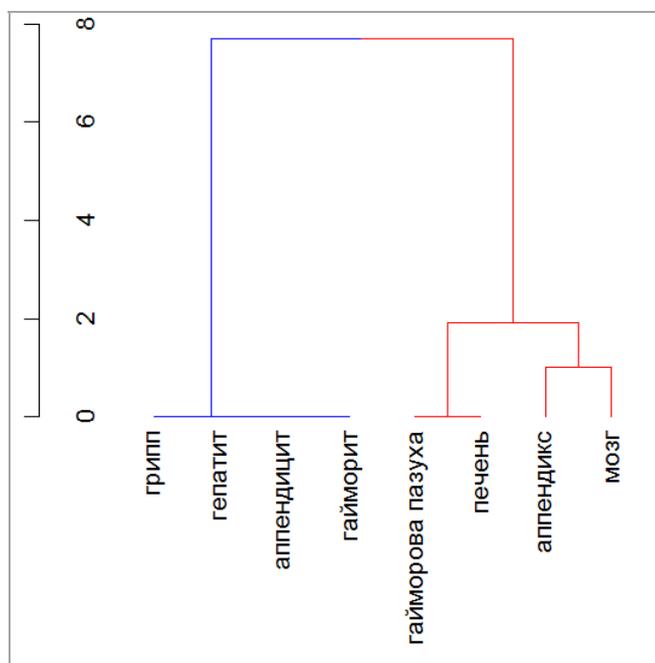


Рис. 6. Кластеры объектов, выделенные на основе иерархической кластеризации

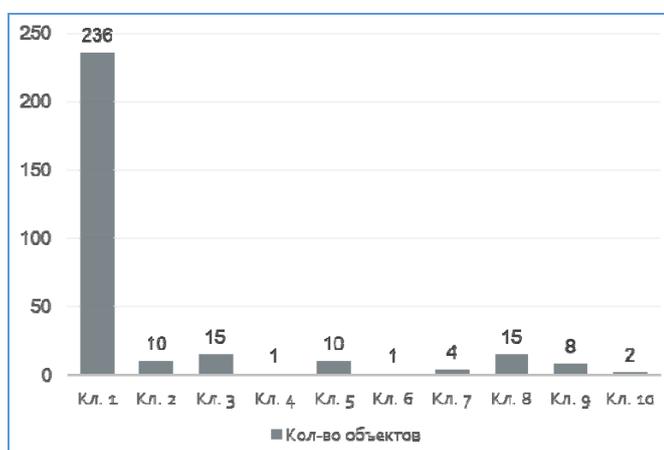


Рис. 7. Пример распределения объектов по кластерам

На рис. 7 представлена типичная диаграмма распределения объектов по кластерам. Типичной ее делает наличие одного гигантского кластера на фоне множества мелких. Эксперименты показывают, что мелкие кластеры, как правило, достаточно устойчивы и действительно состоят из однородных объектов, в то время как гигантский кластер, как правило, представляет собой собрание всех «неясных» объектов, не отнесенных ни к одному из устойчивых кластеров.

### Восстановление пропусков в объектно-признаковых матрицах

Как указывал один из основоположников статистической обработки естественного языка Дж. Ципф, разреженность – это извечная проблема при работе с текстовыми корпусами [8]. Разреженность негативно влияет на результаты кластеризации, искажает их, так что определенная часть терминов в результате такой кластеризации оказывается не в своей категории. Для борьбы с разреженностью данных в этом проекте мы используем подходы на основе слабого машинного обучения. Термин «слабое обучение» возник сравнительно недавно для обозначения методов, в которых обучение, называемое бутстраппингом (англ. bootstrapping – применительно к информационным технологиям «самозагрузка»), ведется на небольшом множестве позитивных примеров [9; 10].

Например, системе, извлекающей из текстов названия болезней, может потребоваться предоставить небольшое количество примеров названий. После этого система ищет в текстах вхождения этих названий, и пытается определить некоторые общие контексты их появления (шаблоны). Затем система по выявленным шаблонам пытается отыскать новые названия болезней, например, в сети Интернет. Процесс обучения является итеративным, что позволяет обнаруживать всё новые шаблоны и новые экземпляры названий. В итоге многократные повторения этого процесса позволят собрать большое количество названий болезней и большое количество шаблонов их употребления в текстах.

В нашем случае в качестве обучающих примеров мы используем самые «сильные» (близко расположенные к центру) объекты кластера и определяем контексты их появления в текстах коллекции. В обнаруженном контексте мы заменяем «сильный» объект кластера на «слабый», например на объект, категория которого вызывает у нас сомнение, и пытаемся отыскать вхождения «слабого» объекта в рамках заданного контекста в сети Интернет. Для этой цели в нашей системе разработан поисковый робот, который осуществляет просмотр тематических веб-страниц и ищет вхождения слабого объекта. Мы формируем задания этому роботу по сбору веб-страниц из сети Интернет, описывая схемы, позволяющие произвести правильный парсинг страниц. Описание схем производится в полуавтоматическом режиме: основная структура веб-страницы загружается в систему, и необходимо настроить только некоторые параметры разметки страниц, например указать контейнер основного текста, последовательность страниц для обхода и т. д.

Таким образом, вслед за авторами работы [11] мы используем Интернет как сверхбольшой и сверхценный корпус текстов и оцениваем вероятность использования слабого объекта в контексте, опираясь на количество откликов (hits) поисковой машины на запрос:

$$\text{Score}(\text{Weak Object}, \text{Context}) = \frac{\text{hits}(\text{Context with Weak Object})}{\text{hits}(\text{Weak Object})}.$$

В табл. 2 представлены кластеры, которые получились после уменьшения разреженности в исходном наборе данных, состоящем из 119 строк (объектов) и 616 столбцов (признаков). Жирным шрифтом выделены слова, которые, с точки зрения эксперта, являются релевантными основному содержанию кластера, подчеркнуты слова, ошибочно включенные в данный кластер. Не все кластеры, с точки зрения эксперта, объединяют однородные объекты, тем не менее эти результаты демонстрируют принципиальную возможность использования предлагаемого подхода. Ошибочное появление слов в том или ином кластере можно объяснить, во-первых, разреженностью полученной матрицы, во-вторых, неполнотой исходных корпу-

сов текстов, в-третьих, тем, что решение задачи кластеризации неоднозначно и существенно зависит от выбора метода кластеризации и критерия оценки качества.

Таблица 2

Кластеры, полученные после уменьшения разреженности

Номер кластера и тематика	Термины
1 – симптомы, признаки, проявления	" <u>активность</u> " "действие" "образование" " <u>период</u> " "повреждение" "признак" "проявление" "развитие" "симптом" "состояние" " <u>фаза</u> " " <u>функция</u> "
2 – действия общего характера, методы	"борьба" "введение" "возникновение" "выделение" "выявление" "использование" "контакт" "передача" "повышение" "прием" "применение" "проведение" "работа" "распространение" "снижение" "течение" "увеличение"
3 – препараты и воздействующие факторы	"антиген" "вирус" "возбудитель" "кровь" "препарат" "фактор"
4 – методы медицинского контроля и лечения	" <u>диагностик</u> " "диагностика" "исследование" "контроль" "лечение" "определение" "профилактика" " <u>фон</u> "
5 – группирующие слова	"вид" "группа" "ряд" "система" "тип" "форма" "характер"
6 – структурные элементы организма	" <u>боль</u> " "клетка" "орган" "ткань" " <u>человек</u> " "эритроцит" "лейкоцит" "бактерия"
7 – характеристики, проявления симптомов и болезней	"выраженность" "гиперемия" " <u>день</u> " " <u>дифтерия</u> " "длительность" " <u>зависимость</u> " "заражение" "исчезновение" "картина" "конец" "момент" "наличие" "нарастание" "отсутствие" "патогенез" "подозрение" "попадание" "появление" "причина" "продолжительность" "проникновение" "случай" " <u>смерть</u> " "срок" "стадия" "тенденция" "тяжесть" "уменьшение" " <u>употребление</u> "
8 – количественные, измеряемые характеристики	"возможность" "время" "интенсивность" "количество" "концентрация" " <u>локализация</u> " " <u>место</u> " " <u>особенность</u> " "показатель" " <u>результат</u> " "содержание" "способность" "степень" "уровень" " <u>частота</u> " "число" "этиология"
9 – слова, обозначающие заболевание	"болезнь" "заболевание" "изменение" "нарушение" "поражение" "процесс"
10 – области	" <u>кожа</u> " " <u>лицо</u> " "область" " <u>организм</u> " "основа" "очаг" "поверхность" "структура" "часть"

На рис. 8, 9 представлены результаты иерархической кластеризации объектов, примененной соответственно к кластерам 2 и 8. Как видно из рисунка, объекты внутри кластера можно, в свою очередь, также разделить на семантические подгруппы в соответствии с дифференцирующими признаками.

### Извлечение формальных понятий и их иерархий

На рис. 10 представлена решетка понятий, сформированная на основе кластера 6. Как видно из этой решетки, лейкоциты, эритроциты и бактерии – это клетки. Причем, отличие эритроцитов и лейкоцитов от бактерий заключается в том, что они могут оседать, у них есть ядро, и они подвержены анизоцитозу, а отличие бактерий в том, что они могут быть патогенными. Клетка, орган и ткань согласно диаграмме Хассе – это понятия одного уровня. Хотя на самом деле и органы, и ткани состоят из клеток, поэтому клетка должна стоять уровнем ниже, предлагаемый подход позволяет выделять только отношения иерархии “is-a” и не позволяет выделять отношения вида “part-of”.

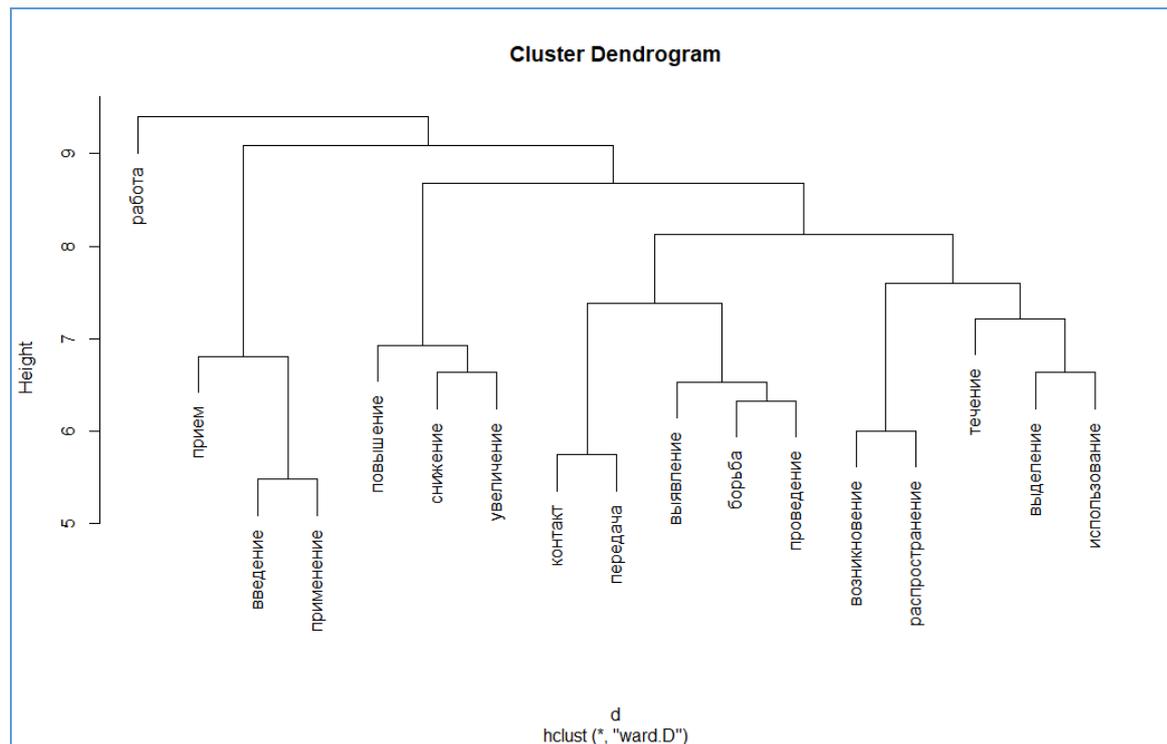


Рис. 8. Результаты повторной кластеризации, примененной к кластеру объектов, обозначающих действие

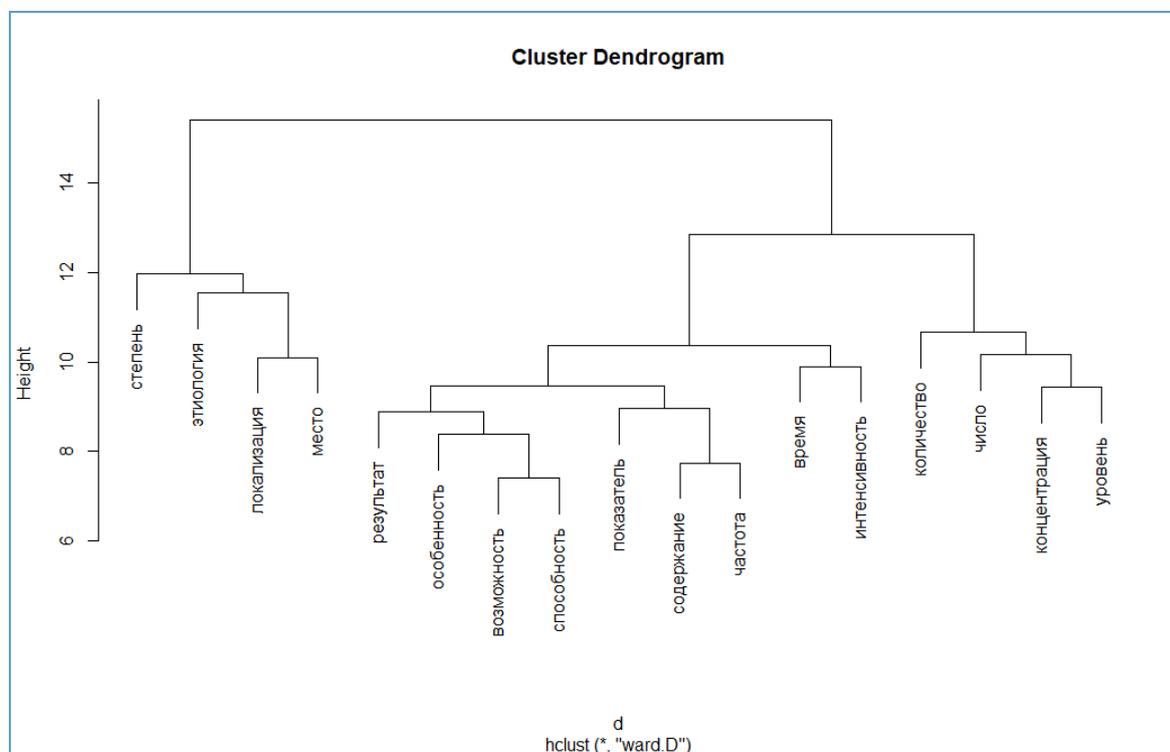


Рис. 9. Результаты повторной кластеризации, примененной к кластеру объектов, обозначающих количественные, измеряемые характеристики

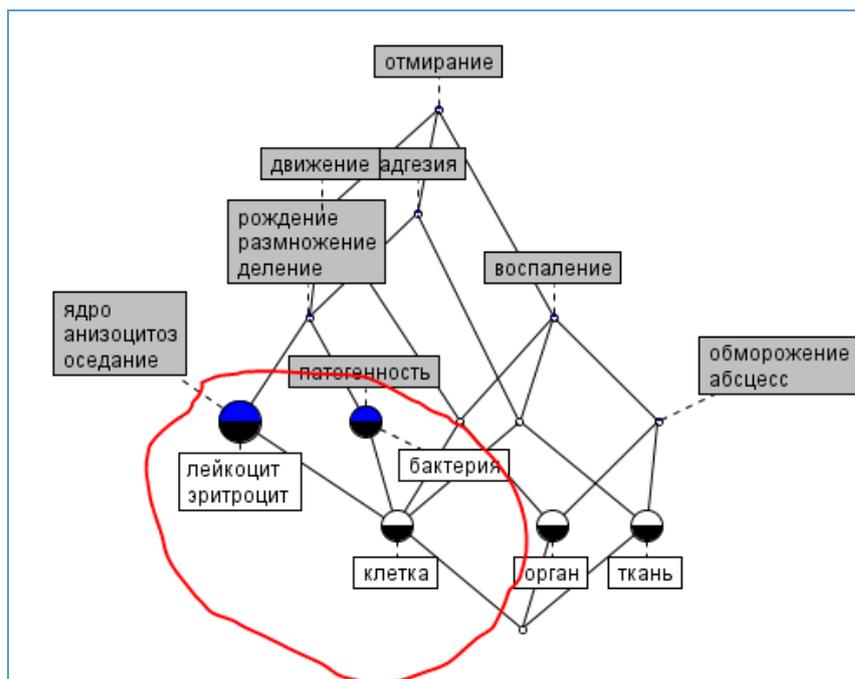


Рис. 10. Решетки понятий для кластера 6

## Заключение

Предложенный в статье новый способ извлечения понятий из текстов предметной области, основанный на комбинации анализа формальных понятий и бутстрап-технологии информационного поиска с использованием данных сети Интернет продемонстрировал высокую эффективность при работе с корпусами текстов, в которых специальная терминология сильно разрежена. Таким образом, при правильном выборе исходных шаблонов поиска бутстраппинг, основанный на использовании открытых ресурсов Интернета как ценных источников знаний, превращается в эффективный инструмент поддержки концептуального моделирования.

В своих будущих исследованиях мы планируем апробировать данный подход на текстах других тематик, а также попробовать формировать формальный контекст на основе новых шаблонов, например шаблонов «существительное + существительное», но с использованием предлогов. По-прежнему остается актуальным вопрос определения количества кластеров: пока мы решаем эту задачу эмпирическим путем.

## Список литературы

1. *Игнатов Д. И.* Анализ формальных понятий: от теории к практике // Доклады всероссийской научной конференции АИСТ'12 «Анализ изображений, сетей и текстов». 16–18 марта 2012 г. Национальный открытый университет «ИНТУИТ». Екатеринбург, 2012. С. 3–15.
2. *Ganter B., Wille R.* Formal concept analysis: mathematical foundations. Springer Science & Business Media, 2012. 284 p.
3. *Кузнецов О. С., Обьедков С. А.* Алгоритмы построения множества всех понятий формального контекста и его диаграммы Хассе // Изв. РАН. Теория и системы управления. 2001. № 1. С. 120–129.
4. *Hwang Y. S., Finch A., Sasaki Y.* Improving statistical machine translation using shallow linguistic knowledge // Computer Speech & Language. 2007. Vol. 21. No. 2. P. 350–372.
5. *Crysmann B. et al.* An integrated architecture for shallow and deep processing // Proceedings of the 40<sup>th</sup> annual meeting on association for computational linguistics. Association for Computational Linguistics, 2002. P. 441–448.

6. PullEnti / К. И. Кузнецов. 2013. URL: <http://www.pullenti.ru/Default.aspx> (дата обращения 07.01.2018).
7. Kozerenko E., Kuznetsov K., Morozova Yu., Romanov D. Semantic Proximity Establishment in the Tasks of Knowledge Extraction and Named Entities Recognition // Proc. of the 2017 Int. Conf. on Artificial Intelligence. 2017. P. 339–344.
8. Zipf G. Selective Studies and the Principle of Relative Frequency in Language. Cambridge, 1932.
9. Nadeau D., Turney P., Matwin S. Unsupervised Named-Entity Recognition: Generating Gazetteers and Resolving Ambiguity // Advances in Artificial Intelligence. 2006. P. 266–277.
10. Schapire R. E. The boosting approach to machine learning: An overview // Nonlinear estimation and classification. New York: Springer, 2003. P. 149–171.
11. Vieira K. et al. Finding seeds to bootstrap focused crawlers // World Wide Web. 2016. Vol. 19. No. 3. P. 449–474.

Материал поступил в редколлегию 24.08.2018

**A. B. Nugumanova, E. M. Bayburin, M. E. Mansurova, V. B. Barakhnin**

<sup>1</sup> Sarsen Amanzholov East-Kazakhstan State University  
34 Tritsatsoy Gvardeiskoy Divizii Str., Ust-Kamenogorsk, 070002, Republic of Kazakhstan

<sup>2</sup> Al-Farabi Kazakh National University  
71 al-Farabi ave., Almaty, 050040, Republic of Kazakhstan

<sup>3</sup> Institute of Computational Technologies SB RAS  
6 Academician Lavrentiev Ave., Novosibirsk, 630090, Russian Federation

<sup>4</sup> Novosibirsk State University  
1 Pirogov St., Novosibirsk, 630090, Russian Federation

{anugumanova, ebaiburin}@ykgu.kz, mansurova.madina@gmail.com, bar@ict.nsc.ru

## **AUTOMATIC EXTRACTION OF FORMAL LATTICES FROM MEDICAL TEXTS BASED ON THE COMBINATION OF THE FORMAL CONCEPT ANALYSIS AND BOOTSTRAPPING TECHNOLOGIES**

The article considers a new way of concept extraction from the subject domain texts based on combination of formal concept analysis and bootstrap technology of information retrieval. Formal concept analysis is a powerful way of automatically deriving the domain concepts, but it is designed for high quality input data, without missing and inaccuracies. Obtaining such datasets directly from texts is difficult because of the strong sparsity of the text corpora. Accordingly, it seems promising to improve the quality of input data with bootstrapping, a technology that provides an intelligent search for fragmented information on the Internet. In this paper, we show the steps of implementing the way of automatically concept extraction from medical texts based on the filling of blanks in highly sparse matrices of the joint occurrence of terms. The input data for formal concept analysis is represented in the form of an object-feature table that reflects the distribution of attributes over the objects of the domain. The purpose of this paper is to show that with proper selection of initial search patterns, bootstrapping based on the use of open Internet resources as valuable sources of knowledge, turns into an effective tool for supporting conceptual modeling.

*Keywords:* formal concept analysis, bootstrapping, information extraction, surface linguistic analysis, information retrieval.

### **References**

1. Ignatov D. I. Analysis of formal concepts: from theory to practice. *Reports of the Russian scientific conference AINT'12 "Analysis of images, networks and texts."* March 16–18, 2012. National Open University "INTUIT". Ekaterinburg, 2012, p. 3–15.

2. Ganter B., Wille R. Formal concept analysis: mathematical foundations. Springer Science & Business Media, 2012, 284 p.
3. Kuznetsov O. S., Obedkov S. A. Algorithms for constructing the set of all concepts of a formal context and its Hasse diagram. *Izvestiya RAN. Theory and control systems*, 2001, no. 1, p. 120–129. (in Russ.)
4. Hwang Y. S., Finch A., Sasaki Y. Improving the statistical translation of shallow linguistic knowledge. *Computer Speech & Language*, 2007, vol. 21, no. 2, p. 350–372.
5. Crysmann B. et al. An integrated architecture for shallow and deep processing. *Proceedings of the 40<sup>th</sup> annual meeting on association for computational linguistics. Association for Computational Linguistics*, 2002, p. 441–448.
6. PullEnti / K. I. Kuznetsov. 2013. URL: <http://www.pullenti.ru/Default.aspx> (date of access 07.01.2018).
7. Kozerenko E., Kuznetsov K., Morozova Yu., Romanov D. Semantic Proximity Establishment in the Tasks of Knowledge Extraction and Named Entities Recognition. *Proc. of the 2017 Int. Conf. on Artificial Intelligence*, 2017, p. 339–344.
8. Zipf G. *Selective Studies and the Principle of Relative Frequency in Language*. Cambridge, 1932.
9. Nadeau D., Turney P., Matwin S. Unsupervised Named-Entity Recognition: Generating Gazetteers and Resolving Ambiguity. *Advances in Artificial Intelligence*, 2006, p. 266–277.
10. Schapire R. E. The boosting approach to machine learning: An overview. *Nonlinear estimation and classification*. New York, Springer, 2003, p. 149–171.
11. *Vieira K. et al.* Finding seeds to bootstrap focused crawlers. *World Wide Web*, 2016, vol. 19, no. 3, p. 449–474.

Received 24.08.2018

*For citation:*

Nugumanova A. B., Bayburin E. M., Mansurova M. E., Barakhnin V. B. Automatic Extraction of Formal Lattices from Medical Texts Based on The Combination of the Formal Concept Analysis and Bootstrapping Technologies. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 4, p. 140–152. (in Russ.) DOI 10.25205/1818-7900-2018-16-4-140-152

**П. А. Титов**

*Институт вычислительной математики  
и математической геофизики СО РАН  
пр. Академика Лаврентьева, 6, Новосибирск, 630090, Россия*

*tapawel@gmail.com*

## **МОДЕЛИРОВАНИЕ УПРУГИХ ВОЛН В СРЕДАХ СО СЛОЖНОЙ ТОПОГРАФИЕЙ СВОБОДНОЙ ПОВЕРХНОСТИ \***

Моделирование – неотъемлемая часть исследования процессов распространения сейсмических волн в различных средах. Широко используемый способ разбиения области на прямоугольные ячейки обладает недостатком: при сложном строении свободной поверхности (например, гора) возникают эффекты при отражении волны от этой поверхности, связанные с тем, что граница аппроксимируется ступенчатой функцией для численного решения задачи. В данной работе предложен иной подход к дискретизации области: построение криволинейной сетки, хорошо согласующейся с геометрией свободной поверхности. Предложен алгоритм численного моделирования на основе численного решения линейной 2D-системы теории упругости, записанной в смещениях, с использованием криволинейной сетки и пошагового метода Лагерра. Представлены результаты моделирования. Также были реализованы две параллельные версии алгоритма, проведены расчеты на различных многоядерных системах (на классической многопроцессорной архитектуре, а также архитектуре с использованием сопроцессоров Intel Xeon Phi – "РСК ПетаСтрим"). Представлены сравнительные тесты ускорений на разных архитектурах, а также описаны особенности распараллеливания алгоритма.

*Ключевые слова:* теория упругости, упругие волны, криволинейная поверхность, криволинейные сетки, метод Лагерра, Intel Xeon Phi, PetaStream.

### **Введение**

Эффективным инструментом исследования процессов распространения волн в различных моделях сложно построенных сред является математическое моделирование. Одним из известных методов численного моделирования распространения упругих волн в случае сложно построенных 2D-сред является разностный метод. При использовании данного метода неизменно сталкиваются с проблемой возникновения дифракционных волн при отражении волны от свободной поверхности. Связано это с тем, что для применения данного метода необходимо разбиение физической области на квадратные или прямоугольные ячейки, вследствие чего свободная поверхность заменяется ступенчатой функцией. Пример использования разностного метода приведен на рис. 1 и 2.

Подобные дифракционные волны вносят погрешность в решение. Таким образом, была поставлена задача об избавлении от дифракционных волн. Для этого в работе предлагается

---

\* Работа выполнена в рамках государственного задания ИВМиМГ СО РАН (проект № 0315-2016-0009), а также при поддержке грантов РФФИ № 16-07-00434 и № 16-01-00455 А.

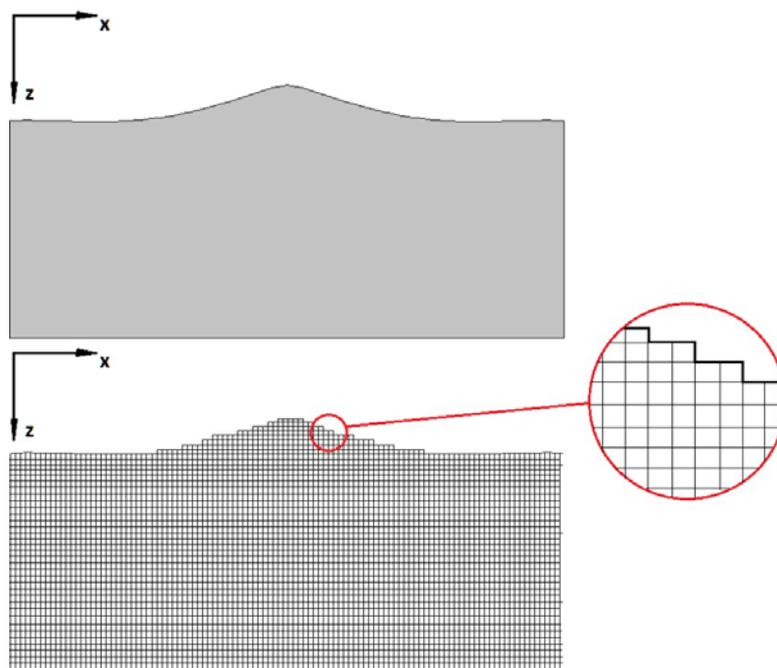


Рис. 1. Область и соответствующее ей разбиение на квадратные ячейки

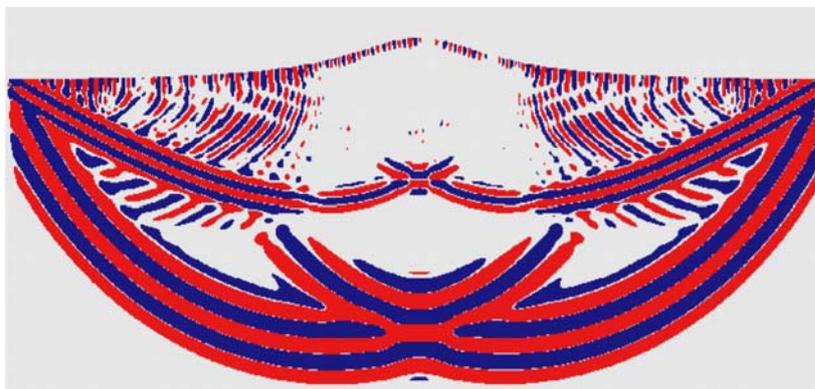


Рис. 2. При отражении от свободной поверхности видны дифракционные волны, идущие вслед за отраженными волнами (возле поверхности) (Рисунок предоставлен Д. А. Караевым, ИВМиМГ СО РАН)

использовать криволинейные сетки. В последнее время все большую популярность приобретает моделирование с использованием криволинейных сеток. Это гибкий инструментари, нашедший применение в численном решении сингулярно-возмущенных задач, моделировании потоков плазмы в камере Токмака и др. Способы построения сеток, а также их применение для решения упомянутых задач приведены в [1].

Данная работа особенно важна для дальнейшей разработки 3D-алгоритма, нацеленного на развитие методов исследования областей, характерной особенностью которых является сложное строение топографии свободной поверхности. К таковым можно отнести магматические вулканы, а также территории арктического шельфа и Сибири (Баженовская свита), богатые залежами углеводородов. Существующие методы работы с такого рода задачами не позволяют получить желаемую точность моделирования ввиду большой трудозатратности и особенностей применяемых методов, таких как метод конечных элементов (МКЭ).

### Построение криволинейной сетки

Существует много методов построения криволинейной сетки. Некоторые из них описаны в работах [1–12]. Криволинейная сетка в «физической» области получается в результате взаимно-однозначного отображения равномерной сетки «расчетной» области. Таким образом, необходимо установить взаимно-однозначное соответствие между «физической» и «расчетной» областью. «Физическая» область находится в пространстве переменных  $(x, y)$ , а «расчетная» – в пространстве переменных  $(q, r)$ . Для этого был использован метод трансфинитной интерполяции [1, с. 53]. Автором были подобраны управляющие функции, позволяющие определять свойства криволинейной сетки.

В данной работе важным моментом построения сетки является ее ортогональность к свободной поверхности, что накладывает ограничение на форму свободных поверхностей, с которыми можно работать при помощи данного метода построения сетки, а именно: поверхность должна быть гладкой функцией  $y = G(x)$  первого порядка на промежутке, на котором она построена для данной области, а также для любых двух разных  $y_1$  и  $y_2$  не существует  $x$ , такого что  $y_1 = G(x)$  и  $y_2 = G(x)$ . Пример на рис. 3.

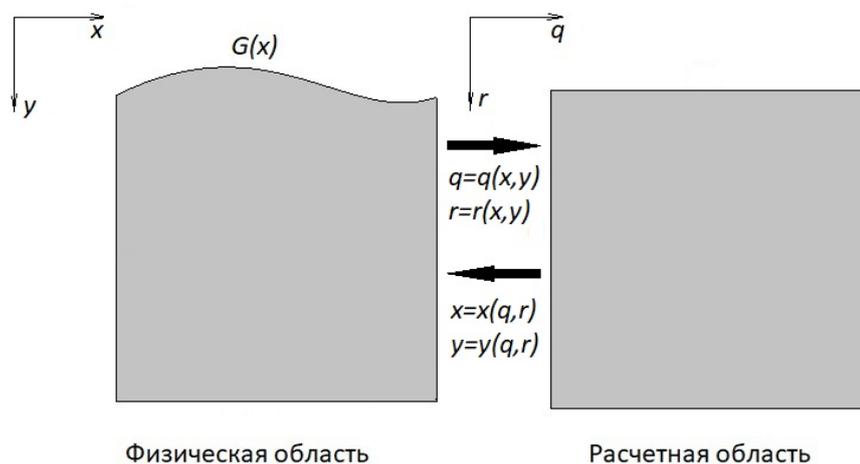


Рис. 3. Пример областей

Таким образом, задачу, поставленную в переменных  $(x, y)$ , можно переписать в переменных  $(q, r)$ .

### Постановка и преобразование задачи

Моделирование проводится на основе численного решения полной линейной системы теории упругости, записанной в смещениях. В «физической» области в переменных  $(x, y)$  линейная система уравнений в смещениях принимает вид

$$\begin{aligned} \rho \frac{\partial^2 u}{\partial t^2} &= \frac{\partial}{\partial x} \left( (2\mu + \lambda) \frac{\partial u}{\partial x} + \lambda \frac{\partial w}{\partial y} \right) + \frac{\partial}{\partial y} \left( \mu \left( \frac{\partial w}{\partial x} + \frac{\partial u}{\partial y} \right) \right) + F_x = 0, \\ \rho \frac{\partial^2 w}{\partial t^2} &= \frac{\partial}{\partial x} \left( \mu \left( \frac{\partial w}{\partial x} + \frac{\partial u}{\partial y} \right) \right) + \frac{\partial}{\partial y} \left( (2\mu + \lambda) \frac{\partial u}{\partial x} + \lambda \frac{\partial w}{\partial y} \right) + F_y = 0. \end{aligned} \tag{1}$$

Граничные условия на  $\partial\Gamma$  и начальные условия:

$$u|_{\partial\Gamma} = w|_{\partial\Gamma} = 0, \quad \frac{\partial w}{\partial t} = \frac{\partial u}{\partial t} = 0. \quad (2)$$

Условия на свободной поверхности  $\partial S$ :

$$\begin{aligned} n_x \left( (2\mu + \lambda) \frac{\partial u}{\partial x} + \lambda \frac{\partial w}{\partial y} \right) + n_y \left( \mu \left( \frac{\partial w}{\partial x} + \frac{\partial u}{\partial y} \right) \right) &= 0, \\ n_x \left( \mu \left( \frac{\partial w}{\partial x} + \frac{\partial u}{\partial y} \right) \right) + n_y \left( (2\mu + \lambda) \frac{\partial u}{\partial x} + \lambda \frac{\partial w}{\partial y} \right) &= 0, \end{aligned} \quad (3)$$

где  $u(x, y)$ ,  $w(x, y)$  – смещения вдоль координатных осей  $OX$  и  $OY$ ;  $\lambda(x, y)$ ,  $\mu(x, y)$  – параметры Ламэ;  $\rho(x, y)$  – плотность;  $\partial\Gamma$  – граница области, не включающей свободную поверхность;  $\partial S$  – свободная поверхность;  $[n_x, n_y]^T$  – единичный вектор нормали к свободной поверхности  $\partial S$ ;  $F_x, F_y$  – правая часть, отвечает за работу источника возмущений в среде. В данной работе используется источник типа «центр давления», работает в течение 2-х секунд.

$$\begin{aligned} F_x &= \begin{cases} \frac{\partial \delta(x - x_0, y - y_0)}{\partial x} (\sin(\pi(t-1)) + 0,8 \sin(2\pi(t-1)) + 0,2 \sin(3\pi(t-1))), & 0 \leq t \leq 2, \\ 0, & t \geq 2, \end{cases} \\ F_y &= \begin{cases} \frac{\partial \delta(x - x_0, y - y_0)}{\partial y} (\sin(\pi(t-1)) + 0,8 \sin(2\pi(t-1)) + 0,2 \sin(3\pi(t-1))), & 0 \leq t \leq 2, \\ 0, & t \geq 2. \end{cases} \end{aligned}$$

Аналогично работе [13] после преобразования системы (1)–(3) получаем задачу вида

$$\begin{aligned} J\rho \frac{\partial^2 u}{\partial t^2} &= \frac{\partial}{\partial q} \left[ Jq_x \left[ (2\mu + \lambda)(q_x \partial_q + r_x \partial_r)u + \lambda(q_y \partial_q + r_y \partial_r)w \right] + \right. \\ &\quad \left. + Jq_y \left[ \mu \left( (q_x \partial_q + r_x \partial_r)w + (q_y \partial_q + r_y \partial_r)u \right) \right] \right] + \\ &\quad + \frac{\partial}{\partial r} \left[ Jr_x \left[ (2\mu + \lambda)(q_x \partial_q + r_x \partial_r)u + \lambda(q_y \partial_q + r_y \partial_r)w \right] + \right. \\ &\quad \left. + Jr_y \left[ \mu \left( (q_x \partial_q + r_x \partial_r)w + (q_y \partial_q + r_y \partial_r)u \right) \right] \right] + F_x = B_1(u, w) + F_x, \end{aligned} \quad (4)$$

$$\begin{aligned} J\rho \frac{\partial^2 w}{\partial t^2} &= \frac{\partial}{\partial q} \left[ Jq_x \left[ \mu \left( (q_x \partial_q + r_x \partial_r)w + (q_y \partial_q + r_y \partial_r)u \right) \right] + \right. \\ &\quad \left. + Jq_y \left[ (2\mu + \lambda)(q_x \partial_q + r_x \partial_r)u + \lambda(q_y \partial_q + r_y \partial_r)w \right] \right] + \\ &\quad + \frac{\partial}{\partial r} \left[ Jr_x \left[ \mu \left( (q_x \partial_q + r_x \partial_r)w + (q_y \partial_q + r_y \partial_r)u \right) \right] + \right. \\ &\quad \left. + Jr_y \left[ (2\mu + \lambda)(q_x \partial_q + r_x \partial_r)u + \lambda(q_y \partial_q + r_y \partial_r)w \right] \right] + F_y = B_2(u, w) + F_y. \end{aligned}$$

Начальные и граничные условия:

$$u|_{\partial\Gamma} = w|_{\partial\Gamma} = 0, \quad \frac{\partial w}{\partial t} = \frac{\partial u}{\partial t} = 0. \quad (5)$$

На свободной границе:

$$\begin{aligned} \hat{r}_x \left[ (2\mu + \lambda)(q_x \partial_q + r_x \partial_r)u + \lambda(q_y \partial_q + r_y \partial_r)w \right] + \\ + \hat{r}_y \left[ \mu \left( (q_x \partial_q + r_x \partial_r)w + (q_y \partial_q + r_y \partial_r)u \right) \right] = 0, \\ \hat{r}_x \left[ \mu \left( (q_x \partial_q + r_x \partial_r)w + (q_y \partial_q + r_y \partial_r)u \right) \right] + \\ + \hat{r}_y \left[ (2\mu + \lambda)(q_x \partial_q + r_x \partial_r)u + \lambda(q_y \partial_q + r_y \partial_r)w \right] = 0, \end{aligned} \tag{6}$$

где  $\hat{r}_x = \frac{r_x}{\sqrt{r_x^2 + r_y^2}}$ ,  $\hat{r}_y = \frac{r_y}{\sqrt{r_x^2 + r_y^2}}$ ,  $q_x = \frac{y_r}{J}$ ,  $r_x = -\frac{y_q}{J}$ ,  $q_y = -\frac{x_r}{J}$ ,  $r_y = \frac{x_q}{J}$ ,  $J = q_x r_y - q_y r_x$  – Якобиан преобразования.

### Пошаговый метод преобразований Лагерра по времени

Впервые данный метод был применен в работе [14]. В отличие от классических преобразований Фурье и Лапласа применение преобразования Лагерра приводит к решению задачи для системы дифференциальных уравнений, не зависящей от параметра разделения, роль которого в данном случае играет номер проекции Лагерра. В результате применения этого преобразования получается задача для проекций Лагерра с правой частью, рекуррентно зависящей от проекций меньших номеров. После применения конечно-разностных аппроксимаций по пространственным переменным, решение исходной задачи сводится к решению системы линейных алгебраических уравнений со многими правыми частями, для решения которых можно использовать различные современные подходы к решению линейных систем.

В работе рассматривается модификация метода, а именно преобразование Лагерра используется на последовательности конечных интервалов по времени. Полученное решение в конце одного временного отрезка используется в качестве начальных данных для решения задачи на следующем временном отрезке и т. д. Связано это с тем, что для получения решения с хорошей точностью через длительный промежуток времени необходимо очень большое количество проекций Лагерра, что критично скажется во время численной реализации. Подробно пошаговый метод преобразования Лагерра описан в работе [15].

При использовании данного подхода появляется необходимость выбора 4-х параметров: количество проекций Лагерра  $N$  масштабного множителя, необходимого для аппроксимации решения функциями Лагерра  $h$ , экспоненциального коэффициента весовой функции  $\mu$ , используемой для нахождения решения на конечном временном интервале, и длительности этого интервала  $\tau$ . Способ выбора этих параметров предложен в работе [15]. Мы же ограничимся лишь основными терминами и непосредственно видом задачи после примененного к ней преобразования Лагерра.

Полиномы Лагерра имеют вид

$$L_0(t) = 1, \tag{7}$$

$$L_1(t) = 1 - t, \tag{8}$$

$$L_{k+1}(t) = \frac{1}{k+1} \left( (2k+1-t)L_k(t) - kL_{k-1}(t) \right), \quad k > 1, \tag{9}$$

$$\dot{L}_k(t) = -\sum_{i=0}^{k-1} L_i(t), \quad k \geq 1. \tag{10}$$

Функции Лагерра

$$l_n(t) = e^{-\frac{t}{2}} L_n(t), \quad n \geq 0, \quad (11)$$

образуют полную ортонормированную систему в  $L_0(0, \infty)$ .

Преобразование

$$v_n = \int_0^\infty v(t) l_n(ht) dt \quad (12)$$

называется прямым преобразованием Лагерра, а  $v_n$  – проекциями Лагерра функции  $v(t)$ .

Обратное преобразование Лагерра:

$$v(t) = \sum_{n=0}^{\infty} v_n l_n(ht). \quad (13)$$

Идея метода состоит в получении приближенного значения

$$v(\tau) \approx \sum_{n=0}^N v_n l_n(\tau h) \quad (14)$$

и значения

$$V(\tau) = e^{\mu\tau} v(\tau), \quad \dot{V}(\tau) = \mu V(\tau) + e^{\mu\tau} \dot{v}(\tau) \quad (15)$$

выбираем в качестве начальных данных для решения задачи при  $t \geq \tau$ . Иначе говоря, теперь вместо нахождения неизвестной функции  $V(t)$  мы ищем значения проекций Лагерра для этой функции.

### Применение преобразования Лагерра к задаче (3)–(6)

Учитывая уравнения (7)–(15), задача (3)–(6) сводится к решению линейной системы уравнений:

$$\left\{ \begin{array}{l} \left[ \left( v + \frac{h}{2} \right)^2 + A \right] v_0 = \left( \frac{h^2}{2} + 2vh \right) v^0 + hv^1 + f_0, \\ \left[ \left( v + \frac{h}{2} \right)^2 + A \right] v_1 = -(h^2 + 2vh)v_0 + \left( \frac{3}{2}h^2 + 2vh \right) v^0 + hv^1 + f_1, \\ \dots \\ \left[ \left( v + \frac{h}{2} \right)^2 + A \right] (v_n - 2v_{n-1} + v_{n-2}) + 2vh(v_{n-1} - v_{n-2}) + h^2 v_{n-1} = f_n - 2f_{n-1} + f_{n-2}, \\ n \geq 2. \end{array} \right. \quad (16)$$

На  $\partial\Gamma$ :

$$v_n|_{\partial\Gamma} = 0, \quad n = 0, \dots, N. \quad (17)$$

На  $\partial S$ :

$$\hat{r}_x \left[ (2\mu + \lambda)(q_x \partial_q + r_x \partial_r) u_n + \lambda(q_y \partial_q + r_y \partial_r) w_n \right] + \quad (18)$$

$$\begin{aligned}
 & +\hat{r}_y \left[ \mu \left( (q_x \partial_q + r_x \partial_r) w_n + (q_y \partial_q + r_y \partial_r) u_n \right) \right] = 0, \\
 \hat{r}_x \left[ \mu \left( (q_x \partial_q + r_x \partial_r) w_n + (q_y \partial_q + r_y \partial_r) u_n \right) \right] + \\
 & +\hat{r}_y \left[ (2\mu + \lambda) (q_x \partial_q + r_x \partial_r) u_n + \lambda (q_y \partial_q + r_y \partial_r) w_n \right] = 0.
 \end{aligned}$$

В (16) вместо  $v_n$  подставляются  $u_n$ ,  $w_n$  –  $n$ -е проекции Лагерра для  $u$  и  $w$  соответственно.

Вместо оператора  $A$  подставляются  $\frac{B_1}{J\rho}$  для  $u_n$  и  $\frac{B_2}{J\rho}$  для  $w_n$  соответственно.  $N$  – число проекций Лагерра,  $v$ ,  $h$  – параметры преобразования Лагерра. Таким образом вместо изначальной системы из 2-х уравнений, записанной в смещениях, получаем систему из  $2(N+1)$  уравнений, записанную в проекциях Лагерра.

Далее для полученной системы строится разностный аналог на равномерной сетке размером  $M \times K$  ячеек. Разностная схема для  $B_1(u, w)$ ,  $B_2(u, w)$  описана в работе [13, уравнения (3.3), (3.4)], за исключением дискретных граничных условий на свободной поверхности  $\partial S$  и на  $\partial \Gamma$ , которые предложены автором данной работы. Упростив (18), получим на  $\partial S$ :

$$\begin{aligned}
 \alpha_1 \frac{\partial u_n}{\partial q} + \alpha_2 \frac{\partial u_n}{\partial r} + \alpha_3 \frac{\partial w_n}{\partial q} + \alpha_4 \frac{\partial w_n}{\partial r} &= 0, \\
 \beta_1 \frac{\partial u_n}{\partial q} + \beta_2 \frac{\partial u_n}{\partial r} + \beta_3 \frac{\partial w_n}{\partial q} + \beta_4 \frac{\partial w_n}{\partial r} &= 0.
 \end{aligned} \tag{19}$$

Из (19) получаем разностный аналог:

$$\begin{aligned}
 u_{n_{i,0}} = u_{n_{i,1}} - 1/2 \left( \frac{\beta_1 \alpha_4 - \alpha_1 \beta_4}{\alpha_2 \beta_4 - \beta_2 \alpha_4} \right)_{i,1/2} & \left( 3/2 (u_{n_{i+1,1}} - u_{n_{i-1,1}}) - 1/2 (u_{n_{i+1,2}} - u_{n_{i-1,2}}) \right) - \\
 & - 1/2 \left( \frac{\beta_3 \alpha_4 - \alpha_3 \beta_4}{\alpha_2 \beta_4 - \beta_2 \alpha_4} \right)_{i,1/2} \left( 3/2 (w_{n_{i+1,1}} - w_{n_{i-1,1}}) - 1/2 (w_{n_{i+1,2}} - w_{n_{i-1,2}}) \right), \\
 w_{n_{i,0}} = w_{n_{i,1}} - 1/2 \left( \frac{\beta_1 \alpha_2 - \alpha_1 \beta_2}{\alpha_4 \beta_2 - \beta_4 \alpha_2} \right)_{i,1/2} & \left( 3/2 (u_{n_{i+1,1}} - u_{n_{i-1,1}}) - 1/2 (u_{n_{i+1,2}} - u_{n_{i-1,2}}) \right) - \\
 & - 1/2 \left( \frac{\beta_3 \alpha_2 - \alpha_3 \beta_2}{\alpha_4 \beta_2 - \beta_4 \alpha_2} \right)_{i,1/2} \left( 3/2 (w_{n_{i+1,1}} - w_{n_{i-1,1}}) - 1/2 (w_{n_{i+1,2}} - w_{n_{i-1,2}}) \right).
 \end{aligned}$$

Здесь используются так называемые «фиктивные» узлы.

На  $\partial \Gamma$  из (17) получаем разностный аналог:

$$u_{n_{i,j}} = u_{n_{M,j}} = w_{n_{i,j}} = w_{n_{M,j}} = 0, \quad u_{n_{i,K}} = w_{n_{i,K}} = 0, \quad n = 0, \dots, N.$$

В итоге получаем разностный аналог системы (16)–(18) со вторым порядком аппроксимации по пространственным и временным переменным.

Для численного решения был задействован метод простой итерации. Подробно суть метода изложена в [16].

Выбор этого метода обусловлен тем, что система для численного решения из  $2(N+1)$  уравнений, переписанная в матричном виде, будет обладать матрицей с большим диагональным преобладанием, что обеспечит быструю сходимость метода.

Далее, на примере  $v_{n,j}$  (вместо него подставляются  $u_{n,j}$  и  $w_{n,j}$ ) опишем общий алгоритм действий.

0. На интервале  $[0, \tau]$  начальные данные выглядят так:  $V_{i,j}^0 = V_{i,j}(0)$ ,  $V_{i,j}^1 = \dot{V}_{i,j}(0)$ .

$V_{i,j}(0)$  и  $\dot{V}_{i,j}(0)$  берутся из начальных данных задачи.

1. Получив все значения  $v_{n,j}$  при помощи метода простой итерации, находим

$$v_{i,j}(\tau) = \sum_{n=0}^N v_{n,j} l_n(\tau h).$$

2. Отсюда, используя формулы  $\dot{v}(t) = c(t) = \sum_{n=0}^N c_n l_n(ht)$ ,  $c_0 = h \left( \frac{v_0}{2} - v(0) \right)$ ,  $c_n = c_{n-1} + \frac{h}{2}(v_n + v_{n-1})$ ,  $n = 1, 2, \dots$ , найдем  $\dot{v}_{i,j}$ .

3. Затем  $V_{i,j}(\tau) = e^{\mu\tau} v_{i,j}(\tau)$ ,  $\dot{V}_{i,j}(\tau) = \mu V_{i,j}(\tau) + e^{\mu\tau} \dot{v}_{i,j}(\tau)$  и, наконец, значения  $V_{i,j}^0 = V_{i,j}(\tau)$ ,  $V_{i,j}^1 = \dot{V}_{i,j}(\tau)$  берутся в качестве начальных данных для следующего временного интервала.

Присваиваем  $v_{i,j}^0 = V_{i,j}^0$ ,  $v_{i,j}^1 = -\mu V_{i,j}^0 + V_{i,j}^1$ .

4.  $v_{i,j}(\tau)$  и  $\dot{v}_{i,j}(\tau)$  будут начальными данными для следующего временного интервала. Далее, аналогично находим  $v_{i,j}(2\tau)$ ,  $\dot{v}_{i,j}(2\tau)$  и т. д.

### Программная параллельная реализация алгоритма

При помощи языка Fortran и пакета MPI было создано 2 варианта параллельной программы и проведены численные расчеты на кластере ССКЦ НКС-30Т (на узле G7 2 процессора Intel Xeon E5670 и 24 Гб оперативной памяти. <http://www.sscs.icmmg.nsc.ru/hardware.html>) с классической многопроцессорной архитектурой, а также на кластере МВС-10П МП с архитектурой РСК ПетаСтрим (8 модулей по 8 сопроцессоров Intel Xeon Phi 7120D, 61 ядро, 244 потока, RAM 16 Gb DDR5. На каждом модуле также стоит процессор Intel Xeon E5-2667 v2. Твердотельные накопители Intel SSD DC S3500. <http://www.jssc.ru/resources/hpc/>). Декомпозиция области для реализации на CPU и на Intel Xeon Phi представлена на рис. 4.

Особенностью архитектуры РСК ПетаСтрим является то, что пользователю доступны для работы только сопроцессоры Intel Xeon Phi. Между собой они обмениваются при помощи MPI-пересылок, а внутри каждого Phi также осуществляется распараллеливание по потокам при помощи OpenMP. Подробнее с особенностями программирования для Intel Xeon Phi можно ознакомиться в работах [17–21]. Далее будут приведены сравнительные тесты ускорений на примере конкретной модели среды.

Для решения системы линейных уравнений используется метод простых итераций, что позволяет «прореживать» обмены между процессами, так как это не влияет на сам факт сходимости итерационного процесса, а только на ее скорость. В ходе расчетов для модели, описанной далее, выяснилось, что относительно временных затрат на расчеты выгоднее сделать суммарно вдвое большее число итераций и вдвое меньшее число обменов. Другими словами, соседние процессы обмениваются раз в четыре итерации. Это ускорило работу программы в среднем в 1,5 раза, при этом точность решения пострадала незначительно – в тестовых расчетах расхождение решений по амплитуде, полученных с прореживанием обменов и без прореживания, составило порядка 0,001 %.

### Математическая модель

Область – однородная среда с горой на поверхности (рис. 5).

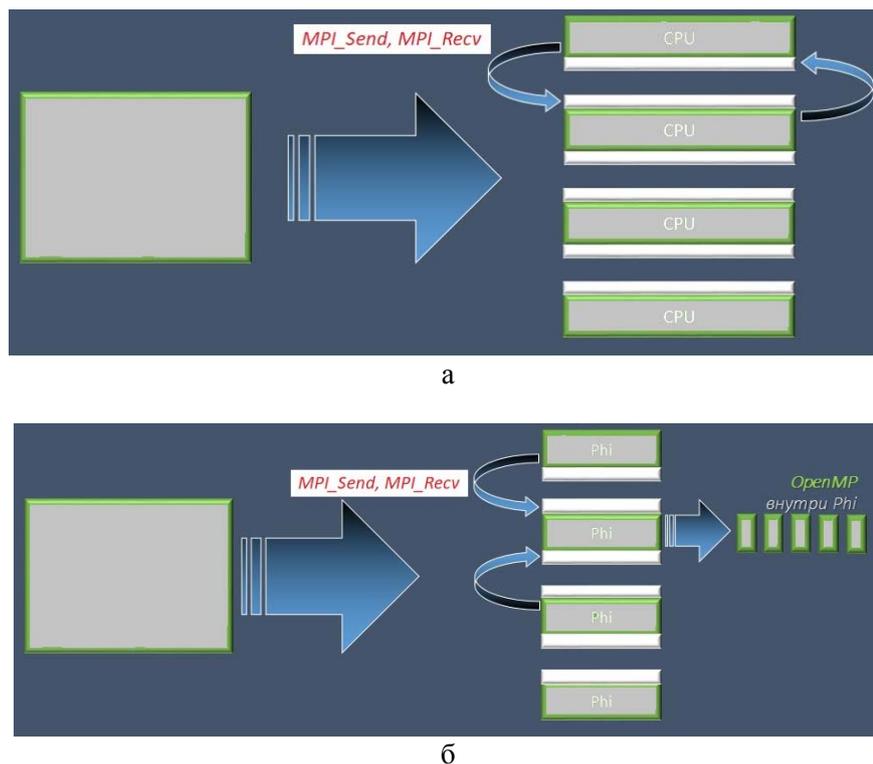


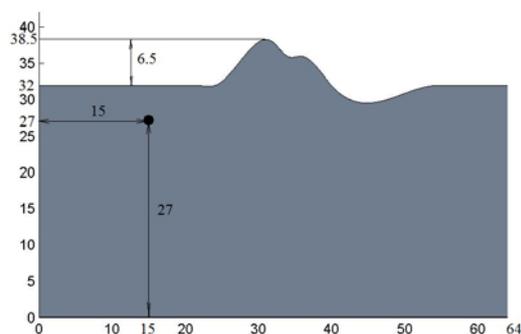
Рис. 4. Декомпозиция расчетной области для реализации на CPU (а) и Intel Xeon Phi (б)

Параметры среды :  $\rho = 1, \lambda = 0,5, \mu = 0,25$ . На рис. 6 размеры области даны в километрах. Соответствующая ей «расчетная» область – прямоугольник размером  $64 \times 32$  км, дискретизация области –  $5120 \times 2060$  ячеек. Источник типа «центр давления» работает в течение 2-х секунд. В уравнениях он представлен двумя компонентами  $(F_x, F_y)$ :

$$F_x = \begin{cases} \frac{\partial \delta(x-x_0, y-y_0)}{\partial x} (\sin(\pi(t-1)) + 0,8 \sin(2\pi(t-1)) + 0,2 \sin(3\pi(t-1))), & 0 \leq t \leq 2, \\ 0, & t \geq 2, \end{cases}$$

$$F_y = \begin{cases} \frac{\partial \delta(x-x_0, y-y_0)}{\partial y} (\sin(\pi(t-1)) + 0,8 \sin(2\pi(t-1)) + 0,2 \sin(3\pi(t-1))), & 0 \leq t \leq 2, \\ 0, & t \geq 2. \end{cases}$$

Рис. 5. Область с «горкой». Точка – расположение источника



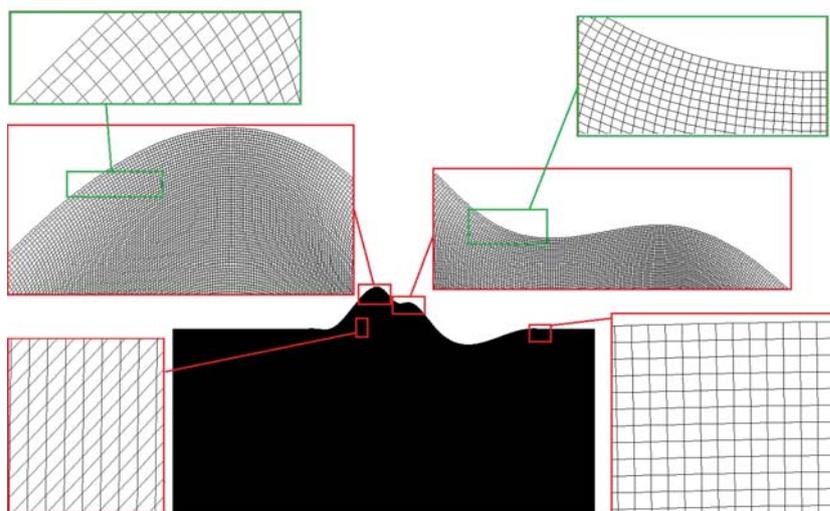


Рис. 6. Область и соответствующая ей криволинейная сетка

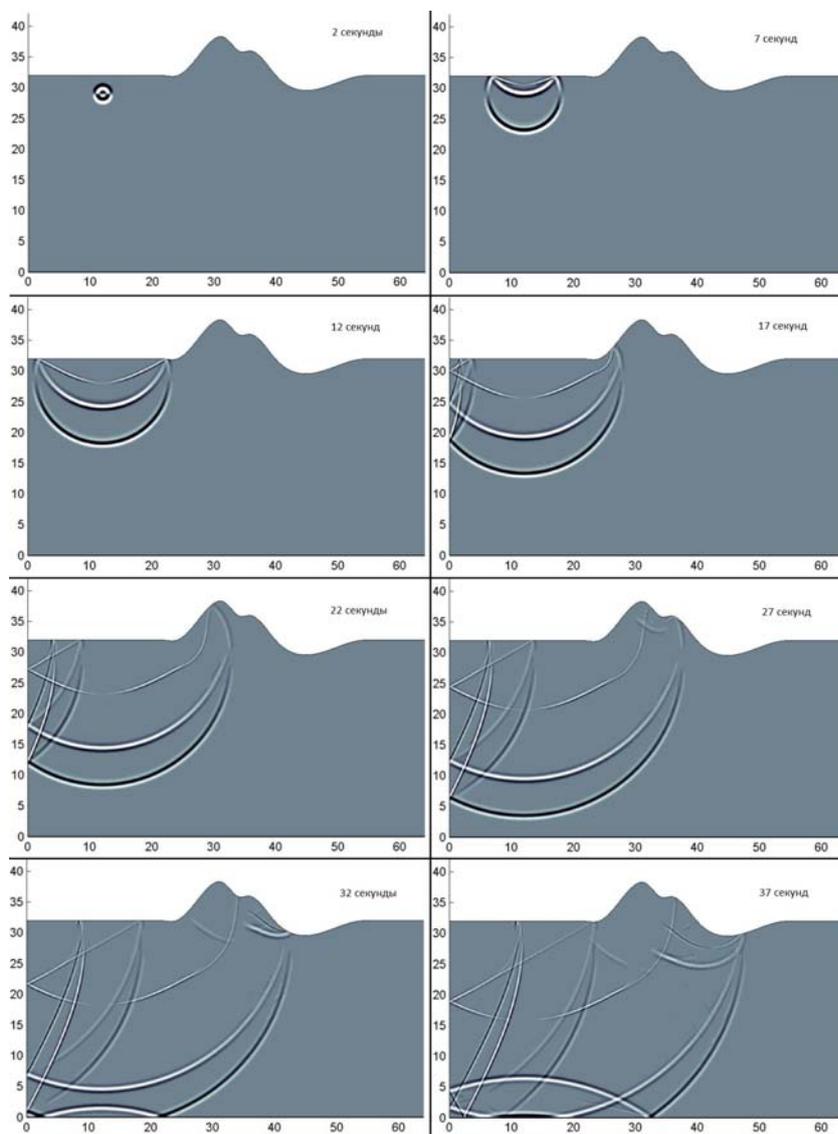


Рис. 7. Снимки волнового поля в последовательные моменты времени 2, 7, 12, 17, 22, 27, 32, 37 с

Важным свойством криволинейной сетки является ее ортогональность к свободной поверхности (рис. 6). Следует также упомянуть, что для двух разных форм поверхности не всегда можно использовать один и тот же алгоритм построения сетки. Зачастую возникает необходимость подбирать параметры управляющих функций занова, чтобы не допустить «перехлеста» ячеек криволинейной сетки. Аналитический метод построения позволяет делать это быстро. Программный генератор сетки написан на языке Fortran и также предложен автором.

### Результаты численного моделирования

На рис. 7 можно наблюдать отсутствие дифракций при отражении волны от поверхности, что подтверждает эффективность подхода, предложенного в данной работе.

### Результаты тестов

Далее приведены сравнительные тесты для разных архитектур (рис. 8). Модель описана в разделе «Математическая модель». Тесты для CPU выполнены на сервере G7 (2 × 6-ядерных Intel Xeon E5670 и 24 Гб оперативной памяти на узле) кластера ССКЦ НКС-30Т, при этом технология hyperthreading отключена. Для Intel Xeon Phi – на кластере МВС-10П. В случае Xeon Phi не проводилось никаких дополнительных процедур оптимизации вроде выравнивания массивов или изменения числа потоков на ядро (по умолчанию их 4). В тесте на масштабируемость алгоритма на каждый сопроцессор Phi приходился объем вычислений, соответствующий нашей модели (5 120 × 2 060 ячеек).

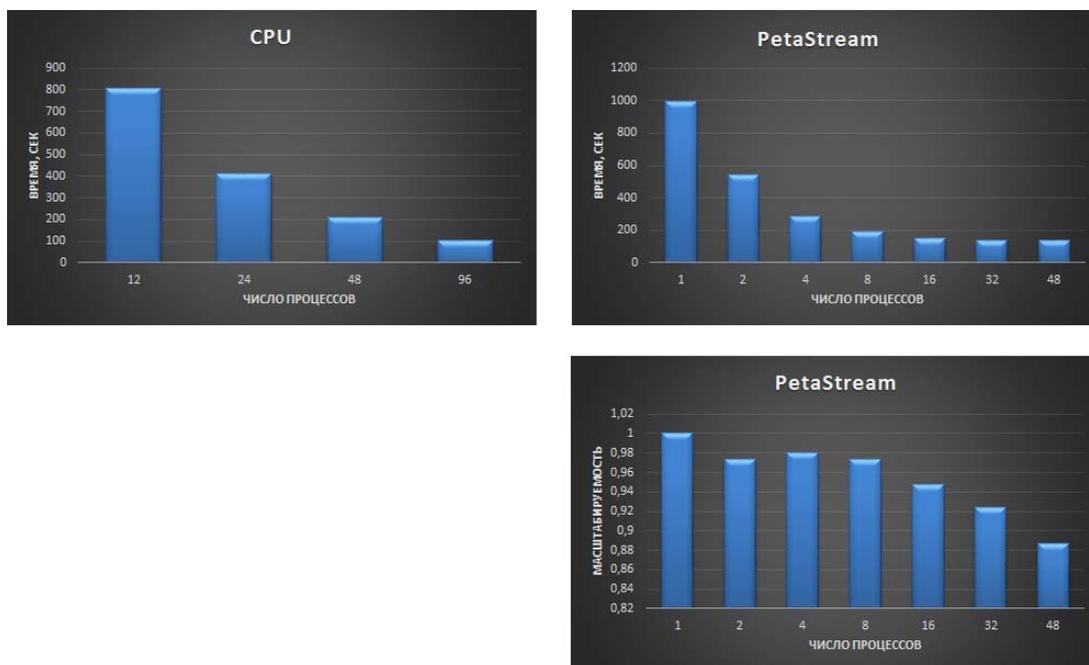


Рис. 8. Тесты. Слева – для CPU, справа – для PetaStream

Как видно из результатов, даже один узел с 12 ядрами CPU (на 1 ядро – 1 MPI-поток) справляется быстрее, чем один Intel Xeon Phi. При этом заметно, что в случае PetaStream алгоритм быстро «уходит в насыщение», что может свидетельствовать о недостаточной оптимизации MPI-обменов между соседними Phi на уровне программного обеспечения кластера.

## Заключение

В работе представлен алгоритм и результаты численного моделирования с использованием этого алгоритма. Также непосредственно автором создано 2 параллельных реализации алгоритма, программно реализован генератор криволинейной сетки. Проведены сравнительные тесты для разных архитектур на примере конкретной модели.

Новизна работы заключается в следующем: предложен новый способ аппроксимации граничных условий, впервые моделирование сейсмических полей проводилось с использованием локально-ортогональной криволинейной сетки, построенной аналитическим способом.

Проведенные тесты показали отсутствие дифракции при отражении от криволинейной свободной поверхности. В дальнейшем эта работа послужит основой для полноценной 3D-задачи, результаты которой могут быть применены для решения реальных задач.

Автор выражает благодарность В. Н. Мартынову за консультации в ходе выполнения работы, а также ССКЦ РАН и МСЦ РАН за предоставленное оборудование для численных расчетов.

## Список литературы

1. Лисейкин В. Д., Рычков А. Д., Кофанов А. В. Технология адаптивных сеток для решения прикладных задач: Монография / Новосибир. гос. ун-т. Новосибирск, 2011. 160 с. ISBN 978-5-94356-981-4
2. Хакимзянов Г. С., Шокин Ю. И. Лекции по разностным схемам на подвижных сетках. Часть 2. Задачи для уравнений в частных производных с двумя пространственными переменными. Казань, 2006.
3. Васева И. А., Кофанов А. В., Лисейкин В. Д., Лиханова Ю. В., Харитончик А. М. Технология построения пространственных адаптивных сеток // Журнал вычислительной математики и математической физики. 2010. Т. 50, № 1. С. 1–19.
4. Васева И. А., Лисейкин В. Д. Применение метода конечных элементов для построения адаптивных сеток // Вычислительные технологии. 2011. Т. 16, № 6. С. 30–35
5. Глассер А., Лисейкин В. Д., Китаева И. А. Контролирование свойств разностных сеток с помощью мониторинга метрики // Журнал вычислительной математики и математической физики. 2005. Т. 45, № 8. С. 1466–1483.
6. Глассер А., Лисейкин В. Д., Шокин Ю. И., Васева И. А., Лиханова Ю. В. Построение разностных сеток с помощью уравнений Бельтрами и диффузии. Новосибирск, Наука, 2006.
7. Лисейкин В. Д., Морочков Ю. И., Васева И. А., Лиханова Ю. В. Некоторые аспекты построения адаптивных сеток // Журнал вычислительной математики и математической физики. 2008. Т. 48, № 9. С. 1638–1658.
8. Лисейкин В. Д., Шокин Ю. И., Васева И. А., Лиханова Ю. В. Технология построения разностных сеток. Новосибирск: Наука, 2009.
9. Liseikin V. D. A computational differential geometry approach to grid generation. Berlin: Springer, 2004.
10. Liseikin V. D. A computational differential geometry approach to grid generation. 2<sup>nd</sup> ed. Berlin: Springer, 2007.
11. Liseikin V. D. Grid generation method. 2<sup>nd</sup> ed. Berlin: Springer, 2010.
12. Vaseva I. A., Liseikin V. D., Likhonova Yu. V., Morokov Yu. N. An elliptic method for construction of adaptive spatial grids // Russ. J. Numer. Anal. Math. Modelling. 2009. Vol. 1. P. 65–78.
13. Daniel Appelo, N. Anders Petersson. A Stable Finite Difference Method for the Elastic Wave Equation on Complex Geometries with Free Surfaces // Communications in Computational Physics. 2009. Vol. 5. No. 1. P. 84–107.
14. Mikhailenko B. G. Spectral Laguerre method for the approximate solution of time dependent problem // Appl. Math. Lett. 1999. Vol. 12. P. 105–110.

15. Демидов Г. В., Мартынов В. Н. Пошаговый метод решения эволюционных задач с использованием функций Лагерра // Сибирский журнал вычислительной математики. 2010. Т. 13, № 4. С. 413–422.
16. Фадеев Д. К., Фадеева В. Н. Вычислительные методы линейной алгебры. М.: Физматгиз, 1960.
17. Reinders James, Jeffers James. Intel Xeon Phi Coprocessor High Performance Programming. Morgan Kaufman Publ. Inc, 2013.
18. Klemm Michael. Programming for the Intel Xeon Phi Coprocessor. Intel Corporation, 2013.
19. Rezaur Rahman. Intel Xeon Phi Coprocessor Architecture and Tools: The Guide for Application Developers. Apress, 2013.
20. Parallel Programming and Optimization with Intel Xeon Phi Coprocessors. Colfax, 2013
21. McCool Michael, Reinders James, Robison Arch. Structured Parallel Programming: Patterns for Efficient Computation. Morgan Kaufman Publ. Inc, 2013.

Материал поступил в редколлегию 16.10.2018

**P. A. Titov**

*Institute of Computational Mathematics and Mathematical Geophysics SB RAS  
6 Academician Lavrentiev Ave., Novosibirsk, 630090, Russian Federation*

*tapawel@gmail.com*

## **ELASTIC WAVES MODELING IN MEDIA WITH COMPLEX FREE SURFACE TOPOGRAPHY**

Modeling is an integral part of the study of seismic wave propagation processes in various media. The widely used method of dividing a region into rectangular cells has a drawback: with a complex free surface structure (for example, a mountain), effects arise when a wave is reflected from this surface, due to the fact that the boundary is approximated by a step function for the numerical solving of the problem. In this paper, we propose a different approach to the discretization of the domain: the construction of a curvilinear grid that agrees well with the geometry of the free surface. An algorithm for numerical simulation based on a numerical solution of the linear 2D-system of the elasticity, written in displacements, using a curvilinear grid and the stepwise Laguerre method, is proposed. Results of the simulation are presented. Two parallel versions of the algorithm were also implemented, calculations were carried out on different multicore systems (on the classical multiprocessor architecture, as well as on the architecture using the Intel Xeon Phi co-processors – "RSC PetaStream"). Comparison tests of accelerations on different architectures are presented. Also features of the algorithm parallelization are described.

*Keywords:* 2D, elasticity, elastic waves, curvilinear surface, curvilinear grids, Laguerre method, Intel Xeon Phi, PetaStream, MPI, OpenMP.

### **References**

1. Liseikin V. D., Rychkov A. D., Kofanov A. V. Adaptive grid technology for applications. Novosibirsk, Novosibirsk State University, 2011, 160 p. ISBN 978-5-94356-981-4 (in Russ.)
2. Khakimzyanov G. S., Shokin Yu. I. Lectures on finite difference schemes on moving grids. Problems for partial differential equations with two space variables. Kazan, Publishing center of KazNU named after Al Farabi, 2006. (in Russ.)
3. Vaseva I. A., Kofanov A. V., Liseikin V. D., Yu. V., Kharitonchik A. M. Constructive technology of spatial adaptive grids. *Journal of Computational Mathematics and Mathematical Physics*, 2010, vol. 50, № 1, p. 1–19. (in Russ.)
4. Vaseva I. A., Liseikin V. D. Finite element method for adaptive grids construction *Computational Technologies*, 2011, vol. 16, № 6, p. 30–35. (in Russ.)

5. Glasser A., Liseikin V. D., Kitaeva I. A. Differential grids properties controlling via monitor metric. *Journal of Computational Mathematics and Mathematical Physics*, 2005, vol. 45, № 8, p. 1466–1483. (in Russ.)
6. Glasser A., Liseikin V. D., Shokin Yu. I., Vaseva I. A., Likhanova Yu. V. Differential grids construction via Beltrami and diffusion equations. Novosibirsk, Nauka, 2006. (in Russ.)
7. Liseikin V. D., Morokov Yu. I., Vaseva I. A., Likhanova Yu. V. Some aspects of adaptive grids construction. *Journal of Computational Mathematics and Mathematical Physics*, 2008, vol. 48, № 9, p. 1638–1658. (in Russ.)
8. Liseikin V. D., Shokin Yu. I., Vaseva I. A., Likhanova Yu. V. Differential grids construction technology. Novosibirsk, Nauka, 2009. (in Russ.)
9. Liseikin V. D. A computational differential geometry approach to grid generation. Berlin, Springer, 2004.
10. Liseikin V. D. A computational differential geometry approach to grid generation. 2<sup>nd</sup> ed. Berlin, Springer, 2007.
11. Liseikin V. D. Grid generation method. 2<sup>nd</sup> ed. Berlin: Springer, 2010.
12. Vaseva I. A., Liseikin V. D., Likhanova Yu. V., Morokov Yu. N. An elliptic method for construction of adaptive spatial grids. *Russ. J. Numer. Anal. Math. Modelling*, 2009, vol. 1, p. 65–78.
13. Daniel Appelo, N. Anders Petersson. A Stable Finite Difference Method for the Elastic Wave Equation on Complex Geometries with Free Surfaces. *Communications in Computational Physics*, 2009, vol. 5, no. 1, p. 84–107.
14. Mikhailenko B. G. Spectral Laguerre method for the approximate solution of time dependent problem. *Appl. Math. Lett.*, 1999, vol. 12, p. 105–110.
15. Demidov G. V., Martynov V. N. Stepwise method for solving evolution problems using Laguerre functions. *Siberian Journal of Numerical Mathematics*, 2010, vol. 13, № 4, p. 413–422. (in Russ.)
16. Fadeev D. K., Fadeeva V. N. Computational methods of linear algebra. Moscow, Fizmatgiz, 1960, p. 214–220. (in Russ.)
17. James Reinders, James Jeffers. Intel Xeon Phi Coprocessor High Performance Programming. Morgan Kaufman Publ. Inc, 2013.
18. Michael Klemm. Programming for the Intel Xeon Phi Coprocessor. Intel Corporation, 2013.
19. Rezaur Rahman. Intel Xeon Phi Coprocessor Architecture and Tools: The Guide for Application Developers. Apress, 2013.
20. Parallel Programming and Optimization with Intel Xeon Phi Coprocessors. Colfax, 2013.
21. Michael McCool, James Reinders, Arch Robison. Structured Parallel Programming: Patterns for Efficient Computation. Morgan Kaufman Publ. Inc, 2013.

Received 16.10.2018

*For citation:*

Titov P. A. Elastic Waves Modeling in Media with Complex Free Surface Topography. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 4, p. 153–166. (in Russ.) DOI 10.25205/1818-7900-2018-16-4-153-166

## СВЕДЕНИЯ ОБ АВТОРАХ

---

- Абрамский Михаил Михайлович** – старший преподаватель кафедры программной инженерии, Высшая школа информационных технологий и интеллектуальных систем, Казанский (Приволжский) федеральный университет
- Алымов Алексей Олегович** – магистрант геолого-геофизического факультета Новосибирского государственного университета
- Байбурин Ержан Мухаметкалиевич** – инженер Восточно-Казахстанского государственного университета им. С. Аманжолова, Усть-Каменогорск, Казахстан
- Балков Евгений Вячеславович** – кандидат технических наук, старший научный сотрудник лаборатории электромагнитных полей Института нефтегазовой геологии и геофизики им. А. А. Трофимука СО РАН, Новосибирск
- Барамия Денис Александрович** – аспирант факультета информационных технологий Новосибирского государственного университета
- Барахнин Владимир Борисович** – доктор технических наук, доцент, ведущий научный сотрудник Института вычислительных технологий СО РАН, профессор Новосибирского государственного университета
- Баталин Кирилл Вячеславович** – магистрант факультета информационных технологий Новосибирского государственного университета
- Боронина Марина Андреевна** – кандидат физико-математических наук, научный сотрудник Института вычислительной математики и математической геофизики СО РАН, Новосибирск
- Вшивков Виталий Андреевич** – доктор физико-математических наук, профессор, главный научный сотрудник Института вычислительной математики и математической геофизики СО РАН, Новосибирск
- Гаранина Наталья Олеговна** – кандидат физико-математических наук, старший научный сотрудник Института систем информатики им. А. П. Ершова СО РАН, Новосибирск
- Глинских Вячеслав Николаевич** – профессор РАН, доктор физико-математических наук, заведующий лабораторией, главный научный сотрудник Института нефтегазовой геологии и геофизики им. А. А. Трофимука СО РАН, доцент Новосибирского государственного университета, профессор Новосибирского государственного технического университета
- Горбенко Николай Иванович** – кандидат физико-математических наук, и. о. заведующего лабораторией Института вычислительной математики и математической геофизики СО РАН, Новосибирск
- Журавлев Сергей Сергеевич** – младший научный сотрудник Института вычислительных технологий СО РАН, Новосибирск
- Зюбин Владимир Евгеньевич** – доктор технических наук, ведущий научный сотрудник Института автоматики и электрометрии СО РАН, Новосибирск
- Карин Юрий Григорьевич** – научный сотрудник лаборатории электромагнитных полей Института нефтегазовой геологии и геофизики им. А. А. Трофимука СО РАН, Новосибирск
- Лаврентьев Михаил Михайлович** – доктор физико-математических наук, профессор, заместитель директора Института автоматики и электрометрии СО РАН, Новосибирск
- Лысаков Константин Фёдорович** – кандидат технических наук, заведующий кафедрой физического факультета Новосибирского государственного университета, научный сотрудник Института автоматики и электрометрии СО РАН, Новосибирск
- Лях Татьяна Викторовна** – инженер-программист Института автоматики и электрометрии СО РАН, Новосибирск

- Мамеев Никита Сергеевич** – магистрант факультета информационных технологий Новосибирского государственного университета
- Мансурова Мадина Есимхановна** – кандидат физико-математических наук, доцент, и. о. профессора Казахского национального университета имени аль-Фараби, Алматы, Казахстан
- Манштейн Александр Константинович** – доктор технических наук, ведущий научный сотрудник лаборатории электромагнитных полей Института нефтегазовой геологии и геофизики им. А. А. Трофимука СО РАН, Новосибирск
- Махина Екатерина Дмитриевна** – магистрант факультета информационных технологий Новосибирского государственного университета
- Молородов Юрий Иванович** – кандидат физико-математических наук, старший научный сотрудник Института вычислительных технологий СО РАН, доцент Новосибирского государственного университета
- Мызников Павел Владимирович** – аспирант факультета информационных технологий Новосибирского государственного университета
- Нечаев Олег Валентинович** – кандидат физико-математических наук, старший научный сотрудник Института нефтегазовой геологии и геофизики им. А. А. Трофимука СО РАН, Новосибирск
- Никитенко Марина Николаевна** – кандидат технических наук, старший научный сотрудник Института нефтегазовой геологии и геофизики им. А. А. Трофимука СО РАН, Новосибирск
- Нугуманова Алия Багдатовна** – PhD, и. о. доцента Восточно-Казахстанского государственного университета им. С. Аманжолова, Усть-Каменогорск, Казахстан
- Облаухов Константин Константинович** – ассистент, факультет информационных технологий Новосибирского государственного университета
- Околнишников Виктор Васильевич** – доктор технических наук, ведущий научный сотрудник Института вычислительных технологий СО РАН, Новосибирск
- Пальчунов Дмитрий Евгеньевич** – доктор физико-математических наук, ведущий научный сотрудник Института математики СО РАН, Новосибирск
- Панин Григорий Леонидович** – научный сотрудник лаборатории электромагнитных полей Института нефтегазовой геологии и геофизики им. А. А. Трофимука СО РАН, Новосибирск
- Попова Кристина Юрьевна** – магистрант факультета информационных технологий Новосибирского государственного университета
- Романов Дмитрий Борисович** – заместитель директора по региональному развитию ООО «Конструкторского Бюро Электрометрии», Москва
- Рудометов Сергей Валерьевич** – кандидат технических наук, научный сотрудник Института вычислительных технологий СО РАН, Новосибирск
- Рыжаков Игорь Дмитриевич** – магистрант факультета информационных технологий Новосибирского государственного университета
- Тимерханов Тимур Ильдарович** – студент Высшей школы информационных технологий и интеллектуальных систем, Казанский (Приволжский) федеральный университет
- Титов Павел Андреевич** – младший научный сотрудник лаборатории суперкомпьютерного моделирования Института вычислительной математики и математической геофизики СО РАН, Новосибирск

---

**Фадеев Денис Игоревич** – младший научный сотрудник лаборатории электромагнитных полей Института нефтегазовой геологии и геофизики им. А. А. Трофимука СО РАН, Новосибирск

**Ходорченко Дарья Андреевна** – магистрант механико-математического факультета Новосибирского государственного университета

**Шадрин Михаил Юрьевич** – доцент физического факультета Новосибирского государственного университета, научный сотрудник Института автоматизации и электротехники СО РАН, Новосибирск

**Шакиров Станислав Рудольфович** – кандидат физико-математических наук, заведующий конструкторско-технологическим отделом Института вычислительных технологий СО РАН, Новосибирск

**Яхьяева Гульнара Эркиновна** – кандидат физико-математических наук, доцент, доцент кафедры общей информатики факультета информационных технологий Новосибирского государственного университета

## ИНФОРМАЦИЯ ДЛЯ АВТОРОВ

---

Авторы представляют статьи на русском или английском языке объемом от 0,5 авторского листа (20 тыс. знаков) до 1 авторского листа (40 тыс. знаков), включая иллюстрации (1 иллюстрация форматом  $190 \times 270$  мм = 1/6 авторского листа, или 6,7 тыс. знаков).

Публикации, превышающие указанный объем, допускаются к рассмотрению только после индивидуального согласования с редакцией журнала.

### **Требования к оформлению основного текста и иллюстративных материалов**

Текст рукописи должен быть представлен в редколлегию в виде файла MS Word. Гарнитура Times New Roman, размер шрифта 14, межстрочный интервал 1,5, размеры полей – стандартные значения текстового процессора. Форматирование – выравнивание по ширине страницы, переносы слов отключены, красной строки нет. Не допускается ручное форматирование абзацев (пробелами, лишними переводами строк, разрывами страниц).

Подзаголовки набираются полужирным шрифтом. Для выделения частей текста и отдельных слов и словосочетаний допускается использование курсива или полужирного шрифта. Подчеркивание, разрядка, изменение основного кегля и выделение цветом и т. п. не используются.

Все иллюстрации к рукописи статьи должны быть приложены в отдельных файлах. При этом в тексте может содержаться как включенное изображение с указанием имени файла, так и только указание. Все иллюстрации, содержащие схемы, графики, алгоритмы, скриншоты и другие изображения должны быть представлены в максимально высоком качестве, без каких-либо потерь и искажений (.jpg, .tif).

Все иллюстрации должны иметь подрисуночную подпись.

Нумерация последовательная и неразрывная от начала статьи. Не допускается использование других наименований, кроме «рис.» и усложнение нумерации (например, «рис. 3.2.»). Ссылка на иллюстрацию в тексте должна быть приведена в круглых скобках: (рис. 1).

Формулы должны быть набраны с использованием редактора **MathType**. Кегль основных символов – 11, греческие символы набираются прямым шрифтом, латинские – курсивом. Нумеруются только те формулы, на которые автор ссылается в тексте.

Ссылки на литературу указываются цифрами в квадратных скобках. Список литературы нумеруется в порядке цитирования и оформляется в соответствии с ГОСТом на библиографическое описание. Ссылки на неопубликованные работы, а также на интернет-ресурсы (кроме электронных изданий, поддающихся библиографическому описанию) оформляются в виде сноски.

### **Дополнительные материалы**

К статье обязательно прилагаются следующие материалы.

- Информация об авторе / авторах на русском и английском языках:
  - ФИО полностью,
  - сведения об ученой степени и ученом звании,
  - должность и место работы, почтовый адрес,
  - электронный адрес,
  - контактный телефон.
- Индекс УДК (Универсальной десятичной классификации).
- Название статьи на русском языке и его авторский перевод на английский язык.
- Аннотация статьи на русском и английском языках.
- Ключевые слова (не более 10–15), на русском и английском (Keywords) языках.
- Список литературы на русском и английском (References) языках.

### **Доставка материалов**

Материалы предоставляются в редакцию только по электронной почте:

inftech@vestnik.nsu.ru