

# **СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ**

**Научный журнал Российской академии наук  
(издается под руководством Отделения нанотехнологий  
и информационных технологий РАН)**

Издается с 1989 года

Журнал выходит ежеквартально

## **Учредитель:**

**Федеральный исследовательский центр  
«Информатика и управление» Российской академии наук**

## **РЕДАКЦИОННЫЙ СОВЕТ**

академик РАН И. А. Соколов — председатель Редакционного совета

академик РАН Г. И. Савин

академик РАН А. Л. Стемпковский

член-корреспондент РАН Ю. Б. Зубарев

профессор Ш. Долев (S. Dolev, Beer-Sheva, Israel)

профессор Ю. Кабанов (Yu. Kabanov, Besancon, France)

профессор В. Ротарь (V. Rotar, San-Diego, USA)

профессор М. Финкельштейн (M. Finkelstein, Bloemfontein, South Africa)

## **РЕДАКЦИОННАЯ КОЛЛЕГИЯ**

академик РАН И. А. Соколов — главный редактор

профессор, д.ф.-м.н. С. Я. Шоргин — заместитель главного редактора

д.т.н. В. Н. Захаров

проф., д.ф.-м.н. А. И. Зейфман

проф., д.т.н. В. Д. Ильин

проф., д.т.н. К. К. Колин

проф., д.ф.-м.н. В. Ю. Королев

к.ф.-м.н. Р. В. Разумчик

д.ф.-м.н. В. И. Сеницын

проф., д.т.н. И. Н. Сеницын

проф., д.ф.-м.н. В. Г. Ушаков

д.ф.-м.н. А. К. Горшенин — отв. секретарь

к.ф.-м.н. С. А. Христочевский

## **Редакция**

к.ф.-м.н. Е. Н. Арутюнов

к.ф.-м.н. Р. В. Разумчик

С. Н. Стригина

© Федеральный исследовательский центр «Информатика  
и управление» Российской академии наук, 2022

Журнал включен в базу данных Russian Science Citation Index (RSCI),  
интегрированную с Web of Science

Журнал входит в систему Российского индекса научного цитирования (РИНЦ)

Журнал включен в базу данных CrossRef (систему DOI — Digital Object Identifier),  
в базу данных Ulrich's periodicals directory

и в информационную систему «Общероссийский математический портал Math-Net.Ru»

Журнал реферируется в «Реферативном журнале» ВИНИТИ  
и в системе Google Scholar

Журнал включен в сформированный Минобрнауки России Перечень рецензируемых научных  
изданий, в которых должны быть опубликованы основные научные результаты диссертаций  
на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук

**<http://www.ipiran.ru/journal/collected>**

# СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ

Том 32 № 2 Год 2022

## СОДЕРЖАНИЕ

|                                                                                                                                                 |           |
|-------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| Вопросы увеличения пропускной способности и территориальной доступности национальной исследовательской компьютерной сети России                 |           |
| <b><i>А. Г. Абрамов, А. В. Евсеев, А. А. Гончар, Б. М. Шабанов</i></b>                                                                          | <b>4</b>  |
| Машинное обучение для задач экстраполяции с малым объемом данных                                                                                |           |
| <b><i>А. О. Белозеров, А. И. Мазур</i></b>                                                                                                      | <b>13</b> |
| Построение сокращенного дерева достижимости для моделей программ в терминах сетей Петри                                                         |           |
| <b><i>Д. В. Леонтьев, Д. И. Харитонов</i></b>                                                                                                   | <b>23</b> |
| К вопросу о совмещении программного определения сетей с виртуализацией сетевых функций                                                          |           |
| <b><i>В. Б. Егоров</i></b>                                                                                                                      | <b>36</b> |
| О применении экспертных методов при оценке эффективности и качества информационных систем                                                       |           |
| <b><i>А. А. Зацаринный, Ю. С. Ионенков</i></b>                                                                                                  | <b>47</b> |
| Аналитическое моделирование и оценивание нестационарных нормальных процессов в стохастических системах, не разрешенных относительно производных |           |
| <b><i>И. Н. Синицын</i></b>                                                                                                                     | <b>58</b> |
| Некоторые подходы к анализу сетевых DLP                                                                                                         |           |
| <b><i>А. А. Грушо, Н. А. Грушо, М. И. Забежайло, Д. В. Смирнов, Е. Е. Тимонина, С. Я. Шоргин</i></b>                                            | <b>72</b> |
| Архитектуры интеллектуальных агентов сплоченных гибридных интеллектуальных многоагентных систем                                                 |           |
| <b><i>С. В. Листопад, И. А. Кириков</i></b>                                                                                                     | <b>81</b> |

# СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ

Том 32 № 2 Год 2022

## СОДЕРЖАНИЕ

|                                                                                                                                                                                       |            |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| Исследование методов детектирования ключевых точек при создании панорамных изображений<br><b>П. О. Архипов, А. К. Трофименков, М. В. Цуканов, Н. Ю. Носова</b>                        | <b>92</b>  |
| Модель определения необходимого числа оптических дисков, организованных в гибридные структуры, для создания и поддержания долговременных электронных архивов<br><b>А. В. Чернышов</b> | <b>105</b> |
| Модель процесса извлечения новых терминов и тональных слов из текстов<br><b>И. М. Зацман</b>                                                                                          | <b>115</b> |
| Подход к разработке концептуальной схемы базы геоданных с поддержкой многовариантного представления геообъектов<br><b>Д. А. Никишин</b>                                               | <b>128</b> |
| Подход к связыванию записей в технологии поддержки конкретно-исторических исследований, основанный на нечетких множествах<br><b>И. М. Адамович, О. И. Волков</b>                      | <b>137</b> |
| S-моделирование и цифровые технологии экономической деятельности<br><b>В. Д. Ильин</b>                                                                                                | <b>146</b> |
| Сравнительный анализ многопоточных транспортных протоколов<br><b>Е. П. Степанов, Р. Л. Смелянский</b>                                                                                 | <b>155</b> |
| Об авторах                                                                                                                                                                            | <b>171</b> |
| Правила подготовки рукописей статей                                                                                                                                                   | <b>175</b> |
| Requirements for manuscripts                                                                                                                                                          | <b>179</b> |

## ВОПРОСЫ УВЕЛИЧЕНИЯ ПРОПУСКНОЙ СПОСОБНОСТИ И ТЕРРИТОРИАЛЬНОЙ ДОСТУПНОСТИ НАЦИОНАЛЬНОЙ ИССЛЕДОВАТЕЛЬСКОЙ КОМПЬЮТЕРНОЙ СЕТИ РОССИИ\*

*А. Г. Абрамов<sup>1</sup>, А. В. Евсеев<sup>2</sup>, А. А. Гончар<sup>3</sup>, Б. М. Шабанов<sup>4</sup>*

**Аннотация:** Статья посвящена обсуждению планов развития Национальной исследовательской компьютерной сети России (НИКС) на период 2021–2024 гг. в рамках национального проекта «Наука и университеты». Акцент в работе сделан на текущем нормативном правовом статусе НИКС, существенных аспектах разработанного администратором сети и утвержденного профильным министерством порядка управления НИКС, включая цели, задачи и ожидаемые в процессе развития результаты. Внимание уделено также вопросам модернизации магистральной телекоммуникационной инфраструктуры сети и расширению ее территориальной доступности внутри России.

**Ключевые слова:** национальная исследовательская компьютерная сеть; НИКС; национальная научно-образовательная сеть; нормативный правовой статус; телекоммуникационная инфраструктура; региональное развитие

**DOI:** 10.14357/08696527220201

### 1 Введение

Термин «национальная научно-образовательная сеть» (National Research and Education Network, NREN) достаточно давно и прочно вошел в профессиональный лексикон мирового сообщества сферы науки и образования. Цели и задачи таких сетей на всех исторических этапах направлены на максимальное соответствие постоянно растущему уровню развития информационно-коммуникационных технологий (ИКТ), изменяющимся внешним условиям и реальным потребностям науки и образования, которые во многих аспектах выходят за рам-

---

\*Статья публикуется по представлению программного комитета VI Международной научно-практической конференции «Информационные технологии и высокопроизводительные вычисления» (ITHPC-2021). Работа была выполнена в МСЦ РАН в рамках государственного задания по теме FNEF-2022-0014.

<sup>1</sup>Санкт-Петербургское отделение Межведомственного суперкомпьютерного центра Российской академии наук — филиала ФГУ ФНЦ НИИСИ РАН, [abramov@niks.ru](mailto:abramov@niks.ru)

<sup>2</sup>Санкт-Петербургское отделение Межведомственного суперкомпьютерного центра Российской академии наук — филиала ФГУ ФНЦ НИИСИ РАН, [evseev@niks.ru](mailto:evseev@niks.ru)

<sup>3</sup>Межведомственный суперкомпьютерный центр Российской академии наук — филиал ФГУ ФНЦ НИИСИ РАН, [andrey.gonchar@jscs.ru](mailto:andrey.gonchar@jscs.ru)

<sup>4</sup>Межведомственный суперкомпьютерный центр Российской академии наук — филиал ФГУ ФНЦ НИИСИ РАН, [shabanov@jscs.ru](mailto:shabanov@jscs.ru)

ки возможностей коммерческих операторов связи и к тому же не в полной мере отвечают их интересам [1–5].

Уровень функционального и технологического развития NREN в конкретной стране можно рассматривать сквозь призму состояния области науки и образования в целом. Ведущие NREN выступают в качестве магистральных инфраструктур и сервисных платформ для крупных исследовательских проектов, интегрированная инфраструктура NREN ответственна за решение задач высокоскоростной передачи, обработки и хранения научных данных и обеспечения доступа к ним, а также к востребованным ресурсам и сервисам.

В России функции NREN выполняет НИКС нового поколения [6, 7], созданная в 2019 г. Межведомственным суперкомпьютерным центром РАН (МЦЦ РАН) по заданию Минобрнауки России. Инфраструктуру, сервисную экосистему и пользовательскую базу НИКС исходно составили две крупнейшие отраслевые сети страны: университетская сеть RUNNet [8] и сеть РАН — RASNet.

Следует констатировать, что на текущий момент по уровню развития и использования телекоммуникационной инфраструктуры и сервисов НИКС заметно отстает от ведущих NREN, как было отмечено в [7]. В 2021 г. Минобрнауки России утверждены Порядок управления НИКС, а также Концепция и дорожная карта функционирования и развития НИКС на 2021–2024 гг. Функции администратора и оператора НИКС закреплены за МЦЦ РАН. Подчеркнем, что впервые за 25-летнюю историю существования в стране отдельных разрозненных научно-образовательных сетей статус NREN официально закреплен за единой сетью федерального масштаба.

## **2 Нормативный статус Национальной исследовательской компьютерной сети и некоторые аспекты порядка управления сетью**

В круг задач федерального проекта «Развитие инфраструктуры для научных исследований и подготовки кадров» (в составе национального проекта «Наука и университеты») входит осуществление развития НИКС в интересах научных и образовательных организаций страны, в том числе с целью обеспечения доступа к передовой научной инфраструктуре коллективного пользования, путем увеличения пропускной способности магистральной инфраструктуры сети, территориальной доступности и набора сервисов.

В качестве характеристики результата выполнения федерального проекта определено, что НИКС предоставит ведущим научным и образовательным организациям страны возможности для выполнения исследований и разработок по приоритетным направлениям научно-технологического развития (НТР), участия в международных и российских научных проектах, базирующихся на использовании устойчивой и отвечающей современным требованиям отраслевой сети.

Предполагается, что на базе усовершенствованной инфраструктурно-сервисной платформы НИКС будет решен ряд задач, среди которых:

- интеграция существующих и создаваемых суперкомпьютерных центров (СКЦ), отдельных центров коллективного пользования (ЦКП) и функционирующих на их базе уникальных научных установок (УНУ), научно-образовательных центров (НОЦ) и научных центров мирового уровня (НЦМУ);
- организация высокоскоростной телекоммуникационной связности научных и образовательных организаций страны;
- реализация эффективного взаимодействия с крупными международными исследовательскими центрами и проектами за счет сотрудничества с мировыми NREN и сетевыми ассоциациями;
- развертывание цифровой среды взаимодействия и отработки перспективных технологических ИКТ-решений;
- обеспечение применения ИКТ в образовании, включая доступ к электронным образовательным ресурсам, обеспечение технологий дистанционного обучения и научно-образовательной мобильности;
- обеспечение выполнения требований информационной безопасности в целях устойчивого функционирования российского сегмента интернета.

К 2024 г. к НИКС будет подключено не менее 90% ведущих научных и образовательных организаций страны, которые соответствуют одному из следующих установленных критериев [7, 9]:

- (1) организации высшего образования, уполномоченные самостоятельно разрабатывать и утверждать образовательные стандарты (№ 273-ФЗ);
- (2) научные организации первой категории, определенные в соответствии с постановлением Правительства РФ от 8 апреля 2009 г. № 312;
- (3) национальные исследовательские университеты;
- (4) федеральные исследовательские центры;
- (5) университеты, находящиеся в ведении Правительства РФ.

Набор критериев и, как следствие, перечень подключаемых к НИКС организаций могут корректироваться с учетом возможных изменений в действующих документах стратегического планирования. В качестве дополнительных критериев (или заменяющих установленные) можно рассматривать участие организации в Программе «Приоритет 2030», в НОЦ, НЦМУ, наличие СКЦ или ЦКП с возможностью удаленного доступа [9].

Порядком управления утверждены следующие ключевые цели НИКС:

- предоставление возможностей для выполнения исследований и разработок по приоритетным направлениям НТР, применения современных ИКТ при реализации образовательных программ, участия в крупных российских и международных проектах;

- участие в формировании и обеспечении функционирования единого научно-образовательного информационного пространства в целях эффективного использования передовой инфраструктуры научных исследований, разработок и инновационной деятельности;
- обеспечение ведомственной сетевой связности и предоставление набора сервисов, развитие интерфейсов управления собственными ресурсами для интеграции с цифровыми платформами науки и образования.

К основным задачам НИКС, помимо явно обозначенных в федеральном проекте, следует отнести:

- обеспечение гарантированного качества и высокой надежности доступа к результатам научных исследований, массивам открытых данных, научно-образовательным ресурсам и сервисам с использованием собственной ИКТ-инфраструктуры и инфраструктуры зарубежных NREN;
- предоставление целевым пользователям возможностей удаленного доступа к объектам научной инфраструктуры коллективного пользования;
- формирование высокоскоростной сети передачи данных для объединения отраслевых СКЦ в национальную суперкомпьютерную инфраструктуру;
- обеспечение межсетевого взаимодействия российских организаций — участников хранения, обработки и анализа данных экспериментов на существующих и создаваемых установках класса «мегасайенс»;
- предоставление возможностей для формирования новых российских и международных коллабораций, расширения научного взаимодействия;
- оптимизация способов сетевого взаимодействия организаций, ведущих фундаментальные и прикладные исследования в приоритетных областях, с организациями реального сектора экономики;
- апробация, внедрение, сопровождение и предоставление сервисов для пользователей НИКС, осуществление мониторинга востребованности, отслеживание спроса и прогнозирование перспектив;
- создание условий для участия российских производителей программного обеспечения в формировании сервисов НИКС в интересах пользователей в рамках процессов импортозамещения.

### **3 Развитие телекоммуникационной инфраструктуры Национальной исследовательской компьютерной сети**

Актуальная телекоммуникационная инфраструктура НИКС подразделяется на магистральную и инфраструктуру доступа [6–8]. Транспортной основой магистральной инфраструктуры служит опорная сеть, которая обеспечивает связность между федеральными округами и расположенными в них отдельными регионами, а также с зарубежными NREN. Функционирование опорной сети

протяженностью более 5000 км обеспечивают внутрироссийские и зарубежные магистральные узлы связи.

В рамках планов развития инфраструктуры и сервисов НИКС предполагаются к выполнению следующие основные мероприятия [7]:

- расширение территориальной доступности НИКС внутри страны;
- увеличение пропускной способности магистральной инфраструктуры;
- подключение к НИКС ведущих научных организаций и организаций высшего образования страны (в соответствии с критериями);
- подключение к НИКС высокопроизводительных СКЦ коллективного пользования сферы науки и образования;
- подключение к НИКС отдельных ЦКП и УНУ, имеющих возможность проводить исследования в режиме удаленного доступа;
- развитие и модернизация сервисов НИКС уровня NREN.

В 2021 г. стартовали работы по масштабной модернизации сетевой инфраструктуры НИКС, вводу в эксплуатацию новых магистральных узлов связи, расширению территориального присутствия в Сибири, на Дальнем Востоке и на Юге России. В планах развития — построение резервированной сетевой магистрали с пропускной способностью 10...100 Гбит/с, подключение к сети более 300 организаций, интеграция отраслевых СКЦ в единую сеть. Сеть передачи данных обеспечит пропускную способность для пользователей не менее 1 Гбит/с, а для высокопроизводительных СКЦ — до 10 Гбит/с.

В ходе планирования работ следует учитывать имеющиеся различия в уровне развития ИКТ-инфраструктуры регионов, конкурентоспособности локальных рынков телекоммуникационных услуг, доступности телекоммуникационных площадок и другие аспекты. В этих условиях имеются риски возникновения трудностей при подключении организаций к сети в отдельных регионах, в связи с чем особое внимание уделяется налаживанию сотрудничества с региональными научно-образовательными сетями, к которым подключены потенциальные пользователи НИКС. В качестве первого реализованного проекта — интеграция в НИКС компонентов региональной сети РОКСОН, имевшей собственную опорную инфраструктуру и предоставлявшей ИКТ-услуги более 30 институтам РАН, расположенным в Санкт-Петербурге.

## **4 Заключение**

Включение НИКС в национальный проект закрепляет ее особую роль в числе приоритетов НТР страны на ближайшие периоды. Устойчивая адресная поддержка, увеличение объемов финансирования позволило сформировать долгосрочное видение и стратегический план развития НИКС.

Впервые за долгое время представилась возможность преодолеть разрыв в уровне развития НИКС по сравнению с ведущими мировыми NREN, создать

современную сеть для поддержки деятельности в сфере науки и образования, реализовать и предоставлять пользователям лучшую в своем классе сеть и сервисы, обеспечить устойчивую и эффективную цифровую платформу для российского научно-образовательного сообщества.

Предполагается, что создаваемая единая отраслевая цифровая инфраструктура, отвечающая современному уровню достижений отрасли ИКТ, будет учитывать актуальные потребности организаций, подведомственных не только Минобрнауки России, но и другим федеральным органам исполнительной власти, а также межведомственным научно-образовательным и научно-техническим объединениям.

## Литература

1. *Allocchio C., Balint L., Berkhout V., et al.* A history of international research networking: The people who made it happen. — New York, NY, USA: Wiley-VCH, 2010. 317 p.
2. *Lehtisalo K.* The history of NORDUnet: Twenty-five years of networking cooperation in the Nordic countries. — NORDUnet, 2005. 153 p. <http://www.nordu.net/history/book.html>.
3. *Andronico G., Ardizzone V., Barbera R., et al.* e-Infrastructures for e-Science: A global view // *J. Grid Comput.*, 2011. Vol. 9. P. 155–184. doi: 10.1007/s10723-011-9187-y.
4. *Ижванов Ю. Л.* Научно-образовательные компьютерные сети: Прошлое, настоящее и тенденции развития // *Образовательные ресурсы и технологии*, 2017. № 2. С. 17–25.
5. *GEANT: The Case for NRENs.* A repository of resources to support funding, advocacy and the advancement of national and regional R&E networks. <https://caseforrens.org>.
6. *Абрамов А. Г., Гончар А. А., Евсеев А. В., Шабанов Б. М.* Национальная исследовательская компьютерная сеть нового поколения: текущее состояние и концепция развития // *Информационные технологии*, 2021. Т. 27. № 3. С. 115–124. doi: 10.17587/it.27.115-124.
7. *Abramov A. G., Evseev A. V., Gonchar A. A., Telegin P. N., Shabanov B. M.* National Research Computer Network of Russia: Regulatory status and plans for the development of regional telecommunications infrastructure in 2021–2024 // *CEUR Workshop Proc.*, 2021. Vol. 2930. P. 29–37.
8. *Абрамов А. Г., Евсеев А. В.* Концептуальные аспекты создания в Российской Федерации национальной исследовательской компьютерной сети нового поколения // *Информационные технологии*, 2019. Т. 25. № 12. С. 724–733. doi: 10.17587/it.25.724-733.
9. *Абрамов А. Г., Гончар А. А., Евсеев А. В., Шабанов Б. М.* Разработка критериев подключения научных и образовательных организаций к Национальной исследовательской компьютерной сети // *Информационные технологии и вычислительные системы*, 2021. № 2. С. 22–33. doi: 10.14357/20718632210203.

Поступила в редакцию 09.12.21

## ISSUES OF INCREASING THE NETWORK BANDWIDTH AND TERRITORIAL ACCESSIBILITY OF NATIONAL RESEARCH COMPUTER NETWORK OF RUSSIA

A. G. Abramov<sup>1</sup>, A. V. Evseev<sup>1</sup>, A. A. Gonchar<sup>2</sup>, and B. M. Shabanov<sup>2</sup>

<sup>1</sup>St.-Petersburg Department of Joint Supercomputer Center of the Russian Academy of Sciences — Branch of Federal State Institution “Scientific Research Institute for System Analysis of the Russian Academy of Sciences,” 14 Smolenka River emb., St. Petersburg 195251, Russian Federation

<sup>2</sup>Joint Supercomputer Center of the Russian Academy of Sciences — Branch of Federal State Institution “Scientific Research Institute for System Analysis of the Russian Academy of Sciences,” 32a Leninsky Prosp., Moscow 119334, Russian Federation

**Abstract:** The paper is devoted to the discussion of plans for the development of the National Research Computer Network of Russia (NIKS) for the period 2021–2024 within the framework of the National Project “Science and Universities.” The emphasis in the work is made on the current regulatory status of the NIKS project, significant aspects of the management procedure developed by the network administrator and approved by the line Ministry, including goals, objectives, and expected results in the development process. Special attention is paid to the issues of modernizing the backbone telecommunications infrastructure of NIKS and expanding the territorial accessibility of the network within Russia.

**Keywords:** National Research Computer Network; NIKS; National Research and Educational Network; regulatory legal status; telecommunications infrastructure; regional expansion

**DOI:** 10.14357/08696527220201

### Acknowledgments

The work was carried out at JSCC RAS as a part of government assignment (topic FNEF-2022-0014). The paper is published on the proposal of the Program Committee of the VI International Scientific and Practical Conference “Information Technologies and High Performance Computing” (ITHPC-2021).

### References

1. Allocchio, C., L. Balint, V. Berkhout, *et al.* 2010. *A history of international research networking: The people who made it happen*. New York, NY: Wiley-VCH. 317 p.
2. Lehtisalo, K. 2005. *The history of NORDUnet: Twenty-five years of networking cooperation in the Nordic countries*. NORDUnet. 153 p. Available at: <http://www.nordu.net/history/book.html> (accessed March 9, 2022).

3. Andronico, G., V. Ardizzzone, R. Barbera, *et al.* 2011. e-Infrastructures for e-Science: A global view. *J. Grid Comput.* 9: 155–184. doi: 10.1007/s10723-011-9187-y.
4. Izhvanov, Yu. L. 2017. Nauchno-obrazovatel'nye komp'yuternye seti: Proshloe, nastoyashchee i tendentsii razvitiya [Scientific-educational computer networks. Past, present and development trends]. *Obrazovatel'nye resursy i tekhnologii* [Educational Resources and Technologies] 2:17–25.
5. GÉANT: The case for NRENs. A repository of resources to support funding, advocacy and the advancement of national and regional R&E networks. Available at: <https://caseforrens.org> (accessed March 9, 2022).
6. Abramov, A. G., A. A. Gonchar, A. V. Evseev, and B. M. Shabanov. 2021. Natsional'naya issledovatel'skaya komp'yuternaya set' novogo pokoleniya: tekushchee sostoyanie i kontseptsiya razvitiya [The new generation National Research Computer Network: Current status and concept for the development]. *Informatsionnye tekhnologii* [Information Technologies] 27(3):115–124. doi: 10.17587/it.27.115-124.
7. Abramov, A. G., A. V. Evseev, A. A. Gonchar, P. N. Telegin, and B. M. Shabanov. 2021. National Research Computer Network of Russia: Regulatory status and plans for the development of regional telecommunications infrastructure in 2021–2024. *CEUR Workshop Procee.* 2930:29–37.
8. Abramov, A. G., and A. V. Evseev. 2019. Kontseptual'nye aspekty sozdaniya v Rossiyskoy Federatsii natsional'noy issledovatel'skoy komp'yuternoy seti novogo pokoleniya [Conceptual aspects of creating the New Generation National Research Computer Network in the Russian Federation]. *Informatsionnye tekhnologii* [Information Technologies] 25(12):724–733. doi: 10.17587/it.25.724-733.
9. Abramov, A. G., A. A. Gonchar, A. V. Evseev, and B. M. Shabanov. 2021. Razrabotka kriteriev podklyucheniya nauchnykh i obrazovatel'nykh organizatsiy k Natsional'noy issledovatel'skoy komp'yuternoy seti [Development of criteria for connecting scientific and educational organizations to the National Research Computer Network]. *Informatsionnye tekhnologii i vychislitel'nye sistemy* [J. Information Technologies and Computing Systems] 2:22–33. doi: 10.14357/20718632210203.

*Received December 9, 2021*

## Contributors

**Abramov Alexey G.** (b. 1976) — Candidate of Science (PhD) in physics and mathematics, assistant professor, leading scientist, St.-Petersburg Department of Joint Supercomputer Center of the Russian Academy of Sciences — Branch of Federal State Institution “Scientific Research Institute for System Analysis of the Russian Academy of Sciences,” 14 Smolenska River emb., St. Petersburg 195251, Russian Federation; [abramov@niks.su](mailto:abramov@niks.su)

**Evseev Anton V.** (b. 1973) — director, St.-Petersburg Department of Joint Supercomputer Center of the Russian Academy of Sciences — Branch of Federal State Institution “Scientific Research Institute for System Analysis of the Russian

Academy of Sciences,” 14 Smolenka River emb., St. Petersburg 195251, Russian Federation; evseev@niks.su

**Gonchar Andrey A.** (b. 1969) — deputy director, Joint Supercomputer Center of the Russian Academy of Sciences — Branch of Federal State Institution “Scientific Research Institute for System Analysis of the Russian Academy of Sciences,” 32a Leninsky Prosp., Moscow 119334, Russian Federation; andrey.gonchar@jscc.ru

**Shabanov Boris M.** (b. 1954) — Doctor of Science in technology, assistant professor, director, Joint Supercomputer Center of the Russian Academy of Sciences — Branch of Federal State Institution “Scientific Research Institute for System Analysis of the Russian Academy of Sciences,” 32a Leninsky Prosp., Moscow 119334, Russian Federation; shabanov@jscc.ru

## МАШИННОЕ ОБУЧЕНИЕ ДЛЯ ЗАДАЧ ЭКСТРАПОЛЯЦИИ С МАЛЫМ ОБЪЕМОМ ДАННЫХ\*

А. О. Белозеров<sup>1</sup>, А. И. Мазур<sup>2</sup>

**Аннотация:** Предложен новый метод экстраполяции вариационных расчетов, в основе которого лежит обучение большого числа искусственных нейронных сетей (ИНС) с последующей фильтрацией для отбора сетей, обученных наилучшим образом. Метод апробирован на модельной задаче с точным решением и применен к оценке энергии связи ядра  ${}^4\text{He}$  на основе расчетов в модели оболочек без инертного кора (МОБИК) с реалистичским потенциалом Daejeon16. Изучена сходимость результатов при увеличении объема входных данных. Показано, что метод обеспечивает достаточно высокую точность прогноза даже в случае небольших модельных пространств. Предложенный метод можно применять как для нахождения других характеристик ядер, так и для решения иных задач, не связанных с ядерной физикой.

**Ключевые слова:** машинное обучение; методы экстраполяции; энергия основного состояния; модель ядерных оболочек

**DOI:** 10.14357/08696527220202

### 1 Введение

В настоящее время методы машинного обучения широко используются для решения различных задач. Однако у этих методов есть два основных ограничения для реализации. Во-первых, как и большинство других статистических методов, машинное обучение требует для корректной работы большого объема данных. Обычно хорошие наборы данных состоят минимум из 1000 примеров. Второе ограничение относится к типу проблемы. Обычно машинное обучение используется в задачах интерполяции, в то время как его использование в задачах экстраполяции изучено недостаточно.

Модель оболочек без инертного кора [1] в настоящее время является одним из основных методов описания атомных ядер. Этот подход называется *ab initio*, поскольку в качестве входной информации используется только реалистичное нуклон-нуклонное взаимодействие (различные типы реалистичных потенциалов представлены в [2–5]).

---

\*Статья публикуется по представлению программного комитета VI Международной научно-практической конференции «Информационные технологии и высокопроизводительные вычисления» (ITNRC-2021). Работа выполнена при поддержке Министерства науки и высшего образования РФ в рамках проекта № 0818-2020-0005 и гранта РФФИ № 20-02-00357 (частично).

<sup>1</sup>Тихоокеанский государственный университет, aobelozerov@gmail.com

<sup>2</sup>Тихоокеанский государственный университет, mazur@khhb.ru

Результаты расчетов в оболочечных моделях, включая МОБИК, зависят от двух параметров: размера базисного пространства модели, определяемого максимальным числом осцилляторных квантов  $N_{\max}$ , и величины осцилляторной энергии  $\hbar\Omega$ .

Основная проблема МОБИК — резкое увеличение объема требуемых вычислительных ресурсов с увеличением  $N_{\max}$ , что связано с экспоненциальным ростом числа базисных функций. Как следствие, современные суперкомпьютеры позволяют проводить расчеты в МОБИК с  $N_{\max} \leq 20$  только для самых легких ядер.

Более точные результаты могут быть получены с использованием простых методов экстраполяции результатов МОБИК на бесконечно большие модельные пространства, таких как экстраполяция В [6, 7]. Было разработано множество методов для экстраполяции вычислений в более сложных ядерных системах [8–13]. Однако все они не имеют строгого обоснования, а потому вопрос поиска новых методов экстраполяции остается актуальным.

Цель данной работы — разработка нового метода экстраполяции результатов вариационных вычислений на бесконечные модельные пространства с использованием методов машинного обучения, способного работать с малыми наборами данных. Полученный алгоритм протестирован на модельной задаче с известным точным решением и применен к задаче экстраполяции энергии основного состояния  $^4\text{He}$  на случай бесконечно больших модельных пространств.

## 2 Модельная задача

Рассмотрим движение квантовой частицы с приведенной массой  $\mu c^2 = 751$  МэВ в поле сферически-симметричного потенциала в парциальной волне  $l = 1$  с полным угловым моментом  $j = 1/2$ . Радиальное уравнение Шредингера имеет вид:

$$Hu_l(k, r) = Eu_l(k, r). \quad (1)$$

Здесь гамильтониан  $H = T + V$ , где  $T$  — оператор кинетической энергии. В качестве модельного потенциала  $V$  используется потенциал Вудса–Саксона с поверхностными спин-орбитальными силами:

$$V(r) = \frac{V_0}{1 + \exp((r - R)/a)} + \frac{2(\mathbf{l} \cdot \mathbf{s})V_{ls}}{ra} \frac{\exp((r - R)/a)}{(1 + \exp((r - R)/a))^2},$$

где  $V_0 = -30$  МэВ;  $V_{ls} = -10$  МэВ·фм<sup>2</sup>;  $R = 3,08$  фм;  $a = 0,53$  фм.

В осцилляторном базисе уравнение (1) эквивалентно бесконечной системе линейных алгебраических уравнений на волновые функции в осцилляторном представлении  $a_{Nl}(k)$ :

$$\sum_{N'=N_0, N_0+2, \dots, \infty} a_{N'l}(k) (H_{NN'} - \delta_{NN'} E) = 0, \quad N = N_0, N_0 + 2, \dots, \infty. \quad (2)$$

Здесь  $H_{NN'}$  — матричные элементы гамильтониана:

$$H_{NN'} = \int_0^\infty R_{Nl} H R_{N'l} dr ,$$

где  $R_{Nl}$  — радиальная осцилляторная функция;  $N = 2n + l$  — полное (энергетическое) число осцилляторных квантов, где  $n$  — главное квантовое число;  $N_0$  — минимальное число осцилляторных квантов рассматриваемой системы (в данном случае  $N_0 = l$ ).

Для короткодействующих потенциалов уравнение (2) сводится к конечной системе уравнений:

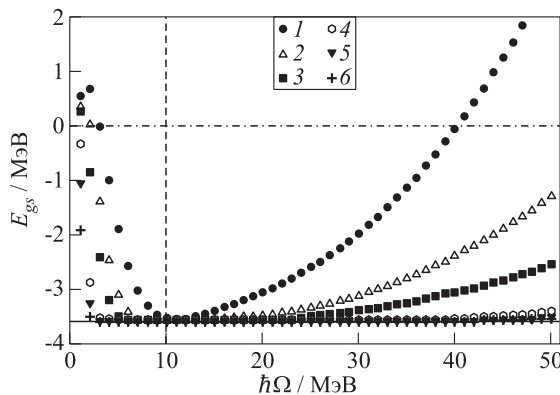
$$\sum_{N'=N_0, N_0+2, \dots, N_0+N_{\max}} a_{N'l}(k) (H_{NN'} - \delta_{NN'} E) = 0 ,$$

$$N = N_0, N_0 + 2, \dots, N_0 + N_{\max} , \quad (3)$$

где  $N_{\max}$  — число квантов возбуждения, задающее модельное пространство.

После диагонализации гамильтониана  $H$  в (3) получаем спектр собственных энергий  $E_\lambda$ . Минимальное значение функции  $E_{gs}(\hbar\Omega)$  для данного модельного пространства ассоциируется с энергией основного состояния ядра  $E_{gs}$ . Отметим, что для модельной задачи размер матрицы гамильтониана с ростом  $N_{\max}$  увеличивается линейно.

Метод имеет два параметра: значение осцилляторной энергии  $\hbar\Omega$  и максимальное число квантов возбуждения  $N_{\max}$ . Были проведены оболочечные расчеты  $E_{gs}(\hbar\Omega)$  в различных модельных пространствах  $N_{\max} = 4 \dots 40$  с шагом в 1 МэВ по  $\hbar\Omega$ . На рис. 1 показаны данные, полученные в отдельных модельных пространствах.



**Рис. 1** Вариационные расчеты в модельной задаче: 1 —  $N_{\max} = 4$ ; 2 — 8; 3 — 12; 4 — 24; 5 — 30; 6 —  $N_{\max} = 38$

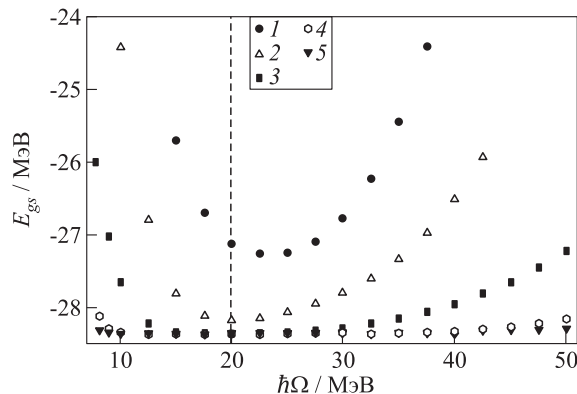
Горизонтальная сплошная линия обозначает точное значение  $E_{gs}^{\text{exp}} = -3,5814$  МэВ энергии основного состояния системы, полученное прямым интегрированием уравнения Шредингера. Как показано в [14], некоторые методы экстраполяции работают лучше, когда используются только данные, лежащие правее минимумов кривых  $E_{gs}(\hbar\Omega)$  (вертикальная штриховая линия). Предыдущие работы авторов статьи подтверждают это утверждение, дополняя его фактом улучшения производительности после использования данных только с  $E_{gs} < 0$  МэВ (горизонтальная штрихпунктирная линия).

### 3 Задача нахождения энергии основного состояния ядра ${}^4\text{He}$

В МОБИК волновая функция разложена по многочастичному осцилляторно-му базису. Как и в случае описанной выше модельной задачи, энергия основного состояния  $E_{gs}$  является функцией максимального числа осцилляторных квантов  $N_{\text{max}}$  и параметра осцилляторной функции  $\hbar\Omega$ . Но, в отличие от модельной задачи, размер многочастичного базиса с увеличением  $N_{\text{max}}$  увеличивается не линейно, а экспоненциально.

На рис. 2 показаны результаты расчетов энергии  $E_{gs}$  основного состояния ядра атома гелия  ${}^4\text{He}$  в МОБИК с реалистическим взаимодействием Daejeon16 [5] для некоторых значений  $N_{\text{max}}$ . Расчеты проводились в модельных пространствах  $N_{\text{max}} = 2 \dots 18$  при  $\hbar\Omega = 8, 9$  и 10 МэВ и далее с шагом 2,5 МэВ до 50 МэВ. Как и в модельной задаче, данные, лежащие слева от минимумов  $E(\hbar\Omega)$  (вертикальная штриховая линия), не использовались в процессе обучения ИНС.

Точное значение энергии основного состояния может быть получено в пределе  $N_{\text{max}} \rightarrow \infty$ . Однако из-за экспоненциального роста числа базисных функ-



**Рис. 2** Энергия основного состояния ядра  ${}^4\text{He}$   $E_{gs}$ , вычисленная в МОБИК с реалистическим взаимодействием Daejeon16: 1 —  $N_{\text{max}} = 2$ ; 2 — 4; 3 — 8; 4 — 14; 5 —  $N_{\text{max}} = 18$

ций вычисление собственных значений матрицы гамильтониана для больших  $N_{\max}$  становится невозможным даже на суперкомпьютерах. Поэтому актуальна проблема экстраполяции результатов, полученных в конечных модельных пространствах, на случай  $N_{\max} \rightarrow \infty$ .

Оценка энергии основного состояния, полученная в [5] путем экстраполяции результатов МОБИК на случай бесконечно больших модельных пространств, составляет  $-28,372$  МэВ; экспериментальное значение энергии связанного состояния —  $-28,296$  МэВ.

## 4 Метод экстраполяции

Как уже упоминалось ранее, методы машинного обучения обычно не используются для задач экстраполяции с малым объемом данных. Тем не менее на их основе может быть построен алгоритм для их решения. Оказывается, для таких задач подходит одновременное обучение множества ИНС.

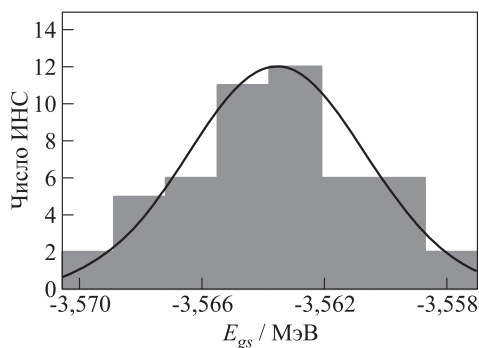
Предсказания, полученные с помощью методов машинного обучения, включают некоторую случайность, связанную с начальными значениями весовых коэффициентов. Чтобы избежать этой случайности, обучается большое число ИНС с одинаковыми гиперпараметрами, но с разными начальными значениями весовых коэффициентов.

Очевидно, что не все обученные ИНС будут давать правильные прогнозы, поэтому важно иметь критерии, позволяющие отфильтровывать заведомо неправильно обученные ИНС.

В модели ядерных оболочек роль такого критерия играет вариационный принцип для энергий связанных состояний, согласно которому энергия основного состояния  $E_{gs}$  должна уменьшаться с ростом  $N_{\max}$  при каждом фиксированном значении  $\hbar\Omega$ .

Для дальнейшей фильтрации, аналогично работе [13], выбирают 50 ИНС с наименьшим среднеквадратичным отклонением предсказанных значений от истинных.

Эмпирически было показано, что зависимость числа ИНС, предсказывающих энергию в каждом интервале, от предсказываемых ими энергий основного состояния  $E_{gs}$  в большинстве случаев представляет собой нормальное распределение. Аппроксимируя их кривой Гаусса, можно получить прогноз метода и его погреш-



**Рис. 3** Гистограмма зависимости числа ИНС, предсказывающих энергию в каждом интервале, от предсказываемых ими энергий основного состояния  $E_{gs}$ : 1 —  $E_{gs}^{\text{пр}} = -3,564$  МэВ;  $\sigma = 0,0029$  МэВ

ность. Среднее значение этого распределения используется как значение, предсказываемое методом, —  $E_{gs}^{pr}$ , а стандартное отклонение распределения  $\sigma$  принимается за погрешность  $\Delta E$  предсказания метода. На рис. 3 приведена такая гистограмма. Гауссоида для данного распределения показана кривой.

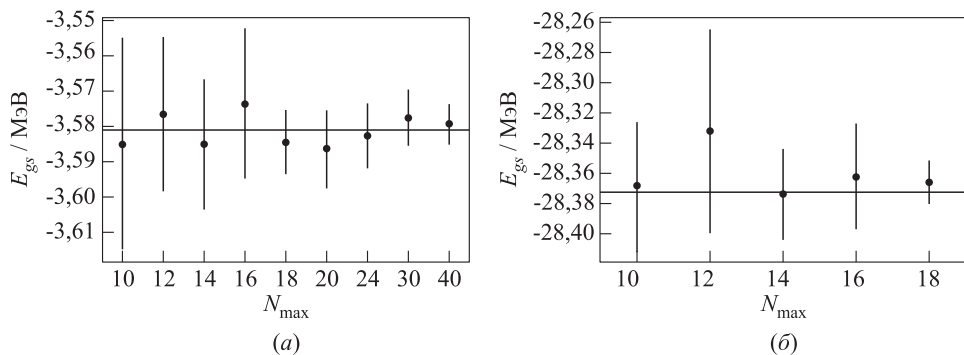
Вышеупомянутый алгоритм был разработан с использованием библиотеки с открытым исходным кодом Keras [15] для языка Python.

## 5 Результаты

С помощью поиска по сетке были получены оптимальные параметры машинного обучения для обеих задач. К этим параметрам относятся: скорость обучения  $\gamma = 0,1$ , размер пакета  $B_S = 64$  и 16 для модельной задачи и задачи нахождения энергии основного состояния ядра  ${}^4\text{He}$  соответственно. Оптимальная архитектура ИНС в обоих случаях — 5 скрытых слоев по 10 нейронов в каждом. Активирующей функцией каждого из слоев, за исключением последнего, является сигмоидальная функция.

Зависимость предсказаний метода от объема данных для модельной задачи показана на рис. 4, а. Точки представляют собой предсказанное значение энергии основного состояния  $E_{gs}$ , вертикальные линии — погрешность предсказания  $\sigma$ . Горизонтальная линия соответствует истинному значению энергии  $E_{gs}^{exp} = -3,581$  МэВ. Процесс обучения проводился на данных вплоть до  $N_{max} = 10, 12, 14$  и так далее. Как видно из рис. 4, а, истинное значение лежит в предсказываемых интервалах для всех размеров набора данных. При увеличении максимального значения  $N_{max}$  предсказания метода изменяются с  $-3,585 \pm 0,030$  до  $-3,579 \pm 0,006$  МэВ для  $N_{max} = 10$  и 40 соответственно.

На рис. 4, б показана аналогичная зависимость для задачи нахождения энергии основного состояния  ${}^4\text{He}$ . Горизонтальная линия — результаты экстраполяции, полученные в [5]. Все предсказанные диапазоны также содержат это



**Рис. 4** Результаты работы алгоритма с разным объемом входных данных для модельной задачи (а) и для задачи нахождения энергии основного состояния  ${}^4\text{He}$  (б);  $\gamma = 0,1$

значение. Предсказания для  $N_{\max} = 10$  и 18 составляют  $-28,368 \pm 0,043$  и  $-28,366 \pm 0,015$  МэВ соответственно.

Разработанный алгоритм хорошо работает даже с небольшими наборами данных. Так, задача нахождения энергии основного состояния  ${}^4\text{He}$  с  $N_{\max} = 10$  содержит около 50 примеров, что является недостаточным объемом данных для корректной работы большинства других алгоритмов на основе машинного обучения.

## 6 Заключение

Предложен новый метод экстраполяции результатов вариационных расчетов в модели ядерных оболочек с осцилляторным базисом на случай бесконечно больших модельных пространств. Метод основан на машинном обучении и позволяет предсказывать значения энергии основного состояния на основе результатов вариационных расчетов, полученных в ограниченном числе модельных пространств  $N_{\max}$  с различными значениями осцилляторного параметра  $\hbar\Omega$ . В разработанном алгоритме используется только программное обеспечение с открытым исходным кодом.

В рамках модельной задачи с известным точным решением исследовалась сходимость результатов при увеличении объема входных данных. Показано, что метод позволяет достичь достаточно высокой точности прогноза даже в случае малых пространств моделей  $N_{\max} \sim 10$ , что особенно важно с точки зрения его применения для анализа результатов, полученных в современных оболочечных подходах. То же утверждение справедливо и для оценки энергии основного состояния ядра  ${}^4\text{He}$ .

Метод позволяет не только получить интересующее значение, но и оценить точность прогнозов. К достоинствам метода можно также отнести универсальность: его можно использовать не только для анализа энергий основного состояния, но и для других характеристик ядер.

## Литература

1. Barrett B. R., Navrátil P., Vary J. P. Ab initio no core shell model // Prog. Part. Nucl. Phys., 2013. Vol. 69. Iss. 1. P. 131–181. doi: 10.1016/j.ppnp.2012.10.003.
2. Stoks V. G. J., Klomp R. A. M., Terheggen C. P. F., de Swart J. J. Construction of high-quality NN potential models // Phys. Rev. C, 1994. Vol. 49. P. 2950–2962. doi: 10.1103/PhysRevC.49.2950.
3. Wiringa R. B., Stoks V. G. J., Schiavilla R. Accurate nucleon–nucleon potential with charge-independence breaking // Phys. Rev. C, 1995. Vol. 51. P. 38–51. doi: 10.1103/PhysRevC.51.38.
4. Machleidt R. High-precision, charge-dependent Bonn nucleon–nucleon potential // Phys. Rev. C, 2001. Vol. 63. Iss. 2. Art. 024001. 32 p. doi: 10.1103/PhysRevC.63.024001.

5. Shirokov A. M., Shin I. J., Kim Y., Sosonkina M., Maris P., Vary J. P. N3LO NN interaction adjusted to light nuclei in ab initio approach // Phys. Lett. B, 2016. Vol. 761. P. 87–91. doi: 10.1016/j.physletb.2016.08.006.
6. Maris P., Vary J. P., Shirokov A. M. Ab initio no-core full configuration calculations of light nuclei // Phys. Rev. C, 2009. Vol. 79. Art. 014308. 15 p. doi: 10.1103/PhysRevC.79.014308.
7. Maris P., Vary J. P. Ab initio nuclear structure calculations of p-shell nuclei with JISP16 // Int. J. Mod. Phys. E, 2013. Vol. 22. Iss. 7. Art. 1330016. 17 p. doi: 10.1142/S0218301313300166.
8. Zhan H., Nogga A., Barrett B. R., Vary J. P., Navrátil P. Extrapolation method for the no-core shell model // Phys. Rev. C, 2004. Vol. 69. Art. 034302. 8 p. doi: 10.1103/PhysRevC.69.034302.
9. Coon S. A., Avetian M. I., Kruse M. K. G., van Kolck U., Maris P., Vary J. P. Convergence properties of ab initio calculations of light nuclei in a harmonic oscillator basis // Phys. Rev. C, 2012. Vol. 86. Art. 054002. doi: 10.1103/PhysRevC.86.054002.
10. Kruse M. K. G., Jurgenson E. D., Navrátil P., Barrett B. R., Ormand W. E. Extrapolation uncertainties in the importance-truncated no-core shell model // Phys. Rev. C, 2013. Vol. 87. Art. 044301. 15 p. doi: 10.1103/PhysRevC.87.044301.
11. More S. N., Ekström A., Furnstahl R. J., Hagen G., Papenbrock T. Universal properties of infrared oscillator basis extrapolations // Phys. Rev. C, 2013. Vol. 87. Art. 044326. 14 p. doi: 10.1103/PhysRevC.87.044326.
12. Furnstahl R. J., More S. N., Papenbrock T. Systematic expansion for infrared oscillator basis extrapolations // Phys. Rev. C, 2014. Vol. 89. Art. 044301. 12 p. doi: 10.1103/PhysRevC.89.044301.
13. Negoita G. A., Vary J. P., Luecke G. R., Maris P., Shirokov A. M., Shin I. J., Kim Y., Ng E. G., Yang C., Lockner M., Prabhu G. M. Deep learning: Extrapolation tool for ab initio nuclear theory // Phys. Rev. C, 2019. Vol. 99. Art. 054308. 11 p. doi: 10.1103/PhysRevC.99.054308.
14. Shirokov A. M., Mazur A. I., Kulikov V. A. On the convergence of oscillator basis calculations // Phys. Atom. Nucl., 2021. Vol. 84. Iss. 2. P. 131–143. doi: 10.1134/S1063778821020149.
15. Keras: The Python Deep Learning library. <https://keras.io>.

*Поступила в редакцию 06.12.21*

---

---

## MACHINE LEARNING FOR EXTRAPOLATION PROBLEMS WITH SMALL DATASET

*A. O. Belozerov and A. I. Mazur*

Pacific National University, 136 Tikhookeanskaya Str., Khabarovsk 680035, Russian Federation

**Abstract:** A new method of extrapolation of variational calculations which is based on training a large number of artificial neural networks with subsequent filtering to select the best networks is proposed. This method is used to evaluate

ground state energy for model problem and to calculate the ground state energy of a  ${}^4\text{He}$  nucleus based on calculations in no-core shell model with realistic potential Daejeon16. The results convergence with increasing number of input data was studied. It is shown that the method allows one to achieve a sufficiently high prediction accuracy even in the case of small model spaces. The method can be applied both in finding various characteristics of nuclei and in solving other problems not related to nuclear physics.

**Keywords:** machine learning; extrapolation methods; ground state energy; nuclear shell model

**DOI:** 10.14357/08696527220202

## Acknowledgments

The paper is published on the proposal of the Program Committee of the 6th International Scientific and Practical Conference “Information Technologies and High Performance Computing” (ITHPC-2021). The work was partially supported by the Russian Foundation for Basic Research (project No. 20-02-00357) and by the Ministry of Science and Higher Education of the Russian Federation (project No. 0818-2020-0005).

## References

1. Barrett, B. R., P. Navrátil, and J. P. Vary. 2013. Ab initio no core shell model. *Prog. Part. Nucl. Phys.* 69(1):131–181. doi: 10.1016/j.pnpnp.2012.10.003.
2. Stoks, V. G. J., R. A. M. Klomp, C. P. F. Terheggen, and J. J. de Swart. 1994. Construction of high-quality NN potential models. *Phys. Rev. C* 49:2950–2962. doi: 10.1103/PhysRevC.49.2950.
3. Wiringa, R. B., V. G. J. Stoks, and R. Schiavilla. 1995. Accurate nucleon–nucleon potential with charge-independence breaking. *Phys. Rev. C* 51:38–51. doi: 10.1103/PhysRevC.51.38.
4. Machleidt, R. 2001. High-precision, charge-dependent Bonn nucleon–nucleon potential. *Phys. Rev. C* 63(2):024001. 32 p. doi: 10.1103/PhysRevC.63.024001.
5. Shirokov, A. M., I. J. Shin, Y. Kim, M. Sosonkina, P. Maris, and J. P. Vary. 2016. N3LO NN interaction adjusted to light nuclei in ab exitu approach. *Phys. Lett. B* 761:87–91. doi: 10.1016/j.physletb.2016.08.006.
6. Maris, P., J. P. Vary, and A. M. Shirokov. 2009. Ab initio no-core full configuration calculations of light nuclei. *Phys. Rev. C* 79:014308. 15 p. doi: 10.1103/PhysRevC.79.014308.
7. Maris, P., and J. P. Vary. 2013. Ab initio nuclear structure calculations of p-shell nuclei with JISP16. *Int. J. Mod. Phys. E* 22(7):1330016. 17 p. doi: 10.1142/S0218301313300166.
8. Zhan, H., A. Nogga, B. R. Barrett, J. P. Vary, and P. Navrátil. 2004. Extrapolation method for the no-core shell model. *Phys. Rev. C* 69:034302. 8 p. doi: 10.1103/PhysRevC.69.034302.

9. Coon, S. A., M. I. Avetian, M. K. G. Kruse, U. van Kolck, P. Maris, and J. P. Vary. 2012. Convergence properties of ab initio calculations of light nuclei in a harmonic oscillator basis. *Phys. Rev. C* 86:054002. 15 p. doi: 10.1103/PhysRevC.86.054002.
10. Kruse, M. K. G., E. D. Jurgenson, P. Navrátil, B. R. Barrett, and W. E. Ormand. 2013. Extrapolation uncertainties in the importance-truncated no-core shell model. *Phys. Rev. C* 87:044301. 15 p. doi: 10.1103/PhysRevC.87.044301.
11. More, S. N., A. Ekström, R. J. Furnstahl, G. Hagen, and T. Papenbrock. 2013. Universal properties of infrared oscillator basis extrapolations. *Phys. Rev. C* 87:044326. 14 p. doi: 10.1103/PhysRevC.87.044326.
12. Furnstahl, R. J., S. N. More, and T. Papenbrock. 2014. Systematic expansion for infrared oscillator basis extrapolations. *Phys. Rev. C* 89:044301. 12 p. doi: 10.1103/PhysRevC.89.044301.
13. Negoita, G. A., J. P. Vary, G. R. Luecke, P. Maris, A. M. Shirokov, I. J. Shin, Y. Kim, E. G. Ng, C. Yang, M. Lockner, and G. M. Prabhu. 2019. Deep learning: Extrapolation tool for ab initio nuclear theory. *Phys. Rev. C* 99:054308. 11 p. doi: 10.1103/PhysRevC.99.054308.
14. Shirokov, A. M., A. I. Mazur, and V. A. Kulikov. 2021. On the convergence of oscillator basis calculations. *Phys. Atom. Nucl.* 84(2):131–143. doi: 10.1134/S1063778821020149.
15. Keras: The Python deep learning library. Available at: <https://keras.io/> (accessed March 28, 2022).

*Received December 6, 2021*

## Contributors

**Belozerov Alexander O.** (b. 2000) — laboratory assistant, Research Laboratory for Modeling Quantum Processes, Pacific National University, 136 Tikhookeanskaya Str., Khabarovsk 680035, Russian Federation; aobelozerov@gmail.com

**Mazur Alexander I.** (b. 1959) — leading scientist, Research Laboratory for Modeling Quantum Processes, Pacific National University, 136 Tikhookeanskaya Str., Khabarovsk 680035, Russian Federation; mazur@khb.ru

## ПОСТРОЕНИЕ СОКРАЩЕННОГО ДЕРЕВА ДОСТИЖИМОСТИ ДЛЯ МОДЕЛЕЙ ПРОГРАММ В ТЕРМИНАХ СЕТЕЙ ПЕТРИ\*

*Д. В. Леонтьев<sup>1</sup>, Д. И. Харитонов<sup>2</sup>*

**Аннотация:** Рассматривается задача построения пространства состояний для анализа поведения императивных программ. При автоматическом построении моделей программ эффект взрыва числа анализируемых состояний составляет основную проблему для поиска ошибок в исходных текстах программ, причем этот взрыв индуцируется за счет состояния множества переменных программы. Предлагается подход к уменьшению числа состояний дерева достижимости моделей программ через отделение модели потока управления программой от моделей переменных и последующего добавления только переменных, влияющих на поток управления, и сокращения состояний этих переменных. Рассмотренный в статье пример показывает, как на практике может быть применен такой подход.

**Ключевые слова:** сети Петри; дерево достижимости; проверка корректности программ; моделирование поведения программ

**DOI:** 10.14357/08696527220203

### 1 Введение

Зависимость современного образа жизни людей от информационных технологий постоянно растет. Растет объем и разнообразие используемого программного обеспечения, которое должно удовлетворять потребностям общества. При этом постоянное появление нового программного обеспечения происходит в условиях, когда процесс качественной проверки отстает от процесса разработки. В результате надежность программного обеспечения, от которого часто зависит жизнь людей, остается под вопросом. Например, из-за ошибки в программном обеспечении аппарата лучевой терапии Therac-25 неправильно рассчитывалась доза облучения и в результате некоторые пациенты получили дозы не в несколько сотен рад, а в 20 000 рад [1] (доза 1000 рад для человека считается несовместимой с жизнью). Другим примером проблем в программном обеспечении могут

---

\*Статья публикуется по представлению программного комитета VI Международной научно-практической конференции «Информационные технологии и высокопроизводительные вычисления» (ИТНПС-2021). Работа выполнена при финансовой поддержке темы госбюджетного задания № 121021700006-0.

<sup>1</sup>Институт автоматизации и процессов управления Дальневосточного отделения Российской академии наук, devozh@dvo.ru

<sup>2</sup>Институт автоматизации и процессов управления Дальневосточного отделения Российской академии наук, demiurg@dvo.ru

послужить две авиакатастрофы Boeing 737 MAX, которые произошли по вине новой системы увеличения маневренных характеристик (MCAS — maneuvering characteristics augmentation system) и повлекли за собой гибель всех пассажиров, находившихся на борту [2]. Из-за ошибок системы автопилот самолета получал неверные данные об угле атаки, вследствие чего считал, что самолет находится на грани сваливания, и опускал нос самолета. Есть еще немало примеров катастрофических ошибок в программном обеспечении, которые могли бы быть обнаружены при надлежащей проверке, что позволило бы избежать трагедий.

В настоящей статье для моделирования программ используются сети Петри, хорошо зарекомендовавшие себя при моделировании параллельных и распределенных систем. В большинстве методов анализа сетей Петри применяется построение графа достижимости, описывающего пространство состояний программы. В сетях Петри, как и в других методах проверки на модели, основной проблемой может стать комбинаторный взрыв числа состояний, так как при построении полного пространства состояний программы необходимо учитывать значение каждой его переменной в каждый момент времени. В настоящей работе предлагается подход к уменьшению числа состояний дерева достижимости моделей программ за счет отделения модели потока управления программой от моделей переменных и последующего добавления только тех переменных, которые влияют на поток управления.

## 2 Пример программы, построение модели

Для построения дерева достижимости в качестве примера рассмотрим программу, изображенную на рис. 1. В этой программе главная функция `main` создает поток `Produce` и вызывает функцию `Consume`. Поток `Produce` вычисляет члены некоторого математического ряда и посылает их в функцию `Consume`, которая вычисляет сумму ряда и определяет момент, когда оба потока должны закончить свое выполнение. Основная функция `main` дожидается завершения обоих потоков и после этого прекращает свое выполнение.

Тестовый пример составлен таким образом, что если поменять местами восьмую и девятую строки в функции `Consume`, то произойдет дедлок в синхронизации потоков — поток `Consume` закончит свое исполнение, в то время как поток `Produce` будет продолжаться. Если рассчитывать на автоматизацию процесса проверки поведения программы, то, несмотря на небольшой код и небольшое число переменных, в зависимости от скорости сходимости выбранного математического ряда число анализируемых состояний программы может достигать произвольно больших чисел.

Построение модели программы автоматизированным способом осуществляется на основе абстрактного семантического графа (ASG — abstract semantic graph), получаемого после синтаксического разбора исходных текстов [3]. При этом формирование М-сети, включающей все значения переменных в исполняемом процессе, нецелесообразно, так как эффект *взрыва числа состояний* сделает

```

1: int main ( )
2: {
3:   pthread_t ThreadId;
4:   double X=0.7;
5:   int errcode = pthread_create (&ThreadId, NULL, Produce, &X);
6:   if (errcode != 0)
7:     return 1;
8:   Consume ( );
9:   void *res;
10:  pthread_join (ThreadId, &res);
11:  return 0;
12: }

```

```

1: double Consume ( )
2: {
3:   bool flagContinue=true;
4:   double value,sum=0.0,lastsum=-1;
5:   while (flagContinue) {
6:     Recv(value);
7:     sum+=value;
8:     flagContinue=(lastsum!=sum);
9:     Send (flagContinue);
10:    lastsum=sum;
11:  }
12:  return sum;
13: }

```

```

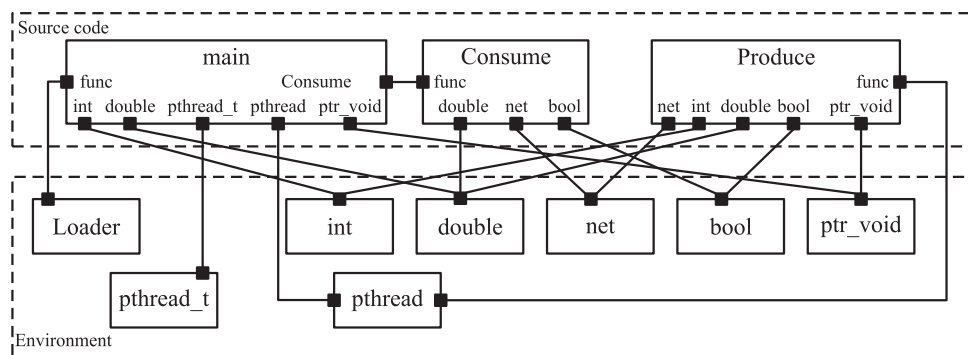
1: #define MakeValue(x,n) x*pow(-1,n)/(2*n-1)
2: void *Produce (void* param)
3: {
4:   bool flagContinue=true;
5:   double value,X=*(double *)param;
6:   int n=1;
7:   while(flagContinue)
8:   {
9:     value=MakeValue(X,n);
10:    n=n+1;
11:    Send(value);
12:    Recv(flagContinue);
13:  }
14:  return NULL;
15: }

```

**Рис. 1** Фрагмент программы на языке C

построение графа достижимости программы фактически невозможным. Поэтому существуют два варианта. Во-первых, можно обесцветить модель программы, полностью удалив из нее данные, но сохранив граф потока управления. Во-вторых, можно дополнительно отделить данные от потока управления программой. Для этого необходимо вынести переменные в отдельные подсети, которые группируются по типам используемых переменными данных. Каждая такая сеть описывает свой тип данных как абстрактный, т. е. выделяя методы и операции, которые можно применить к типу данных, доступ к которым осуществляется через точки доступа.

Исходя из текстов программы, представленных на рис. 1, полная модель программы в композиционном виде должна содержать сети Петри, описывающие функции **main**, **Consume**, **Produce**, а также сети, описывающие среду исполнения. К среде исполнения относится загрузчик программы, сетевое окружение, описывающее функции для пересылки сообщений **Send** и **Recv**, а также библиотека работы с потоками, представленная в программе функциями



**Рис. 2** Композиционное представление модели программы с выделением данных в отдельные сети

`pthread_begin` и `pthread_join`. На рис. 2 представлено композиционное представление модели программы с разделением на модели данных и модели функций, включающее все вышеобозначенные сети.

### 3 Построение дерева достижимости

Для построения дерева достижимости рассмотрим упрощенную модель программы, составленную из функций `Consume`, `Produce` и среды передачи данных, изображенную на рис. 3. В эту модель добавлены стартовые и конечные места для функций, а также обозначена начальная маркировка, позволяющая провести моделирование поведения программы. Переходы модели размечены точками доступа для синхронизации с сетями, моделирующими данные: `bool`, `int`, `double`, `ptr_void`.

Схематично анализ поведения программы можно представить в виде двух шагов: во-первых, необходимо построить пространство состояний, в которых может побывать программа, а во-вторых, проверить на этом пространстве наличие тупиковых состояний, возможность появления необходимых для нормального поведения программы событий (живость переходов), а также отсутствие появления вредных событий, если таковые были заранее обозначены. Представленный ниже классический алгоритм строит итерационно *дерево достижимости*, начиная от начальной маркировки и продолжая добавлять к уже построенным достижимым состояниям сети Петри новые, только если состояния ранее не встречались. Дерево достижимости отличается от графа тем, что в нем не объединяются вершины с одинаковой маркировкой, что позволяет легко построить путь от каждой вершины до стартовой, т. е. начальной маркировки сети.

Классический алгоритм построения дерева достижимости может быть применен к анализируемой модели при условии ее обесцвечивания. С точки зрения



**Рис. 3** Анализируемая модель

---

### Классический алгоритм построения дерева достижимости

```
1: Waiting:={M0};
2: Tree.BeginTree(M0);
3: repeat
4:   remove M from Waiting;
5:   for all (t,M') in Next(M) do
6:     begin
7:       if !Tree.Has(M')    Waiting.Add(M');
8:       Tree.AddArc(M,t,M');
9:     end;
10: until !Waiting.IsEmpty()
```

$Next(M)$  определяется исходя из правил срабатывания переходов следующим образом:  
 $Next(M) = \{(t, M') \mid t \in T \wedge M[t]M'\}$ , где запись  $M[t]M'$  означает, что маркировка  $M'$  получается из маркировки  $M$  после срабатывания перехода  $t$ .

---

анализа это означает, что рассматриваются все ветви исполнения программы, даже такие, которые не могут реально произойти из-за отсутствия необходимого набора значений переменных в пространстве достижимых состояний программы. На рис. 4 представлено дерево достижимости, построенное для обесцвеченной модели. Узел с маркировкой на более темном фоне означает состояние программы, из которого достижимы конечные состояния программы, которыми служат маркировки  $Ce + Pe + N1$ ,  $Ce + P8 + N2$  и  $C3 + Pe + N1$ . Достижение программой первой маркировки отражает факт корректного исполнения программы, достижение последних двух — факт тупиковых состояний программы или, по-другому, факт ее зависания. Анализ обесцвеченной сети позволяет сказать, что, возможно, в коде программы присутствуют варианты некорректного поведения, а также указать путь от начальной маркировки до тупиковых состояний, который далее программист может проверить *вручную*.

Ранее представленный классический алгоритм построения дерева достижимости для сетей Петри может быть естественным образом расширен для М-сетей, однако, как было упомянуто в разд. 2, программа построена таким образом, что при медленной сходимости ряда (в тексте приведен ряд Лейбница) пространство состояний программы будет включать миллиарды значений переменных. При этом с точки зрения поведения программы необходимо всего лишь ответить на вопрос, достижимы ли маркировки  $Ce + Pe + N1$ ,  $Ce + P8 + N2$  и  $C3 + Pe + N1$ .

Основой для построения сокращенного дерева достижимости служит подход к раздельной верификации императивных программ с применением сетей Петри, описанный одним из авторов в 2009 г. [4]. Согласно этому подходу можно проверять работоспособность классов (модулей) в программе, заменяя их модели на протоколы, причем делать это можно поочередно с каждым классом или каждой переменной в программе. Чтобы определить совместимость программы и исполь-

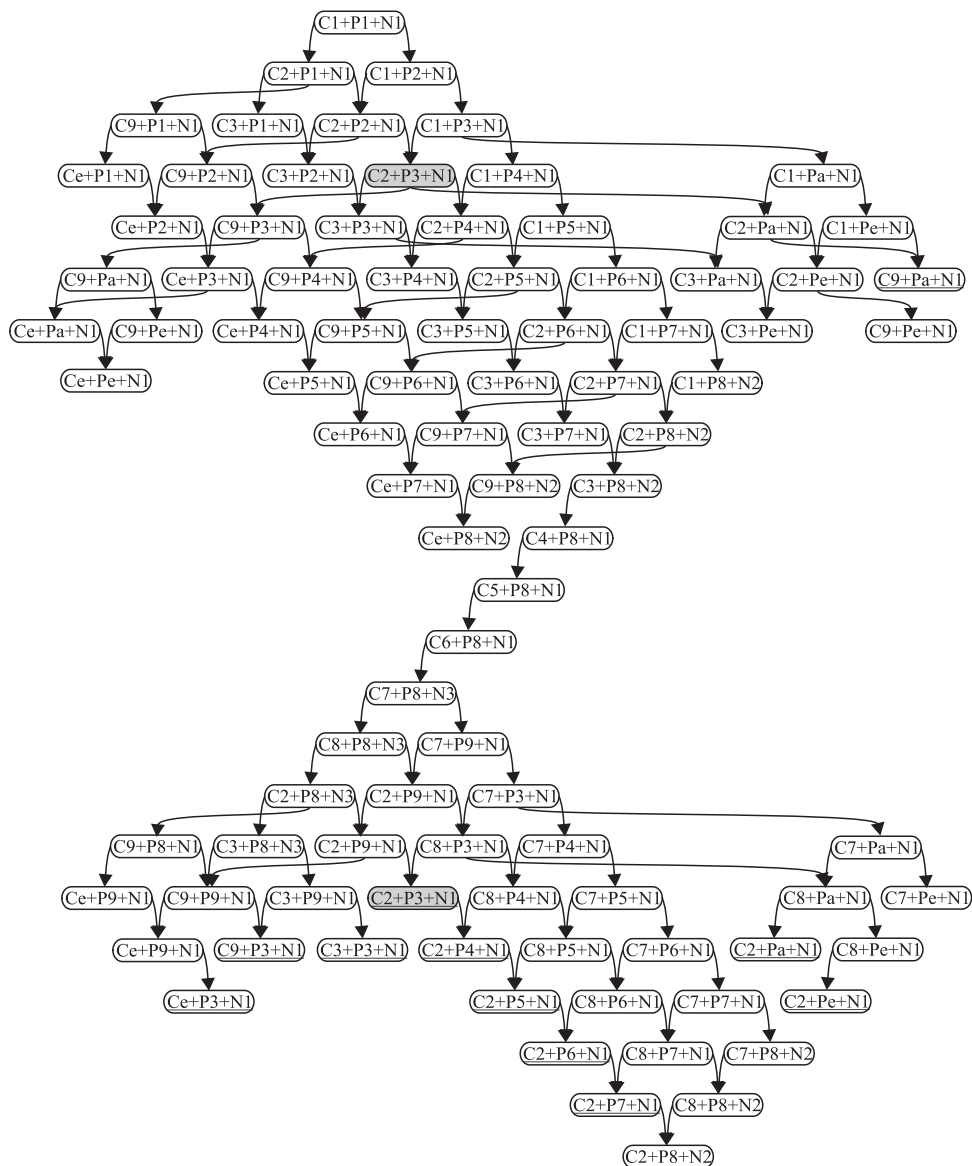
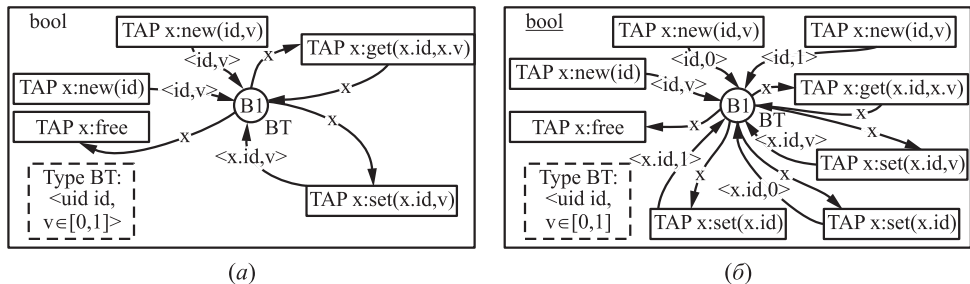


Рис. 4 Дерево достижимости обесцвеченной анализируемой модели

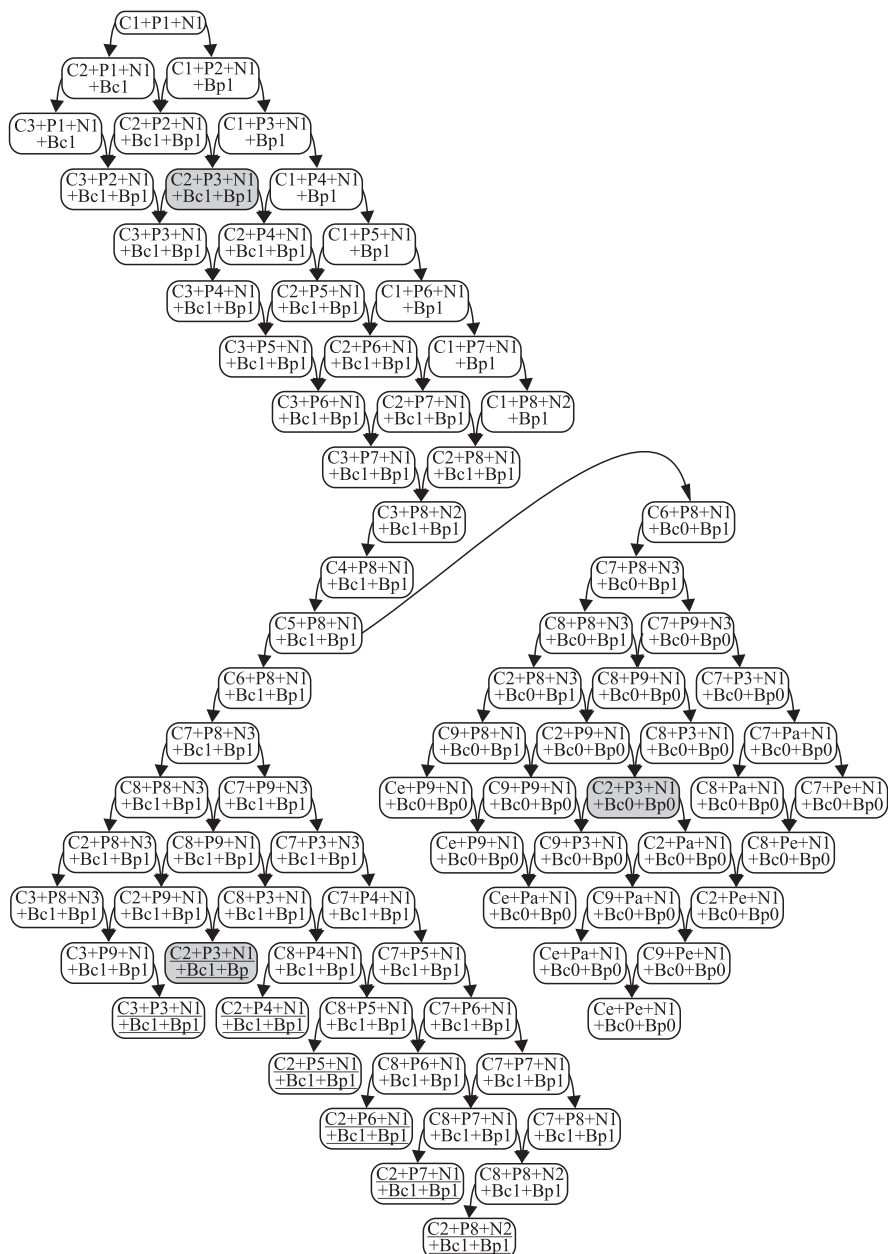


**Рис. 5** Два варианта протокола типа данных: bool (а) и менее детерминированный bool (б)

зуемого ею модуля, необходимо выполнение двух условий. Во-первых, анализ поведения программы, синхронизированной с протоколом модуля, не должен содержать ошибок, а во-вторых, протокол модуля должен описывать поведение, допускающее в качестве одного из вариантов поведения модели реализации модуля. Фактически понятие протокола должно быть менее детерминированным, менее детализированным, т. е. предлагать больше вариантов исполнения, чем реальный модуль. Таким образом получается, что *если иметь набор протоколов разной степени детализации для каждого из основных типов данных, используемых в программе, то можно проверить, насколько программа зависит от конкретной переменной или конкретного типа данных*.

В случае с анализируемой программой можно заменить детерминированный протокол типа данных bool (в котором каждая операция присвоения должна содержать присваиваемое переменной значение) на менее детерминированный протокол, в котором имеются операции присвоения с указанием и без указания присваиваемого переменной значения. На рис. 5 изображены оба варианта модели типа данных bool. На рис. 5, б приведен менее детерминированный протокол, в котором операция присвоения set(x.id) случайным образом присваивает переменной значение 0 или 1.

На рис. 6 изображено дерево достижимости анализируемой модели, синхронизированной с менее детерминированным протоколом типа данных bool. К маркировкам состояний дерева достижимости 4 добавлены токены в месте B1 протокола типа данных, для краткости обозначенные как Bc1, Bc0 для переменной в части сети, относящейся к потоку управления из функции Consume, и Bp1, Bp0 — из функции Produce. Видно, что состояние  $Ce + Pe + N1$  из обесцвеченной сети достигается при маркировке  $Ce + Pe + N1 + Bc0 + Bp0$  в более полной модели. Маркировки  $Ce + P8 + N2$  и  $C3 + Pe + N1$  недостижимы в более полной модели, и это означает, что цель построения пространства состояний достигнута при анализе поведения программы с использованием протокола двух переменных одного типа данных. Таким образом, дерево достижимости на рис. 6 описывает полностью поведение анализируемой программы и при этом содержит минимальный набор данных, необходимый для такого описания.



**Рис. 6** Дерево достижимости анализируемой модели, синхронизированной с протоколом типа данных bool

## 4 Заключение

Сети Петри как средство моделирования распределенных систем не так часто применяется для построения моделей программ [5]. В большинстве случаев их применение ограничено этапом дизайна программных систем и либо не доходит до стадии проверки корректности программ, либо ограничивается довольно примитивными моделями. Между тем, как продемонстрировано в работах авторов, сети Петри могут быть использованы для автоматизированного построения моделей программ [3], где размеры моделей могут достигать сотен тысяч мест и переходов. Технически проверка таких моделей программ может выполняться с применением темпоральных логик и символьной верификации [6], однако построение пространства состояний программы выглядит более перспективным, так как поддается полной автоматизации.

К сожалению, при полном моделировании работы моделей программ неизбежно возникает проблема *взрыва числа состояний* (state explosion problem), так как в общем случае размерность пространства состояний сильно зависит от размера модели, числа моделируемых процессов, числа используемых переменных, а также от количества значений, которые могут принимать переменные. В результате прямой подход к построению пространства состояний перестает работать уже для самых простых программ. Поэтому важным направлением развития методов верификации не только программ, но и распределенных систем в целом становится разработка методов сокращения числа состояний. С конца XX в. среди таких методов выделяют: метод упрямых множеств (stubborn set method), использование двоичных разрешающих диаграмм (symbolic binary decision diagrams), методы, основанные на частичном порядке (partial order methods), а также использование симметрии и эквивалентности на рассматриваемых моделях [7].

В контексте использования сетей Петри при верификации распределенных систем для сокращения числа состояний могут быть применены методы, основанные на частичном порядке, а более конкретно — развертки моделей, введенные МакМилланом в 1993 г. [8], и понятие ветвящихся процессов [9]. В отличие от графа или дерева достижимости, где поведение системы описано в виде всех возможных последовательностей срабатывания переходов, развертки и ветвящиеся процессы описывают поведение системы в рамках истинного параллелизма, разворачивая по-отдельности поведение всех независимо существующих процессов. Таким образом, развертки сетей представляют сокращенное представление дерева достижимости за счет уменьшения числа дублирующихся состояний, описывающих независимо исполняющиеся процессы. Развертки могут применяться и к раскрашенным сетям Петри, более подходящим для описания и анализа программ [10]. Однако, например, развертка рассмотренного в настоящей статье примера по-прежнему будет содержать миллиарды состояний, так как развертки сетей не могут сократить число состояний используемых в программе переменных. Предложенный же в настоящей статье подход направлен на сокращение

числа состояний, связанных с использованными переменными. В дальнейшем авторы надеются адаптировать его и для построения более кратких разверток моделей программ.

В предложенном подходе в описании моделей императивных программ в терминах сетей Петри переменные выделяются в отдельные сети, моделирующие типы данных. При таком выделении появляется возможность построения *обесцвеченного* дерева достижимости для потока управления программой без учета данных, на котором могут быть обнаружены корректные и потенциально опасные пути исполнения программ. Построение протоколов типов данных с различной детализацией позволяет проверить достижимость потенциально опасных состояний, не учитывая значения всех переменных программы. Дерево достижимости, дополненное значениями переменных, позволяющих достоверно определить достижимость всех его состояний, представляет собой сокращенное дерево достижимости программы.

## Литература

1. *Leveson N. G., Turner C. S.* An investigation of the Therac-25 accidents // *Computer*, 1993. Vol. 26. No. 7. P. 18–41.
2. *Mako S., Pilat M., Svab P., Kozuba J., Cicvakova M.* Evaluation of MCAS system // *Acta Avionica J.*, 2019. Vol. 21. No. 1. P. 21–28.
3. *Kharitonov D., Tarasov G.* Modeling function calls in program control flow in terms of Petri nets // *ACSIJ*, 2014. Vol. 3. No. 6. P. 82–91.
4. *Харитонов Д. И.* Раздельная верификация объектно-ориентированных программ с построением протокола C++ класса в терминах сетей Петри // *Моделирование и анализ информационных систем*, 2009. Т. 16. № 1. С. 92–112.
5. *Denaro G., Pezze M.* Petri nets and software engineering // *DAIMI Report Series*, 2003. Vol. 29. No. 1. P. 439–466.
6. *Latvala T., Makela M.* LTL model checking for modular Petri nets // *Applications and theory of Petri nets* / Eds. J. Cortadella, W. Reisig. — *Lecture notes in computer science ser.* — Springer, 2004. Vol. 3099. P. 298–311.
7. *Valmari A.* The state explosion problem // *Lectures on Petri nets I: Basic models* / Eds. W. Reisig, G. Rozenberg. — *Lecture notes in computer science ser.* — Springer, 1998. Vol. 1491. P. 429–528.
8. *McMillan K. L.* Using unfoldings to avoid the state explosion problem in the verification of asynchronous circuits // *Computer aided verification* / Eds. G. von Bochmann, D. K. Probst. — Berlin, Heidelberg: Springer, 1993. P. 164–177.
9. *Engelfriet J.* Branching processes of Petri nets // *Acta Inform.*, 1991. Vol. 28. No. 6. P. 575–591.
10. *Козюра В. Е.* Реализация системы проверки моделей раскрашенных сетей Петри с использованием разверток. — Новосибирск: ИСИ им. А. П. Ершова СО РАН, 2002. 44 р.

Поступила в редакцию 30.11.21

## CONSTRUCTING OF THE BRIEF REACHABILITY TREE FOR PROGRAM MODELS IN TERMS OF PETRI NETS

D. V. Leontyev and D. I. Kharitonov

Institute of Automation and Control Processes, Far-Eastern Branch of the Russian Academy of Sciences, 5 Radio Str., Vladivostok 690041, Russian Federation

**Abstract:** The article deals with the problem of building a state space for analyzing the imperative programs behavior. The state explosion problem of analyzed states number in the automatic program models construction is the main problem for finding errors in the programs source code. This explosion is induced mainly due to the composition of the program variables states. The article proposes an approach to reducing the number of states of the reachability tree of program models by separating the program control flow model from the variable models and then adding only variables that affect the control flow and reducing the states of these variables. The example considered in the article shows how such an approach can be applied in practice.

**Keywords:** Petri nets; reachability tree; checking programs correctness; modeling program behavior

**DOI:** 10.14357/08696527220203

### Acknowledgments

The paper is published on the proposal of the Program Committee of the VI International Scientific and Practical Conference “Information Technologies and High Performance Computing” (ITHPC-2021). The research was supported by state funded program No. 121021700006-0.

### References

1. Leveson, N. G., and C. S. Turner. 1993. An investigation of the Therac-25 accidents. *Computer* 26(7):18–41.
2. Mako, S., M. Pilat, P. Svab, J. Kozuba, and M. Cicvakova. 2019. Evaluation of MCAS system. *Acta Avionica J.* 21(1):21–28.
3. Kharitonov, D., and G. Tarasov. 2014. Modeling function calls in program control flow in terms of Petri nets. *ACSIJ* 3(6):82–91.
4. Kharitonov, D. 2009. Razdel'naya verifikatsiya ob"ektno-orientirovannykh programm s postroeniem protokola C++ klassa v terminakh setey Petri [Separable object-oriented program verification with C++ class protocol definition in terms of Petri nets]. *Modelirovanie i analiz informatsionnykh system* [Modeling and Analysis of Information Systems] 16(1):92–112.
5. Denaro, G., and M. Pezze. 2003. Petri nets and software engineering. *DAIMI Report Series* 29(1):439–466.

6. Latvala, T., and M. Makela. 2004. LTL model checking for modular Petri nets. *Applications and theory of Petri nets*. Eds. J. Cortadella and W. Reisig. Lecture notes in computer science ser. Springer. 3099:298–311.
7. Valmari, A. 1998. The state explosion problem. *Lectures on Petri nets I: Basic models*. Eds. W. Reisig and G. Rozenberg. Lecture notes in computer science ser. Springer. 1491:429–528.
8. McMillan, K. L. 1993. Using unfoldings to avoid the state explosion problem in the verification of asynchronous circuits. *Computer aided verification*. Eds. G. von Bochmann and D. K. Probst. Berlin, Heidelberg: Springer. 164–177.
9. Engelfriet, J. 1991. Branching processes of Petri nets. *Acta Inform.* 28(6):575–591.
10. Kozyura, V. E. 2002. *Realizatsiya sistemy proverki modeley raskrashennykh setey Petri s ispol'zovaniem razvertok* [Implementation of a system for checking models of colored Petri nets using developments]. Novosibirsk: A. P. Ershov IIS SB RAS. 44 p.

*Received November 30, 2021*

## Contributors

**Leontyev Denis V.** (b. 1992) — scientist, Institute of Automation and Control Processes, Far-Eastern Branch of the Russian Academy of Sciences, 5 Radio Str., Vladivostok 690041, Russian Federation; devozh@dvo.ru

**Kharitonov Dmitriy I.** (b. 1973) — Candidate of Science (PhD) in technology, senior scientist, Institute of Automation and Control Processes, Far-Eastern Branch of the Russian Academy of Sciences, 5 Radio Str., Vladivostok 690041, Russian Federation; demiurg@dvo.ru

## К ВОПРОСУ О СОВМЕЩЕНИИ ПРОГРАММНОГО ОПРЕДЕЛЕНИЯ СЕТЕЙ С ВИРТУАЛИЗАЦИЕЙ СЕТЕВЫХ ФУНКЦИЙ

*В. Б. Егоров<sup>1</sup>*

**Аннотация:** Многие фирмы, поставляющие программные продукты на сетевой рынок, предлагают средства для программного определения сетей (software-defined networking, SDN) и виртуализации сетевых функций (network functions virtualization, NFV). Частое соседство в публикациях и рекламе обоих акронимов, вплоть до их объединения в SDN/NFV, создает впечатление, что за ними стоит что-то близкородственное, если вообще не одно и то же. Выяснение же действительного соотношения сетевых технологий SDN и NFV затрудняется сложностями с пониманием самих предметов сопоставления. В публикациях по этим предметам закрепились часто повторяемые утверждения, многие из которых при внимательном рассмотрении оказываются сомнительными. Настоящая статья ставит целью показать, что декларируемая схожесть обеих технологий во многом иллюзорна, они развиваются в разных направлениях и разными темпами: NFV востребована и активно распространяется в провайдерских сетях, в то время как SDN нашла воплощение главным образом в сетях крупных интернет-компаний и интерес к ней последнее время не растет, чему есть целый ряд объективных причин. В рекламном соседстве обеих технологий, вплоть до видимости их объединения комплексным акронимом SDN/NFV, заинтересованы главным образом фирмы-поставщики сетевых программных продуктов по сугубо маркетинговым соображениям.

**Ключевые слова:** виртуализация сетевых функций; программно определяемая сеть; NFV; SDN

**DOI:** 10.14357/08696527220204

### 1 Введение

В рекламе и презентациях многих ведущих фирм, поставляющих программное обеспечение (ПО) на сетевой рынок, присутствуют продукты для программного определения сетей и виртуализации сетевых функций. Причем частое соседство обоих акронимов в разного рода публикациях, вплоть до их объединения в SDN/NFV, создает впечатление родства или как минимум тесной взаимосвязи стоящих за ними технологий.

Поскольку обе рассматриваемые сетевые технологии входят в топ модных трендов, вполне понятна активная реклама фирмами продуктов, имеющих отно-

---

<sup>1</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук, VEgorov@ipiran.ru

шение к этим технологиям, с целью захвата перспективного рынка. Объединение двух модных акронимов в одном рекламном ролике и даже в одном продукте может преследовать цель и экономии на рекламе, и охвата максимального числа потенциальных покупателей без раздувания номенклатуры предлагаемых продуктов.

Однако не только в мире рекламы, но и в экспертной среде можно встретить рассуждения о некой схожести SDN и NFV, основными чертами которой, в частности, называют [1]:

- упрощение развертывания сетей и управление ими;
- обещание замены дорогого проприетарного оборудования на ПО, работающее на недорогих массово выпускаемых серверах;
- ожидание снижения капитальных и эксплуатационных затрат.

В настоящей статье ставится под сомнение обоснованность названных упрощений, обещаний и ожиданий, а также оцениваются действительная степень близости технологий SDN и NFV и перспектива их объединения в некую комплексную сетевую технологию SDN/NFV.

## **2 Сложности с пониманием предметов сопоставления**

В относящихся к SDN и NFV публикациях закрепились некоторые часто повторяемые утверждения, многие из которых при внимательном рассмотрении не могут не вызвать сомнений. Так, согласно мнению специалистов Центра прикладных исследований компьютерных сетей, «SDN и NFV относятся к технологиям компьютерных сетей нового поколения, обе могут сосуществовать в одной сетевой среде и имеют много одинаковых характеристик и компонентов. Оба подхода имеют схожую цель: уменьшить сложность сетей, обеспечить масштабируемость и автоматизацию управления, повысить мощности физической инфраструктуры сетей с наложением виртуальной, упростить развертывания, программно устранять неполадки и, конечно, снизить капитальные и операционные затраты» [2]. Если опустить постулируемую в этом заявлении априорно очевидную возможность существования сетевых технологий в сетевой среде, остается совершенно непонятным, каким образом SDN или NFV могут уменьшить сложность сетей. Если под не имеющим смысла термином «мощности физической инфраструктуры сетей» подразумевается их общая пропускная способность (а что можно подразумевать еще?), то наложение виртуальных сетей на физическую инфраструктуру ни на йоту не увеличивает «мощность» последней. Ее фактическая пропускная способность может только снизиться вследствие неизбежных накладных расходов отражения манипуляций с удобными виртуальными абстракциями на куда менее податливую инфраструктурную реальность или из-за недостаточной реактивности SDN-контроллера. Снижение капитальных и операционных затрат от внедрения SDN не подтверждено никакими реальными практиками, и, по-видимому, в значительной мере на декларативном уровне оно остается и для NFV.

Единственной безусловной общей чертой обоих, SDN и NFV, подходов к построению сетей остается программная автоматизация управления как физической сетевой инфраструктурой, так и виртуальными образованиями, действительно потенциально способная, устранив оператора и динамично управляя ресурсами и сервисами по запросам, более эффективно подстраивать их под текущие задачи.

Можно найти попытки выявить общность SDN и NFV в некоем абстрагировании: «Коренная общность SDN и NFV в том, что в обоих случаях используется абстрагирование: SDN разделяет сетевые функции продвижения пакетов и управления сетью, а NFV абстрагирует те и другие функции от оборудования, на котором они выполняются. Таким образом, SDN и NFV используют виртуализацию для абстрагирования сетевой инфраструктуры в ПО и последующей реализации в ПО аппаратной платформы» [3]. Но в чем заключается абстрагирование при разделении сетевых функций, которые разделены в любой сети без всякого абстрагирования? Как вообще можно абстрагировать функцию, которая сама по себе уже есть абстракция? Получается, что при внимательном рассмотрении и тут никакой «коренной общности SDN и NFV» найти не удастся, а всего лишь подчеркивается неудачность термина NFV, на деле имеющего в виду виртуализацию не сетевых функций, что бессмысленно, а сетевых устройств: нельзя виртуализировать коммутацию, но давно общей практикой стали виртуальные коммутаторы [4].

Морока с терминологией в области «виртуализации сетевых функций» носит, к сожалению, массовый характер даже в научных кругах. Примером может служить следующее определение: «NFV — технология виртуализации физических сетевых элементов телекоммуникационной сети, когда сетевые функции исполняются программными модулями, работающими на стандартных серверах...» [5]. Здесь сначала совершенно справедливо акцентируется виртуализация не функций, а «физических сетевых элементов», но далее почему-то «программными модулями» исполняются не функции этих физических сетевых элементов, что должно было бы следовать из предыдущего акцента, а опять-таки «сетевые функции». Между тем такая основополагающая сетевая функция, как передача данных между абонентами сети, находящимися на разных серверах, не может «исполниться» никакими «программными модулями» при отсутствии соответствующих физических сетевых элементов.

Если стремление фирм-поставщиков сетевого ПО любой ценой объединить SDN и NFV можно оправдать маркетинговыми соображениями, то сложнее понять такие заявления теоретиков-экспертов, как «NFV может обеспечить SDN инфраструктурой, на которой работает программное обеспечение SDN» [5]. Но все ПО SDN работает в одном-единственном месте — контроллере сети, и этот инфраструктурный компонент не может быть обеспечен никакой виртуализацией, включая NFV, а только поставщиком соответствующего серверного оборудования. Таким образом, пересечения, а тем более слияния SDN с NFV не наблюдается и в этом случае.

Единственную реальную схожесть SDN и NFV можно увидеть в нацеленности на замену дорогих проприетарных сетевых устройств дешевым унифицированным оборудованием: «цели NFV близки к целям SDN в части использования обычных стандартных серверов и коммутаторов» [5]. Однако и это скорее видимость близости. Одно и то же оборудование применяется по-разному и с различными несоприкасающимися целями: SDN предполагает использование «стандартных серверов» для контроллеров сети, NFV — для эмуляции традиционных сетевых устройств. А «стандартные коммутаторы» вообще никому не нужны: SDN предполагает их замену на специальные OpenFlow-коммутаторы, в то время как NFV заменяет их виртуальными.

### 3 Сопоставление SDN и NFV

Для оценки схожести и различия SDN и NFV полезно сопоставить наиболее важные их характеристики. Такое сопоставление, сделанное, в частности, в таблице, наглядно показывает, как мало общего между этими технологиями. Одним из немногих объединяющих обе колонки таблицы пунктов оказываются поставщики сетевого ПО в качестве инициаторов инноваций. Поэтому неудивительно, что именно у них наиболее выражена тенденция к объединению SDN и NFV, пусть всего лишь как средству универсализации продуктов и рекламы для удовлетворения более широкого круга потенциальных покупателей.

В действительности у SDN и NFV очень мало общего, а их декларируемая схожесть в целом иллюзорна. Обе технологии развиваются самостоятельно по

Сопоставление SDN и NFV

| Характеристика      | SDN                                                                                     | NFV                                                                                      |
|---------------------|-----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------|
| Основное назначение | Централизация и программируемость управления сетью                                      | Замена специализированного оборудования сетевых устройств на ПО                          |
| Стратегия           | Физическое разделение плоскостей данных и управления, консолидация плоскости управления | Программная реализация сетевых функций в типовых массовых серверах (виртуальных машинах) |
| Основные применения | Датацентры, корпоративные сети больших предприятий, облачные платформы                  | Сети телекоммуникационных операторов и сервис-провайдеров                                |
| Инициаторы          | Владельцы датацентров, поставщики сетевого ПО                                           | Провайдеры телеком-сервисов, поставщики ПО                                               |
| Целевые устройства  | Универсальные серверы и специальные коммутаторы                                         | Универсальные серверы, виртуальные устройства                                            |
| Уровни OSI          | 2...4                                                                                   | 4...7                                                                                    |
| Протоколы           | OpenFlow и аналогичные                                                                  | Нет                                                                                      |
| Стандартизатор      | ONF                                                                                     | ETSI                                                                                     |

разным направлениям и разными темпами. Эту объективную реальность не могут изменить ни реклама ведущих игроков сетевого рынка, ни звучащие в ее поддержку призывы ангажированных экспертов к объединению SDN и NFV: «Можно использовать технологию SDN без использования NFV, и наоборот. Но эти технологии взаимодополняющие, и их раздельное использование ограничивает возможности выигрыша для бизнеса. В конечном счете для получения конкурентных преимуществ объединение технологий SDN и NFV становится необходимым» [6]. В такого рода призывах бездоказательные утверждения о взаимодополняемости и безосновательные обещания конкурентных преимуществ не могут дезавуировать существо дела, вынужденно констатируемое в первой фразе цитаты. Это существо находит и другие выражения: «Может ли существовать SDN без NFV? Конечно, может... Может ли существовать NFV без SDN — ответ тот же» [7] или «цели NFV могут быть достигнуты и без использования механизмов SDN, средствами, которые уже имеются во многих дата-центрах» [5]. К последнему еще необходимо добавить, что если NFV не нуждается в SDN, то SDN тем более может обойтись без NFV, поскольку появилась и обрела популярность значительно раньше последней.

#### **4 Попытки и трудности объединения SDN и NFV**

Более-менее естественное объединение различных сетевых технологий, SDN и NFV в частности, теоретически могло бы произойти в рамках гипотетического единого «сетевцентрического» подхода, основанного на общих, открытых и совместимых стандартах. Вот как, например, представляет условия воплощения такого подхода сайт TAdviser:

- универсализация телекоммуникационного оборудования и сетевых сервисов благодаря общим спецификациям сетевой инфраструктуры;
- упрощение внедрения технологии NFV благодаря общим рекомендациям и шаблонам;
- стандартизация интерфейсов прикладного программирования (application programming interfaces, API), которая позволит программным продуктам разных поставщиков взаимодействовать друг с другом.

В этом перечне условий ключевая роль отводится стандартизации: нужны общие спецификации, готовые шаблоны и стандартные API. Уловив эту модальность, в игру включились стандартизирующие организации. Так, ETSI (European Telecommunications Standards Institute) в спецификации NFV Release 2 сделал попытку встроить SDN-контроллер в инфраструктуру NFV в нескольких разных качествах [8]:

- физической сетевой функции (physical network function);
- виртуализированной сетевой функции (virtualized network function);

- менеджера виртуализированной инфраструктуры (virtualized infrastructure manager);
- контроллера в инфраструктуре виртуализации сетевых функций (network function virtualization infrastructure);
- контроллера в системе поддержки операций и бизнеса (operation/business support system).

В этой связи одной из целей дальнейших исследований ETSI определяет возможный интерфейс между контроллером SDN и оркестровщиком NFV. Тогда конечной целью усилий ETSI можно предположить возможность администрировать и оркестровать виртуальные и физические сети SDN/NFV на единой платформе [9].

Но, несмотря на активную рекламу продавцов и усилия стандартизаторов, в последнее время в хоре дифирамбов обоим рассматриваемым технологиям и их обещанному поставщиками и ожидаемому потребителями объединению все чаще слышны критические голоса. Постепенно накапливающийся опыт выявил основные факторы, мешающие развитию технологического альянса SDN/NFV [10]:

- фрагментарность решений. Наряду с ONF (Open Networking Foundation) и ETSI имеются другие вырабатывающие стандарты организации. Они работают без должной координации, а иногда в режиме конкуренции;
- замешательство в среде вендоров. Они противятся SDN/NFV из нежелания отказываться от годами наработанных решений и боязни того, что вместо покупки у них оборудования пользователи будут просто скачивать виртуализированные устройства с какого-нибудь GitHub. Некоторые из них создают собственные конкурирующие продукты SDN и NFV, способные работать только на их оборудовании;
- организационная трансформация. Переход к NFV/SDN требует организационной перестройки, но устоявшиеся бюрократические структуры успешно сопротивляются организационным новшествам. Персонал тоже нередко противится инновациям из-за нежелания переучиваться;
- недостаточная проработанность бизнес-стратегий. Ожидаемые от технических преимуществ SDN/NFV экономические выгоды трудно просчитать заранее из-за отсутствия подтвержденных практикой бизнес-моделей. Переход на новую технологию — процесс затратный, а окупаемость стратегии SDN/NFV может занять многие годы.

Особенно настораживает тот факт, что интерес к SDN перестал расти с 2008 г. [7]. Опасения и даже пессимистические прогнозы в отношении будущего технологии SDN/NFV, ранее крайне осторожно высказывавшиеся отдельными авторами, на конференции OpenStack summit 2018 года прозвучали открытым текстом и были неожиданно поддержаны со стороны стандартизирующих организаций. Неудивительно, что в этих условиях некоторые крупные компании,

такие как AT&T или Verizon, выразили намерение приступить к разработкам собственных решений в области SDN/NFV.

## 5 Разные пути SDN и NFV

Концепцию SDN породили теоретики, а ее практическое применение буквально навязывалось владельцам датацентров. В противовес SDN идея NFV была рождена и успешно продвигается практиками — консорциумом сервис-провайдеров, «которые обнаружили, что аппаратные решения сетевых устройств препятствуют быстрому развертыванию новых сетевых сервисов и проблему можно решить с помощью технологий виртуализации и NFV в частности» [11]. Отсюда вытекает главное различие между SDN и NFV: у технологии NFV есть свой заинтересованный потребитель, и она развивается, в то время как SDN своего массового потребителя не нашла и стагнирует из-за целого ряда неопределенностей [1]:

- относительные успехи крупных компаний, таких как Google, Facebook или Amazon, не удастся воспроизвести в условиях датацентров меньших масштабов;
- ПО SDN остается областью большой неопределенности и потенциальных рисков при небольшом числе надежно удерживающихся на рынке поставщиков;
- растущий интерес компаний к уходу в «облака» уменьшает у них стимулы к коренному обновлению своей устаревающей инфраструктуры и ПО;
- появление многосерверных шасси с широкополосным сетевым соединением, способным обслуживать десятки и сотни серверов, а также перспектива деагрегированных серверов решают многие сетевые проблемы и объективно уменьшают потребность в SDN.

В дополнение ко всем неопределенностям очень больно по «имиджу» SDN может ударить появившееся в последнее время понимание того, что практически нереализуемым оказался один из ее главных козырей — широко разрекламированная возможность замены разнообразного дорогого проприетарного сетевого оборудования на простые и дешевые унифицированные коммутаторы. По ряду объективных причин в реальности имеющиеся сегодня SDN-сети остаются все теми же IP-сетями с традиционными маршрутизаторами в узлах, лишь дополненными поддержкой протокола OpenFlow [12]. Наконец, несмотря на обилие рекламы SDN, никто из ведущих поставщиков сетевого оборудования и ПО не заявляет программное определение сетей главной целью своего технологического развития [13].

В итоге остаются нерешенными такие базовые вопросы, как пути перехода к SDN, перенос идей SDN в транспортные сети и реализация сетевых сервисов в «облачной» парадигме вычислений [14]. Время играет на NFV, но, похоже,

против SDN. А тем самым и против альянса этих технологий, все больше напоминающего басенную «упряжку» лебедя с раком и щукой.

## 6 Заключение

Все попытки выявить черты схожести технологий SDN и NFV, включая и их недлинный перечень во введении к настоящей статье, обходят стороной технические аспекты. Схожесть отмечается лишь в целях внедрения и ожидаемых экономических выигрышах, которые в конечном счете всегда одинаковы для всех технологических новшеств: снижение расходов и увеличение доходов. И это не случайно. Содержательное сопоставление технологий SDN и NFV затрудняется неоднозначностью понимания самих предметов сопоставления как экспертным сообществом, так и участниками сетевого рынка. Декларируемые схожие черты SDN и NFV при внимательном рассмотрении оказываются лишь видимостью схожести, а попытки обосновать необходимость их объединения неубедительны.

Сопоставление технологий выявляет у них мало содержательно общего. Они преследуют разные цели, развиваются разными путями, распространяются с разной скоростью и не испытывают тяги друг к другу. В их объединении, на сегодняшний день искусственном, заинтересованы главным образом крупные фирмы-поставщики сетевого ПО по маркетинговым соображениям и, может быть, стандартизирующие организации — по престижным. Поэтому реклама фирм-поставщиков сетевого ПО пестрит предложениями SDN/NFV и дышит оптимизмом. Однако для рядового пользователя, который, вероятно, был бы не прочь получить готовый комплексный работоспособный продукт, радикально меняющий технические характеристики и улучшающий экономические показатели его сети, все далеко не так оптимистично. Готовых программных решений для превращения его сети в SDN-сеть на рынке нет, и перспективы их появления остаются весьма туманными. Ситуация с NFV-решениями лучше, но и тут комплексные продукты, такие, например, как инфраструктура виртуализации сетевых функций с менеджером виртуализированной инфраструктуры, предлагаются только двумя программными платформами: открытой OpenStack и vSphere компании VMware.

В итоге технологии SDN и NFV идут каждая своим путем, причем первая, стартовав раньше, все больше отстает от второй, что также усложняет их союз и оставляет проблематичным появление в ближайшем будущем доступных массовому пользователю комплексных продуктов для SDN, а тем более SDN/NFV.

## Литература

1. *Fruehe J.* The battle of SDN vs. NFV // MoorInsights&Strategy. <https://moorinsightsstrategy.com/the-battle-of-sdn-vs-nfv>.
2. «Ростелеком» и ЦПИКС стали партнерами в области развития SDN и NFV // Сnews, 24 февраля 2015 г. <https://www.arccn.ru/media/1449/>.

3. Tittel E. SDN vs. NFV: What's the difference? // Cisco. <https://www.cisco.com/c/en/us/solutions/software-defined-networking/sdn-vs-nfv.html>.
4. Егоров В. Б. Некоторые неявные вопросы сетевой виртуализации // Системы и средства информатики, 2017. Т. 27. № 3. С. 88–98.
5. Шалагинов А. SDN и NFV: как это работает на сети оператора связи // Telecom&IT, 27 декабря 2015 г. <https://shalaginov.com/2015/12/27/sdn-%D0%B8-nfv-%D0%BA%D0%B0%D0%BA-%D1%8D%D1%82%D0%BE-%D1%80%D0%B0%D0%B1%D0%BE%D1%82%D0%B0%D0%B5%D1%82-%D0%BD%D0%B0-%D1%81%D0%B5%D1%82%D0%B8-%D0%BE%D0%BF%D0%B5%D1%80%D0%B0%D1%82%D0%BE%D1%80%D0%B0/>.
6. Ganczareno J. The difference between SDN and NFV — a simple guide // Codilime, July 28, 2020.
7. Сушков А. SDN & NFV и при чем тут Облака // Хабр, 28 ноября 2016 г.
8. Odini M.-P. SDN in NFV architectural framework // SDN.IEEE.org, May 2016.
9. Real-time OSS — SDN/NFV orchestration: OSS system for SDN/NFV network orchestration in telecoms // Comarch. [https://www.comarch.com/telecommunications/oss-solutions/sdn-nfv/?gclid=EAlaIqobChMlotaH-9uQ7AIVFSwYCh35BgIAEAAAYAiAAEglYYPD\\_BwE&gclid=aw.ds](https://www.comarch.com/telecommunications/oss-solutions/sdn-nfv/?gclid=EAlaIqobChMlotaH-9uQ7AIVFSwYCh35BgIAEAAAYAiAAEglYYPD_BwE&gclid=aw.ds).
10. Шалагинов А. Шесть лет SDN/NFV — откуда пессимизм? // ИКСмедиа, 25 декабря 2018 г. <http://www.iksmedia.ru/blogs/post/5555404-Shest-let-SDNNFV-otkuda-pessimizm.html>.
11. Pate P. NFV and SDN: What's the difference? // SDXcentral, March 30, 2020. <https://www.sdxcentral.com/articles/contributed/nfv-and-sdn-whats-the-difference/2013/03/>.
12. Егоров В. Б. К вопросу о соотношении программно определяемых и традиционных IP-сетей // Системы и средства информатики, 2022. Т. 32. № 1. С. 73–82.
13. Барсков А. SDN: кому и зачем это надо? // Ж. сетевых решений/LAN, 2012. № 12. Ст. 13033012.
14. Логинов С. С. Об уровнях управления в программно-конфигурируемой сети (SDN) // Т-Comm: телекоммуникации и транспорт, 2017. Т. 11. № 3. С. 50–55.

*Поступила в редакцию 13.10.21*

---

---

## ON COMBINING THE SOFTWARE-DEFINED NETWORKING WITH THE NETWORK FUNCTIONS VIRTUALIZATION

*V. B. Egorov*

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** Many firms delivering software to the networking market offer products for the software-defined networking (SDN) and network functions virtualization (NFV). In advertising, the frequent neighborhood of both acronyms, up to their merging into SDN/NFV, creates the impression that the technologies

behind them are closely related, if not the same at all. The ascertainment of the actual interrelation between SDN and NFV is complicated by difficulties with understanding the objects themselves. In publications on the subject, some common statements have entrenched, many of them turning out to be dubious at a captious exploration. The article impugns the assumed in various publications affinity of both technologies. They originated independently developing in diverse directions and at different rates: NFV is actively demanded in provider networks, while SDN has been implemented mainly by some big internet companies, with the interest in the technology not growing last years because of a number of uncertainties. Actually, both acronyms, SDN and NFV, gravitate towards each other, even up to sticking together into SDN/NFV, basically in advertisements of networking software suppliers for purely marketing reasons.

**Keywords:** network functions virtualization; NFV; SDN; software-defined networking

**DOI:** 10.14357/08696527220204

## References

1. Fruehe, J. The battle of SDN vs. NFV. *MoorInsights&Strategy*. Available at: <https://moorinsightsstrategy.com/the-battle-of-sdn-vs-nfv/> (accessed March 14, 2022).
2. Rostelekom i TsPIKS stali partnerami v oblasti razvitiya SDN i NFV [Rostelekom and ARCCN became partners in the area of SDN and NFV development]. February 24, 2015. *Cnews*. Available at: <https://www.arccn.ru/media/1449/> (accessed March 14, 2022).
3. Tittel, E. SDN vs. NFV: What's the difference? *Cisco*. Available at: <https://www.cisco.com/c/en/us/solutions/software-defined-networking/sdn-vs-nfv.html> (accessed March 14, 2022).
4. Egorov, V. 2017. Nekotorye neyavnye voprosy setevoy virtualizatsii [Some implicit issues of the network virtualization]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 27(3):88–98.
5. Shalaginov, A. December 27, 2015. SDN i NFV: Kak eto rabotaet na seti operatora svyazi [SDN and NFV: How does it work in the telecom provider network]. *Telecom&IT*. Available at: <https://shalaginov.com/2015/12/27/sdn-%D0%B8-nfv-%D0%BA%D0%B0%D0%BA-%D1%8D%D1%82%D0%BE-%D1%80%D0%B0%D0%B1%D0%BE%D1%82%D0%B0%D0%B5%D1%82-%D0%BD%D0%B0-%D1%81%D0%B5%D1%82%D0%B8-%D0%BE%D0%BF%D0%B5%D1%80%D0%B0%D1%82%D0%BE%D1%80%D0%B0/> (accessed March 14, 2022).
6. Ganczarenko, J. July 28, 2020. The difference between SDN and NFV — a simple guide. *Codilime*. Available at: <https://codilime.com/the-difference-between-sdn-and-nfv-a-simple-guide/> (accessed March 14, 2022).
7. Sushkov, A. November 28, 2016. SDN & NFV i pri chem tut oblaka [SDN & NFV and what the clouds have to do with them]. *Habr*. Available at: <https://habr.com/ru/company/billing/blog/316324/> (accessed March 14, 2022).
8. Odini, M.-P. May 2016. SDN in NFV architectural framework. *SDN.IEEE.org*. Available at: <https://sdn.ieee.org/newsletter/may-2016/sdn-in-nfv-architectural-framework> (accessed March 14, 2022).

9. Real-time OSS — SDN/NFV orchestration: OSS system for SDN/NFV network orchestration in telecoms. *Comarch*. Available at: <https://www.comarch.com/telecommunications/oss-solutions/sdn-nfv/> (accessed March 14, 2022).
10. Shalaginov, A. December 25, 2018. Shest' let SDN/NFV — otkuda pessimizm? [Six years of SDN/NFV — why pessimistic?]. *IKSmedia*. Available at: <http://www.iksmedia.ru/blogs/post/5555404-Shest-let-SDNNFV-otkuda-pessimizm.html> (accessed March 14, 2022).
11. Pate, P. March 30, 2020. NFV and SDN: What's the difference? *SDXcentral*. Available at: <https://www.sdxcentral.com/articles/contributed/nfv-and-sdn-whats-the-difference/2013/03/> (accessed March 14, 2022).
12. Egorov, V. B. 2019. K voprosu o sootnoshenii programmno opredelyaemykh i traditsionnykh IP-setey [Interrelation between the software-defined and conventional IP-networks]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 32(1):73–82.
13. Barskov, A. 2012. SDN: komu i zachem eto nado? [SDN: Whom and what for is it necessary?] *Zh. setevykh resheniy/LAN* [Network Solutions/LAN J.]. Available at: <https://www.osp.ru/lan/2012/12/13033012> (accessed March 14, 2022).
14. Loginov, S. 2017. Ob urovnyakh upravleniya v programmno-konfiguriruемой seti (SDN) [About the control planes in the software-defined networking]. *T-Comm: Telekommunikatsii i transport* [T-Comm] 11(3):50–55.

*Received October 13, 2021*

## Contributor

**Egorov Vladimir B.** (b. 1948)—Candidate of Science (PhD) in technology, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; [VEgorov@ipiran.ru](mailto:VEgorov@ipiran.ru)

## О ПРИМЕНЕНИИ ЭКСПЕРТНЫХ МЕТОДОВ ПРИ ОЦЕНКЕ ЭФФЕКТИВНОСТИ И КАЧЕСТВА ИНФОРМАЦИОННЫХ СИСТЕМ

*А. А. Зацаринный<sup>1</sup>, Ю. С. Ионенков<sup>2</sup>*

**Аннотация:** Статья посвящена вопросу применения экспертных методов при оценке эффективности и качества информационных систем (ИС). Рассмотрены общие вопросы применения экспертных оценок, включая проведение процедуры экспертного опроса и установление согласованности мнений экспертов. Представлены различные подходы к определению числа экспертов, включая формулы для проведения расчетов, и сделаны сравнительные расчеты по этим формулам. Изложен формализованный подход к оценке компетентности экспертов, включающий их научную квалификацию, аргументированность оценок и степень знакомства с оцениваемой проблемой.

**Ключевые слова:** экспертные методы; эффективность; качество; число экспертов; квалификация экспертов; информационная система

**DOI:** 10.14357/08696527220205

### 1 Введение

В настоящее время ИС во многом определяют научно-технический потенциал страны, уровень развития ее народного хозяйства, образ жизни и деятельности человека. При этом ИС относятся к классу сложных больших систем с длительным сроком эксплуатации. Одним из важных направлений их развития и совершенствования является разработка методического аппарата, позволяющего проводить оценку их технического уровня на всех стадиях жизненного цикла: от замысла до снятия с производства и утилизации.

Технический уровень ИС может характеризоваться двумя понятиями — «эффективность» и «качество». Отметим соотношение понятий «качество» и «эффективность»: если эффективность характеризует степень соответствия системы назначению и приспособленности к достижению целей, поставленных при ее создании, то качество представляется как совокупность свойств системы, обуславливающих ее пригодность для использования по назначению [1, 2].

Анализ и условия применения существующего методического аппарата оценки эффективности и качества ИС, а также предложения по его совершенствованию

---

<sup>1</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук, AZatsarinny@ipiran.ru

<sup>2</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук, UIonenkov@ipiran.ru

представлены в работах [3–7]. В методиках, представленных в этих работах, широко используются методы экспертных оценок. Данные методы основываются на знаниях специалистов и накопленном ими опыте при проведении научно-исследовательских и опытно-конструкторских работ, изучении научно-технической литературы, опыте научно-исследовательской и производственной деятельности.

В литературе достаточно подробно представлены процедуры методов экспертных оценок, включая вопросы формирования экспертных групп, форм работы с экспертами, численности и компетентности экспертов, обработки экспертной информации. Основные факторы, определяющие точность экспертных оценок, — это число экспертов в экспертной группе и их компетентность.

В настоящей статье рассматриваются общие вопросы применения методов экспертных оценок, проведено сравнение различных подходов к определению требуемого числа экспертов, а также предложен методический подход к оценке их компетентности.

## 2 Методические подходы к экспертным оценкам

Методы экспертных оценок используются, как правило, при исследовании объектов, анализ развития которых не поддается формализации, т. е. для которых трудно разработать аналитическую модель. При оценке эффективности и качества ИС эти методы применяются для определения весовых коэффициентов и отдельных показателей эффективности.

Экспертная оценка включает [8]:

- составление таблицы экспертного опроса;
- подбор экспертов;
- проведение экспертизы;
- анализ результатов.

Качественный и количественный состав экспертной комиссии должен формироваться с учетом широты проблемы, достоверности оценок, затрат ресурсов и компетентности экспертов.

Формирование коллективного мнения экспертов осуществляется различными методами. Наиболее распространен из них групповой метод, когда все эксперты работают одновременно в составе какого-либо совета, комиссии и т. п. Недостатки данного метода: неодинаковая активность экспертов, влияние авторитетов, приспособление к мнению большинства.

Наряду с групповым методом применяется индивидуальный метод, когда эксперты работают отдельно и независимо. Обобщенное мнение формируется путем статистической обработки мнений отдельных экспертов. В этом случае удастся исключить недостатки, присущие групповому методу, но появляется опасность нивелирования мнений отдельных экспертов, следования за обезличенным статистическим большинством.

Успех экспертного опроса зависит от числа исходных показателей: чем больше показателей подлежит оценке, тем большей совокупности показателей эксперты присваивают одинаковую важность. Иначе говоря, с возрастанием числа показателей эксперты становятся все более безразличными в своих суждениях. Поскольку в общем случае число исходных показателей не может быть сокращено без ущерба для полноты описания системы, полезно введение обобщенных показателей, объединяющих ряд исходных показателей, и их группировка.

Процедура экспертного опроса подразделяется на три стадии.

На первой стадии составляется таблица экспертного опроса и приложение к ней. В таблицу заносятся характеристики и показатели систем, по которым необходимо выяснить мнение специалистов. Каждому специалисту предлагается оценить их важность в баллах, а также указать свою квалификацию, степень знакомства с рассматриваемыми системами, источники аргументации.

На второй стадии определяется круг экспертов и проводится непосредственно экспертный опрос.

На третьей стадии проводится обработка данных: вычисляются средние оценки и устанавливается степень согласованности мнений экспертов.

Сначала оценки, данные каждым экспертом, нормируются по формуле

$$\beta_{kn} = \frac{C_{kn}}{\sum_{n=1}^N C_{kn}},$$

где  $C_{kn}$  — оценка  $k$ -м экспертом важности  $n$ -го показателя (в баллах);  $N$  — число показателей.

Затем вычисляются средние значения оценок каждого показателя

$$\bar{\beta}_n = \frac{1}{K} \sum_{k=1}^K \beta_{kn}$$

и соответствующие среднеквадратические отклонения

$$\sigma = \sqrt{\frac{1}{K} \sum_{k=1}^K (\beta_{kn} - \bar{\beta}_n)^2},$$

где  $K$  — число экспертов.

После этого определяются коэффициенты вариации

$$v_n = \frac{\sigma_n}{\bar{\beta}_n},$$

характеризующие степень согласованности мнений экспертов о важности отдельных показателей. Чем больше величина  $v_n$ , тем меньше согласованность

экспертов в оценке  $n$ -го показателя; при одинаковом (полностью согласованном) мнении экспертов эта величина равна нулю.

Степень согласованности мнений экспертов по всем показателям оценивается коэффициентом конкордации Кенделла [8]:

$$\omega = \frac{12 \sum_{n=1}^N (S_n - \bar{S})^2}{K^2 (N^3 - N) - K \sum_{k=1}^K T_k}.$$

Здесь  $S_n = \sum_{k=1}^K R_{kn}$  — сумма рангов, присвоенных экспертами  $n$ -му показателю, где  $R_{kn}$  — ранг  $n$ -го показателя, назначенный  $k$ -м экспертом;  $\bar{S} = (1/N) \sum_{n=1}^N S_n$  — среднее значение сумм рангов;

$$T_k = \sum_{m=1}^{M_k} (t_{km}^3 - t_{km}),$$

где  $t_{km}$  — число одинаковых рангов  $m$ -го типа в оценках  $k$ -го эксперта,  $M_k$  — количество групп показателей с совпавшими рангами в оценках  $k$ -го эксперта.

Коэффициент конкордации изменяется в пределах от нуля до единицы, причем  $\omega = 1$  соответствует полной согласованности мнений экспертов.

При обработке данных по группам экспертов вычисляется коэффициент ранговой корреляции Спирмена [8]

$$r = 1 - \frac{6 \sum_{n=1}^N (R_{ni} - R_{nj})^2}{N (N^2 - 1)},$$

показывающий связь между оценками, назначенными экспертами  $i$ -й и  $j$ -й групп.

Если согласованность мнений экспертов высокая и неслучайная (коэффициенты конкордации и ранговой корреляции значимы и близки к единице), то найденные весовые коэффициенты и отдельные показатели могут быть использованы для расчета обобщенных оценок эффективности.

### 3 Методические подходы к определению числа экспертов

При проведении экспертных оценок очень важным вопросом является определение числа экспертов. Число экспертов должно быть достаточно большим, чтобы отдельные мнения не имели неправомерно большое значение, однако при очень большом числе экспертов снижается уровень их компетентности, что может привести к снижению точности экспертных оценок.

К экспертам предъявляются следующие требования: глубокие знания в оцениваемой области; наличие научного интереса к оцениваемым аспектам проблемы; наличие производственного или исследовательского опыта по оцениваемой проблеме.

В научно-технической литературе предлагается ряд подходов к определению требуемого числа экспертов, базирующихся на статистических методах.

В частности, в [9] предложена следующая формула для определения требуемого числа экспертов:

$$n = \frac{t^2 \sigma^2 N}{\Delta^2 N + t^2 \sigma^2}, \quad (1)$$

где  $n$  — число экспертов;  $t$  — нормированное отклонение, соответствующее принятому уровню доверительной вероятности;  $\Delta$  — предельная ошибка выборки;  $\sigma^2$  — дисперсия исследуемого признака в генеральной совокупности;  $N$  — объем генеральной совокупности.

Если дисперсия исследуемого признака даже приблизительно не известна, то она принимается равной своему максимуму — 0,25 (0,5 × 0,5) и формула (1) приобретает вид:

$$n = \frac{0,25 t^2 N}{\Delta^2 N + 0,25 t^2}. \quad (2)$$

В [10] для определения числа экспертов использовано соотношение, которое применяется для вычисления погрешности наблюдений:

$$n = \frac{t_p^2}{\varepsilon_1^2}. \quad (3)$$

Здесь  $n$  — число экспертов;  $\varepsilon_1 = \varepsilon/S$  — предельно допустимая относительная ошибка экспертной оценки, где  $S$  — среднеквадратичное отклонение распределения оценок;  $t_p$  — коэффициент Стьюдента, определяющий ширину доверительного интервала и зависимость от величины вероятности оценки  $P$  ( $t_p$  — табулированная величина).

В зависимости от заданной погрешности экспертной оценки и выбранной величины вероятности  $P$  может быть определена возможная численность экспертов (табл. 1) [10].

**Таблица 1** Минимально допустимое число экспертов в группе

| $\varepsilon_1$ | Вероятность оценки $P$ |      |      |      |     |      |     |      |
|-----------------|------------------------|------|------|------|-----|------|-----|------|
|                 | 0,99                   | 0,95 | 0,90 | 0,85 | 0,8 | 0,75 | 0,7 | 0,65 |
| 0,5             | 26                     | 15   | 11   | 8    | 7   | 5    | 4   | 4    |
| 0,3             | 74                     | 43   | 31   | 23   | 19  | 15   | 12  | 10   |

В [11] минимальное число экспертов предлагается определять по формуле:

$$n = 0,5 \left( \frac{3}{\alpha} + 5 \right), \quad (4)$$

где  $0 < \alpha \leq 1$  — параметр, задающий минимальный уровень ошибки экспертизы.

В соответствии с [12], для проведения экспертной оценки необходимо привлечение не менее 7–9 экспертов.

Для сравнения этих подходов и определения требуемого числа экспертов проведем расчеты по представленным выше формулам (2)–(4).

При расчете по формуле (2) примем величину доверительной вероятности, представляющей собой табулированную величину, равной 0,9545 ( $t = 2$ ), при этом  $\Delta = t\sqrt{\sigma^2/N} = 2 \cdot 0,1235 = 0,247$ , а объем генеральной совокупности, анализируемой экспертами,  $N$  примем равным 1500. Тогда число экспертов составит

$$n = \frac{0,25 \cdot 1500 \cdot 2^2}{0,247^2 \cdot 1500 + 0,25 \cdot 2^2} = 16.$$

При использовании формулы (3) для предельно допустимой относительной ошибки экспертной оценки, равной 0,3, число экспертов при вероятности оценки 0,8 составит 19 чел., а при вероятности оценки 0,75 — 15 чел.

При расчете по формуле (4) при  $\alpha = 0,2$  получаем 10 чел., а при  $\alpha = 0,3$  — 8 чел.

Таким образом, в соответствии с проведенными по разным формулам расчетами, для достоверного результата экспертного оценивания (вероятность оценки не менее 0,8) число экспертов должно составлять 10–15 чел., а для получения более высокой достоверности оценок число экспертов должно быть увеличено.

Кроме того, достоверность оценок, полученных в результате экспертного опроса, существенно зависит от компетентности привлекаемых экспертов.

#### 4 Методический подход к оценке компетентности экспертов

Не менее важным вопросом, чем определение числа экспертов в группе, является оценка компетентности экспертов. Следует отметить, что объективные оценки экспертов на практике трудно реализуемы и в основном используются субъективные способы: оценка на основании документов о квалификации, взаимооценка и самооценка.

Для оценки компетентности экспертов предлагается доработанный авторами статьи формализованный подход [10, 13], учитывающий научную квалификацию эксперта ( $K_n$ ), аргументированность его оценок ( $K_a$ ) и степень знакомства с оцениваемой проблемой ( $K_3$ ).

Научная квалификация эксперта ( $K_n$ ) оценивается в соответствии с табл. 2 [10, 13].

**Таблица 2** Примерный уровень научной квалификации эксперта

| Должность                                                  | Значения коэффициента квалификации $K_n$ |               |             |                      |
|------------------------------------------------------------|------------------------------------------|---------------|-------------|----------------------|
|                                                            | Без ученой степени                       | Кандидат наук | Доктор наук | Член-корр., академик |
| Старший научный сотрудник                                  | 0,15                                     | 0,225         | 0,3         | 0,5                  |
| Начальник лаборатории, руководитель группы                 | 0,2                                      | 0,3           | 0,4         | 0,6                  |
| Начальник отдела, заместитель начальника отдела            | 0,25                                     | 0,375         | 0,5         | 0,75                 |
| Руководитель комплекса, заместитель руководителя комплекса | 0,3                                      | 0,5           | 0,6         | 0,9                  |
| Директор, заместитель директора, научный руководитель темы | 0,4                                      | 0,6           | 0,8         | 1,0                  |

**Таблица 3** Оценка аргументации мнения эксперта

| Источники аргументации                           | Коэффициент аргументации |
|--------------------------------------------------|--------------------------|
| Стаж работы                                      | 0,15                     |
| Производственный и научно-исследовательский опыт | 0,2                      |
| Теоретический анализ проблемы                    | 0,15                     |
| Учет отечественных и зарубежных публикаций       | 0,15                     |
| Учет современных тенденций развития ИС           | 0,2                      |
| Наличие научных трудов                           | 0,1                      |
| Личностные характеристики эксперта               | 0,05                     |

В работах [8, 10] предложен ряд показателей для аргументированности оценок эксперта. Но они не в полной мере учитывают состояние и перспективы развития современных ИС.

Исходя из этого, авторами предлагается оценивать аргументированность оценок эксперта ( $K_a$ ) по следующим показателям: стаж работы в данной области; производственный и научно-исследовательский опыт; проведенный теоретический анализ проблемы; учет отечественных и зарубежных публикаций в данной области; учет современных тенденций развития ИС и информационно-телекоммуникационных технологий; наличие научных трудов; личностные характеристики эксперта (эмоциональность и т. д.). Каждая из этих составляющих оценивается соответствующим коэффициентом в диапазоне от 0 до 1, при этом суммарный коэффициент аргументации не должен превышать единицы. Оценка аргументации мнения эксперта представлена в табл. 3.

Степень знакомства эксперта с оцениваемой проблемой ( $K_3$ ) характеризуется выставляемым экспертом баллом самооценки (от 0 до 10), причем максимальному баллу (10) соответствует полное знакомство с оцениваемой проблемой,

а минимальному баллу (0) — полное отсутствие знакомства. Затем проводится нормирование оценки с ее переводом в диапазон от 0 до 1.

Комплексный показатель компетентности  $i$ -го эксперта в простейшем случае, принимая одинаковые значения отдельных составляющих, можно вычислить по формуле:

$$K_i = \frac{K_n + K_a + K_z}{3}.$$

В некоторых работах (см., например, [14, 15]) предлагается более сложный подход, связанный с вычислением весовых коэффициентов отдельных элементов комплексного показателя эффективности с использованием метода анализа иерархий.

Комплексный показатель компетентности эксперта характеризует весомость мнения данного эксперта и показывает, с каким весом следует включать количественные оценки данного эксперта по рассматриваемой проблеме в статистическую обработку.

## 5 Заключение

При оценке эффективности и качества ИС широко используются методы экспертных оценок. Эти методы применяются для решения задач, которые не могут быть решены формализованными методами из-за отсутствия необходимых данных по характеристикам создаваемой системы. В этом случае возникает необходимость использования экспертных оценок, основанных на опыте и знаниях специалистов в исследуемой области.

Процедура проведения экспертных оценок получила широкое развитие, она всесторонне обоснована и максимально формализована. Кроме того, хорошо разработан математический аппарат оценки точности экспертных оценок. Вместе с тем одной из основных проблем при проведении экспертиз остается определение числа экспертов и их компетентности.

В данной статье рассмотрен ряд подходов к определению числа экспертов, проведены расчеты по различным формулам и сделан вывод о том, что число экспертов при вероятности оценки не менее 0,8 должно составлять 10–15 человек, а для получения более высокой достоверности оценок число экспертов должно быть увеличено.

Для оценки компетентности экспертов предложен формализованный подход, учитывающий научную квалификацию эксперта, аргументированность его оценок и степень знакомства с оцениваемой проблемой. Даны предложения по расчету комплексного показателя компетентности экспертов.

## Литература

1. ГОСТ 15467-79. Управление качеством продукции. Основные понятия. Термины и определения. — М.: Стандартинформ, 2009. 21 с.

2. ГОСТ 34.003-90. Информационная технология. Автоматизированные системы. Термины и определения. — М.: Стандартинформ, 2005. 14 с.
3. *Зацаринный А. А., Ионенков Ю. С., Шабанов А. П.* К вопросу о сравнительной оценке эффективности ситуационных центров // Системы и средства информатики, 2013. Т. 23. № 2. С. 155–171.
4. *Зацаринный А. А., Ионенков Ю. С.* К вопросу оценки эффективности автоматизированных систем с использованием метода анализа иерархий // Системы и средства информатики, 2015. Т. 25. № 3. С. 161–178.
5. *Зацаринный А. А., Ионенков Ю. С.* Некоторые аспекты оценки эффективности автоматизированных информационных систем на различных стадиях их жизненного цикла // Системы и средства информатики, 2016. Т. 26. № 3. С. 123–136.
6. *Зацаринный А. А., Ионенков Ю. С., Сучков А. П.* Некоторые аспекты оценки эффективности облачных технологий // Системы и средства информатики, 2018. Т. 28. № 3. С. 104–117.
7. *Зацаринный А. А., Ионенков Ю. С.* Некоторые методические аспекты выбора показателей эффективности информационных систем // Системы высокой доступности, 2019. № 4. С. 19–26.
8. *Окунев Ю. Б., Плотников В. Г.* Принципы системного подхода к проектированию в технике связи. — М.: Связь, 1976. 183 с.
9. *Кузовкова Т. А., Салютин Т. Ю., Шаравова О. И.* Статистика инфокоммуникаций. — М.: Горячая линия-Телеком, 2019. 548 с.
10. *Семенов С. С.* Оценка качества и технического уровня сложных систем: Практика применения метода экспертных оценок. — М.: ЛЕНАНД, 2015. 352 с.
11. *Литвак Б. Г.* Экспертные оценки и принятие решений. — М.: Патент, 1996. 298 с.
12. *Саати Т.* Принятие решений: Метод анализа иерархий. — М.: Радио и связь, 1993. 278 с.
13. *Бобровников Г. Н., Клебанов А. И.* Комплексное прогнозирование создания новой техники. — М.: Экономика, 1989. 205 с.
14. *Чернышева Т. Ю.* Иерархическая модель оценки и отбора экспертов // Доклады ТУСУР, 2009. № 1(19). Ч. 1. С. 168–173.
15. *Марычева П. Г.* Методика оценки компетентности экспертов // Вестник Самарского государственного технического университета. Сер.: Технические науки, 2018. № 4(60). С. 29–40.

*Поступила в редакцию 28.12.21*

---

---

## ON THE USE OF EXPERT METHODS IN EVALUATING EFFECTIVENESS AND QUALITY OF INFORMATION SYSTEMS

*A. A. Zatsarinny and Yu. S. Ionenkov*

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The article is devoted to the application of expert methods in assessing effectiveness and quality of information systems. The general issues of the

application of expert assessments, including the conduct of the expert survey procedure and the establishment of consistency of expert opinions, are considered. Various approaches to determining the number of experts, including formulas for calculations, are presented and comparative calculations based on these formulas are carried out. A formalized approach to assessing the competence of experts is presented, including their scientific qualifications, reasonableness of assessments, and the degree of familiarity with the problem being evaluated.

**Keywords:** expert methods; efficiency; quality; number of experts; qualification of experts; information system

**DOI:** 10.14357/08696527220205

## References

1. GOST 15467-79. 2009. Upravlenie kachestvom produktsii. Osnovnye ponyatiya. Terminy i opredeleniya [Product quality control. Basic concepts. Terms and definitions]. Moscow: Standardinform Publ. 21 p.
2. GOST 34.003-90. 2005. Informatsionnaya tekhnologiya. Avtomatizirovannye sistemy. Terminy i opredeleniya [Information technology. Set of standards for automated systems. Automated systems. Terms and definitions]. Moscow: Standardinform Publ. 14 p.
3. Zatsarinny, A. A., Yu. S. Ionenkov, and A. P. Shabanov. 2013. K voprosu o sravnitel'noy otsenke effektivnosti situatsionnykh tsentrov [Regarding comparative evaluation of situational centers efficiency]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 23(2):155–171.
4. Zatsarinny, A. A., and Yu. S. Ionenkov. 2015. K voprosu otsenki effektivnosti avtomatizirovannykh sistem s ispol'zovaniem metoda analiza ierarkhiy [Regarding automated systems efficiency evaluation using analytic hierarchy process]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(3):161–178.
5. Zatsarinny, A. A., and Y. S. Ionenkov. 2016. Nekotorye aspekty otsenki effektivnosti avtomatizirovannykh informatsionnykh sistem na razlichnykh stadiyakh ikh zhiznennogo tsikla [On aspects of automated information system efficiency evaluation at different stages of lifecycle]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 26(3):123–136.
6. Zatsarinny, A. A., Yu. S. Ionenkov, and A. P. Suchkov. 2018. Nekotorye aspekty otsenki effektivnosti oblachnykh tekhnologiy [Some aspects of cloud computing efficiency estimation]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(3):104–117.
7. Zatsarinny, A. A., and Y. S. Ionenkov. 2019. Nekotorye metodicheskie aspekty vybora pokazateley effektivnosti informatsionnykh sistem [Some methodological aspects of the choice of performance indicators of information systems]. *Sistemy vysokoy dostupnosti [Highly Available Systems]* 15(4):19–26.
8. Okunev, Y. B., and V. G. Plotnikov. 1976. *Printsipy sistemnogo podkhoda k proektirovaniyu v tekhnike svyazi* [Systematic approach principles to design in communication technics]. Moscow: Svyaz'. 183 p.
9. Kuzovkova, T. A., T. Yu. Salyutina, and O. I. Sharavova. 2019. *Statistika infokommunikatsiy* [Statistics of infocommunications]. Moscow: Goryachaya liniya-Telekom. 548 p.

10. Semenov, S. S. 2015. *Otsenka kachestva i tekhnicheskogo urovnya slozhnykh sistem: Praktika primeneniya metoda ekspertnykh otsenok* [Assessment of the quality and technical level of complex systems: The practice of applying the method of expert assessments]. Moscow: LENAND. 352 p.
11. Litvak, B. G. 1996. *Ekspertnye otsenki i prinyatie resheniy* [Expert assessments and decision-making]. Moscow: Patent. 298 p.
12. Saati, T. 1993. *Prinyatie resheniy. Metod analiza ierarkhiy* [Decision making. Analytic hierarchy process]. Moscow: Radio i svyaz'. 278 p.
13. Bobrovnikov, G. N., and A. I. Klebanov. 1989. *Kompleksnoe prognozirovanie sozdaniya novoy tekhniki* [Comprehensive forecasting of the creation of new equipment]. Moscow: Ekonomika. 205 p.
14. Chernysheva, T. Yu. 2009. Ierarkhicheskaya model' otsenki i otbora ekspertov [Hierarchical model of evaluation and selection of experts]. *Doklady TUSUR* [Proceedings of the TUSUR University] 1-1(19):168–173.
15. Marycheva, P. G. 2018. Metodika otsenki kompetentnosti ekspertov [The method of assessment the competence of experts]. *Vestnik Samarskogo gosudarstvennogo tekhnicheskogo universiteta. Ser. Tekhnicheskie nauki* [Vestnik of Samara State Technical University. Technical sciences ser.] 4(60):29–40.

*Received December 28, 2021*

## Contributors

**Zatsarinny Alexander A.** (b. 1951) — Doctor of Science in technology, professor, principal scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; AZatsarinny@ipiran.ru

**Ionenkov Yuriy S.** (b. 1956) — Candidate of Science (PhD) in technology, senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; uionenkov@ipiran.ru

## АНАЛИТИЧЕСКОЕ МОДЕЛИРОВАНИЕ И ОЦЕНИВАНИЕ НЕСТАЦИОНАРНЫХ НОРМАЛЬНЫХ ПРОЦЕССОВ В СТОХАСТИЧЕСКИХ СИСТЕМАХ, НЕ РАЗРЕШЕННЫХ ОТНОСИТЕЛЬНО ПРОИЗВОДНЫХ

*И. Н. Синицын*<sup>1</sup>

**Аннотация:** Разработано методическое обеспечение для аналитического моделирования, оценивания, идентификации и калибровки существенно нестационарных, например ударных, нормальных стохастических систем (СтС), не разрешенных относительно производных, основанное на регрессионной линеаризации. Дан иллюстративный пример. Полученные результаты могут быть обобщены на случай дифференциальных интегродифференциальных систем, стохастически не разрешенных относительно старших производных, в условиях многоэшелонированных ударных воздействий. Исследование выбросов и накопление возмущений также имеет важное прикладное значение.

**Ключевые слова:** аналитическое моделирование; калибровка; оценивание (фильтрация и экстраполяция); стохастическая система, не разрешенная относительно старшей производной (СтС НРОП); ударная СтС НРОП; фильтр и экстраполятор Калмана и Бьюси

**DOI:** 10.14357/08696527220206

### 1 Введение

В [1] рассмотрены конечномерные нелинейные стационарные и нестационарные СтС, не разрешенные относительно производных, в условиях широкополосных гауссовских и негауссовских возмущений. Такие СтС описывают поведение многих технических систем информатики и управления. Для аналитического моделирования и оценки чувствительности к параметрам нормальных (гауссовских) нестационарных и стационарных стохастических процессов (СтП) в таких СтС разработаны два метода аналитического моделирования (МAM): на основе метода статистической линеаризации (МСЛ) и метода нормальной аппроксимации (МНА). В приложении 1 приведены типовые нелинейности, не разрешенные относительно производных, а в приложениях 2 и 3 — два тестовых примера. Алгоритмы положены в основу разрабатываемого инструментального программного обеспечения для решения задач надежности и безопасности технических систем.

---

<sup>1</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук; Московский авиационный институт, sinitsin@dol.ru

Методы параметрического аналитического моделирования для СтП в конечномерных нелинейных стационарных и нестационарных СтС, не разрешенных относительно производных, рассмотрены в [2]. Такие модели описывают поведение многих технических систем информатики и управления, если пренебречь переходными процессами. Для аналитического моделирования и оценки точности и чувствительного алгоритмов разработаны МАМ на основе МНА и МСЛ, методов ортогональных разложений (МОР) и квазиоментов, а также других методов параметризации одно- и многомерных плотностей. Особое внимание уделено МАМ для дискретных СтС по МОР на основе эталонного пуассоновского распределения. Представлены типовые нелинейности, не разрешенные относительно производных, и коэффициенты их статистической линеаризации, а также два тестовых примера.

Вопросы синтеза нелинейных нормальных (гауссовских) субоптимальных фильтров (НСОФ) для дифференциальных СтС, не разрешенных относительно производных, рассмотрены в [3]. Представлены уравнения состояния и наблюдения нелинейных дифференциальных СтС, не разрешенных относительно производных. Синтез НСОФ выполнен при следующих условиях: (1) отсутствуют пуассоновские шумы в наблюдениях; (2) коэффициент при гауссовском шуме не зависит от состояния. Подробно рассмотрен синтез НСОФ при аддитивных шумах в уравнениях состояния и наблюдения. Для иллюстрации методов синтеза НСОФ приведен пример.

В [4] для нелинейных интегродифференциальных СтС (ИДСтС), не разрешенных относительно производных и приводимых к дифференциальным методом сингулярных ядер, разработаны алгоритмы аналитического моделирования нормальных СтП, при этом нелинейность под интегралом может быть разрывной, а также синтеза НСОФ для онлайн-обработки информации в ИДСтС. Подробно рассмотрен тестовый пример с разрывной нелинейностью под знаком интеграла. Предложены алгоритмы оценки качества НСОФ на основе теории чувствительности.

В [5] разработано методическое обеспечение для негладких правых частей уравнений СтС, не разрешенных относительно старшей производной (НРОП). Рассмотрены вопросы аналитического моделирования нормальных процессов на основе нелинейных регрессионных моделей. Особое внимание уделено методам нормальной (гауссовской) фильтрации и экстраполяции. Изучены вопросы условно-оптимальной фильтрации и экстраполяции для СтС НРОП с параметрическими шумами.

Для систем, стохастически не разрешенных относительно производных, в [6] разработаны два подхода к сведению таких систем к детерминированным уравнениям, не разрешенным относительно математических ожиданий и ковариационных характеристик, а также математических ожиданий и координатных функций канонических разложений (КР). После сведения таких систем к детерминированным используются известные результаты. Приведен пример. Рассмотрены вопросы оценивания (фильтрации, экстраполяции и др.), идентификации и ка-

либровки для приведенных моделей. Алгоритмы положены в основу разрабатываемого инструментального программного обеспечения для решения задач надежности и безопасности технических систем.

Рассмотрим развитие [5] на случай типовых существенно нестационарных, в первую очередь ударных, возмущений. В разд. 2 рассмотрены вопросы аналитического моделирования нормальных СтП в скалярных и векторных СтС НРОП с помощью общих, а также основанных на КР нелинейных корреляционных методов. Особое внимание уделено как широкополосным, так и узкополосным и ударным возмущениям. Раздел 3 посвящен вопросам оптимизации, оцениванию, идентификации и калибровке. В разд. 4 приведен пример, иллюстрирующий методы разд. 2 и 3. Заключение содержит выводы и некоторые обобщения.

## 2 Аналитическое моделирование

**2.1.** Рассмотрим нестационарную скалярную дифференциальную систему следующего вида:

$$\varphi \equiv \varphi(t, \Theta, X_t, pX_t, \dots, p^{(l)}X_t, U_t) = 0. \quad (1)$$

Здесь  $\Theta$  — инструментальный вектор параметров;  $p = d/dt$ ;  $X_t = X(t)$  и  $p^{(l)}X_t$  — скалярные СтП, понимаемые в среднеквадратичном смысле;  $U_t$  — скалярный СтП с известными вероятностными характеристиками;  $\varphi$  — нелинейная, в общем случае разрывная функция отмеченных переменных, допускающая линейную регрессионную аппроксимацию относительно старшей производной вида:

$$\varphi \approx \varphi_0 + k_{\bar{X}}^{\varphi} p^{(l)} X_t^0 + k_U^{\varphi} U_t^0. \quad (2)$$

Здесь введены следующие обозначения:

$$\varphi_0 = \varphi_0(t, \Theta, m_t^{\bar{X}}, K_t^{\bar{X}}), \quad \bar{k}_{\bar{X}_l}^{\varphi} = k_{\bar{X}_l}^{\varphi}(t, \Theta, m_t^{\bar{X}}, K_t^{\bar{X}}),$$

где  $\bar{X}_t = [\bar{X}_t^T U_t]^T$ ;  $\bar{X}_t = [X_t \dots X_t^{(l-1)}]^T$ . После введения переменных

$$pX_t = X_1, \dots, pX_t^{(l-1)} = X_t^{(l)}$$

получим векторную дифференциальную систему, разрешенную относительно старшей производной  $\dot{X}_l$ :

$$\dot{X}_1 = X_2, \dots, \dot{X}_{l-1} = X_l; \quad \varphi - \varphi_0 = k_{\bar{X}_l}^{\varphi} p^{(l)} X_l^0 + k_U^{\varphi} U_t^0 = 0$$

или при условии  $k_X^\varphi \neq 0$ :

$$\varphi_0 = 0; \quad (3)$$

$$\dot{X}_1^0 = X_2^0, \dots, \dot{X}_{l-1}^0 = X_l^0, \quad \dot{X}_l^0 = -\frac{\varphi_0}{k_X^\varphi} X_l^0 - \frac{k_U^\varphi}{k_X^\varphi} U_t^0. \quad (4)$$

Решая нелинейную систему (3) совместно с уравнениями (4) по формулам корреляционной теории СтП [7–9], найдем искомые  $m_t^{\bar{X}}$  и  $K_t^{\bar{X}}$ .

Для определения  $K_t^{\bar{X}}$  и  $K^{\bar{X}}(t_1, t_2)$  используются формулы корреляционной теории:

$$K_t^{\bar{X}} = \mathcal{A}_t^2 K_t^U; \quad K^{\bar{X}}(t_1, t_2) = \mathcal{A}_{t_1} \mathcal{A}_{t_2} K^U(t_1, t_2), \quad (5)$$

где  $\mathcal{A}_t$  — линейный оператор, определяемый уравнениями (4);  $K_t^U$  — ковариационная функция возмущения  $U_t$ .

**Утверждение 2.1.** Пусть нестационарная скалярная дифференциальная система (1), не разрешенная относительно старшей производной, допускает линейную регрессионную аппроксимацию вида (2), причем  $K_X^\varphi \neq 0$ . Тогда (1) приводится к нестационарной дифференциальной системе, разрешенной относительно старшей производной, вида (3) и (4) при условии (5).

**2.2.** Для нестационарного скалярного возмущения  $U_t$  с конечной ковариационной функцией  $K_t^0$  и выражаемого КР

$$U_t = m_t^U + \sum_{j=1}^{H_U} V_j u_{jt}^U,$$

где  $m_t^U = MU_t$ ;  $V_j$  — независимые случайные величины с дисперсиями  $D_j$ ;  $u_{jt}^U$  — детерминированные координатные функции, имеем следующие соотношения:

$$\left. \begin{aligned} X_t &= m_t^X + \sum_{j=1}^{H_x} V_j u_{jt}^X; \\ K_t^X &= \sum_{j=1}^{H_x} D_j \left( u_{jt}^X \right)^2; \\ K^X(t_1, t_2) &= \sum_{j=1}^{H_x} D_j u_{jt_1}^X (u_{jt_2}^X)^T; \\ \varphi &\approx \varphi_0 + \sum_{j=1}^{H_\varphi} V_j u_{jt}^\varphi = 0; \\ \varphi_0 &= 0; \\ u_{jt}^\varphi &= 0. \end{aligned} \right\} \quad (6)$$

Здесь  $u_{jt}^\varphi$  — координатные функции линейной регрессионной статистической линеаризации посредством КР [8, 9], зависящие от  $m_t^X$ ,  $m_t^U$ ,  $u_{jt}^U$ ,  $u_{jt}^X$  и  $D_j$ .

**Утверждение 2.2.** Пусть нестационарная скалярная дифференциальная система (1), не разрешенная относительно старшей производной, допускает линейную регрессионную аппроксимацию по методу КР (МКР). Тогда (1) приводится к нестационарной дифференциальной системе, разрешенной относительно старшей производной, вида (6).

**2.3.** Как известно [8], метод гармонической линеаризации можно рассматривать как частный случай МКР при гармонических функциях со случайной равномерно распределенной начальной фазой. Поэтому при  $U_t = 0$  и автоколебаниях на частоте  $\Omega$  и случайных начальных условиях имеют место соотношения:

$$X_t = A_0 + A \sin \Omega t; \quad \varphi \approx \varphi_0 + (q + q' \Omega^{-1} p) (X_t - A_0) = 0, \quad (7)$$

где  $\varphi_0$ ,  $q$  и  $q'$  зависят от  $A_0$ ,  $A$  и  $\Omega$ :

$$\begin{aligned} \varphi_0 &= \frac{1}{2\pi} \int_0^{2\pi} \varphi (A_0 + a \sin \Omega t) d\Omega t = 0; \\ q &= \frac{1}{\pi A} \int_0^{2\pi} \varphi (A_0 + a \sin \Omega t) \sin \Omega t d\Omega t = 0; \\ q' &= \frac{1}{\pi A} \int_0^{2\pi} \varphi (A_0 + a \sin \Omega t) \cos \Omega t d\Omega t = 0. \end{aligned}$$

**Утверждение 2.3.** Пусть в системе (1)  $U_t \equiv 0$  и имеют место гармонические автоколебания на частоте  $\Omega$ . Тогда в основе алгоритма приведения лежат соотношения (7) метода гармонической линеаризации.

**2.4.** Для гармонических возмущений  $U_t = A \sin(\omega t + \chi_m)$  линейная регрессионная модель МКР совпадает с моделью гармонической линеаризации и приводит к следующим результатам:

$$\tilde{X}_t = \tilde{A}_0 + \tilde{A} \sin \psi; \quad \psi = \omega t + \tilde{\psi}; \quad \varphi \approx \varphi_0 + (q + q' \omega^{-1} p) (\tilde{X}_t - \tilde{A}_0) = 0, \quad (8)$$

где  $p = d/dt$ ;  $\varphi_0$ ,  $q$  и  $q'$  зависят от  $\tilde{A}_0$ ,  $\tilde{A}$  и  $\tilde{\psi}$ :

$$\varphi_0 = \frac{1}{2\pi} \int_0^{2\pi} \varphi d\psi = 0; \quad q = \frac{1}{\pi \tilde{A}} \int_0^{2\pi} \varphi \sin \psi d\psi = 0; \quad q' = \frac{1}{\pi \tilde{A}} \int_0^{2\pi} \varphi \cos \psi d\psi = 0.$$

Для  $U_t$ , близких к стационарным, координатные функции выбираются в виде произведений  $e^{-\alpha t} \cos(\omega t + \beta)$ .

Для ударных  $U_t$  они задаются алгоритмически или как решения линейных дифференциальных уравнений с переменными коэффициентами.

**Утверждение 2.4.** Пусть в системе (1) имеют место вынужденные гармонические колебания на частоте  $\omega$ . Тогда в основе алгоритма приведения лежат соотношения (8) метода гармонической линеаризации.

**2.5.** Пусть уравнение (1) векторное ( $\dim \bar{\varphi} = n^{\bar{\varphi}}$ ) и имеет вид:

$$\bar{\varphi} = \bar{\varphi}(t, \Theta, \bar{X}_t, \dot{\bar{X}}_t, \dots, p^{(l)} \bar{X}_t, U_t) = 0, \quad (9)$$

где вектор состояния  $\bar{X}_t$  имеет размерность  $\dim \bar{X}_t = n^{\bar{X}}$ , а вектор возмущений  $U_t$  обладает размерностью  $n^U$ . Тогда имеют место следующие соотношения:

$$\bar{\varphi} \approx \bar{\varphi}_0 + k_{\bar{X}}^{\bar{\varphi}} \bar{X}_t^0 + k_U^{\bar{\varphi}} U_t^0, \quad (10)$$

где

$$k_{\bar{X}}^{\bar{\varphi}} \bar{X}_t^0 + k_U^{\bar{\varphi}} U_t^0 = 0$$

или

$$\begin{aligned} \bar{X}_t^0 &= - \left( k_{\bar{X}}^{\bar{\varphi}} \right)^{-1} k_U^{\bar{\varphi}} U_t^0 = \bar{A}_t U_t^0 \text{ при } \det(k_{\bar{X}}^{\bar{\varphi}}) \neq 0; \\ \bar{\varphi}_0 &= 0; \end{aligned} \quad (11)$$

$$K_t^{\bar{X}} = \bar{A}_t^2 K_t^U; \quad (12)$$

$$K^{\bar{X}}(t_1, t_2) = \bar{A}_{t_2} \bar{A}_{t_1} K^U(t_1, t_2). \quad (13)$$

Здесь  $\bar{\varphi}_0$ ,  $k_{\bar{X}}^{\bar{\varphi}}$ ,  $k_U^{\bar{\varphi}}$  и  $\bar{A}_t$  параметрически зависят от  $m_t^{\bar{X}}$ ,  $m_t^U$ ,  $K_t^{\bar{X}}$ ,  $K_t^U$  и  $K_t^{\bar{X}U}$ .

**Утверждение 2.5.** Пусть нестационарная векторная дифференциальная система (9), не разрешенная относительно старшей производной, допускает линейную регрессионную аппроксимацию (10), причем  $\det k_{\bar{X}}^{\bar{\varphi}} \neq 0$ . Тогда (9) приводится к разрешенной относительно старшей производной системе (11)–(13).

Если нестационарное, например ударное, возмущение  $U_t$  выразить с помощью уравнения формирующего фильтра  $(\Phi\Phi)$  через нестационарный гауссовский (нормальный) белый шум  $V_t$

$$\dot{U}_t = A_{1t}^U U_t + B_t^U V_t, \quad (14)$$

то для составного вектора  $\bar{\bar{X}}_t = [\bar{X}_t^T U_t^T]^T$  придем к соотношениям:

$$\bar{\bar{\varphi}} \approx \bar{\bar{\varphi}}_0 + k_{\bar{\bar{X}}}^{\bar{\bar{\varphi}}} \bar{\bar{X}}_t^0 = 0; \quad (15)$$

$$\left. \begin{aligned} \dot{\bar{X}}_t^0 &= A_{1t}^{\bar{X}} \bar{X}_t + B_t^V V_t; \\ \bar{\varphi}_0 &= 0; \\ \dot{K}_t^{\bar{X}} &= A_{1t}^{\bar{X}} K_t^{\bar{X}} + K_t^{\bar{X}} (A_{1t}^{\bar{X}})^T + B_t^V \nu_t (B_t^V)^T; \\ \frac{\partial K^{\bar{X}}(t_1, t_2)}{\partial t_2} &= K^{\bar{X}}(t_1, t_2) A_{1t_2}^T, \quad K^{\bar{X}}(t_1, t_1) = K_{t_1}^{\bar{X}}. \end{aligned} \right\} \quad (16)$$

Здесь  $A_t^{\bar{X}}$ ,  $A_{1t}^{\bar{X}}$  и  $B_t^V$  зависят параметрически от  $m_t^{\bar{X}}$ ,  $m_t^U$ ,  $K_t^{\bar{X}}$ ,  $K_t^U$  и  $K_t^{\bar{X}U}$ .

**Утверждение 2.6.** Пусть нестационарная векторная дифференциальная система (9), (14), не разрешенная относительно производной, допускает линейную регрессионную аппроксимацию (15). Тогда (9), (14) приводятся к разрешенной относительно старшей производной системе (16).

### 3 Оптимизация, оценивание, идентификация и калибровка

**3.1.** Следуя [10–12], будем считать систему оптимальной, если она при решении поставленной задачи обеспечивает минимум принятого критерия качества, и статистически оптимальной системой будем называть систему, которая обеспечивает минимум статистического критерия качества. В дальнейшем статистически оптимальные системы для краткости будем называть просто оптимальными.

Наиболее часто в практических задачах пользуются точностными критериями, т. е. критерием минимума средней квадратической ошибки (с. к. о.). Обозначим через  $\delta \bar{X}_t = \bar{X}_t - \bar{X}_{TP}$  вектор ошибки нормально распределенного выходного сигнала модели, а через  $m_t^{\delta \bar{X}}$  и  $K_t^{\delta \bar{X}}$  — вектор математических ожиданий и ковариационную матрицу в момент времени  $t$ . Тогда критерий минимума с. к. о. можно представить в следующем векторно-матричном виде:

$$\varepsilon^2 = \text{tr } M_N \left[ \delta \bar{X}_t \delta \bar{X}_t^T \right] = \text{tr } \Gamma_t^{\delta \bar{X}} = \text{tr} \left[ K_t^{\delta \bar{X}} + m_t^{\delta \bar{X}} \left( m_t^{\delta \bar{X}} \right)^T \right],$$

где через  $\text{tr}$  обозначен след соответствующей матрицы;  $M_N$  — математическое ожидание для нормального распределения.

Будем рассматривать ансамбль систем как промышленное изделие, выполненное по известным стандартам и обладающее инструментальными ошибками, описываемыми вектором  $\Theta$ , т. е.  $\delta \bar{X}_t = \delta \bar{X}_t(\Theta)$ . Тогда следует различать методическую ( $\Theta \equiv 0$ ) и инструментальную точность ( $\Theta \neq 0$ ).

Для вероятностного анализа нормальных СтП по априорным данным для  $\delta \bar{X}_t$  и  $\delta \bar{X}_t(\Theta)$  предполагается априорное знание параметров нормального распределения. В этом случае вероятностный анализ наряду с оценкой точности позволяет решать задачу оптимизации по критерию минимума с. к. о.

Во многих практических случаях оптимизация характеристик системы как промышленного изделия не обеспечивает требуемую точность на этапах жизненного цикла. Необходим учет апостериорных (условных) данных после проведения определенных (прямых и косвенных) наблюдений над системой.

Различают оценивание и идентификацию [11, 12]. Оценивание — это процедура оценки полезного сигнала по наблюдению некоторого входного сигнала. Последний может быть случайной функцией переменных, описывающих функционирование системы. Идентификация — это процедура определения оценок характеристик системы по результатам наблюдения ее входных и выходных сигналов. Полезный сигнал при идентификации может представлять собой вектор неизвестных параметров. Последнюю задачу называют калибровкой. Задача идентификации состоит в оценке вектора параметров. Различие задач оценивания и идентификации непринципиальное и заключается в разных условиях наблюдения и особенностях требуемого полезного сигнала. Это различие приводит к некоторой специфике оптимальных алгоритмов оценки. Они подробно изучены в [11, 12].

**3.2.** В задачах нормальной (гауссовской) фильтрации используются два типа уравнений: уравнения для составного вектора состояния

$$\bar{X}_t = [X_t, \dots, X_t, \dots, X_t^{(l)}]^T$$

и для вектора наблюдений

$$\bar{Y}_t = [Y_{1t} \dots Y_{(l-1)t}]^T.$$

Уравнения состояния имеют вид:

$$\varphi(t, \Theta, \bar{X}_t, \bar{Y}_t, \bar{U}_t) = 0, \quad (17)$$

где  $\bar{U}_t = [\bar{U}_1^T \bar{U}_2^T]^T$ . Будем считать:

- (1) нестационарные возмущения  $\bar{U}_t$  связаны со скалярным нормальным белым шумом  $V_t$  уравнениями  $\Phi\Phi$ ;
- (2) уравнение (17) допускает приведение к линейному детерминированному уравнению, разрешенному относительно старшей производной;
- (3) уравнение для вектора наблюдения линейно относительно  $\bar{X}_t$  и  $\bar{Y}_t$ .

Тогда, применяя теорию линейной фильтрации, основанную на приведенных уравнениях вида

$$\dot{\bar{X}}_t = a\bar{Y}_t + a_1\bar{X}_t + a_0 + \psi V; \quad (18)$$

$$\dot{\bar{Y}}_t = b\bar{Y}_t + b_1\bar{X}_t + b_0 + \psi_1 V, \quad (19)$$

придем к искомым фильтрационным уравнениям общего линейного фильтра [11]:

$$\left. \begin{aligned} \dot{\hat{X}}_t &= \left( a\bar{Y}_t + a_1\hat{\bar{X}}_t + a_0 \right) + \beta_t \left[ \dot{Y}_t - (b\bar{Y}_t + b_1\hat{\bar{X}}_t + b_0) \right], \quad \hat{\bar{X}}_{t_0} = \bar{X}_0, \\ \beta_t &= (R_t b_1^T + \psi \nu \psi_1^T) (\psi_1 \nu \psi_1^T)^{-1}; \\ \dot{R}_t &= a_1 R_t + R_t a_1^T + \psi \nu \psi^T - (R_t b_1^T + \psi \nu \psi_1^T) (\psi_1 \nu \psi_1^T)^{-1} (b_1 R_t + \psi_1 \nu \psi^T), \\ R_{t_0} &= R_0 \end{aligned} \right\} \quad (20)$$

при условии  $\det(\psi_1 \nu \psi_1^T) \neq 0$ .

Уравнения (20) при  $a = 0$  и  $b = 0$  представляют собой уравнения фильтра Калмана–Бьюси [7, 8, 11]. Последние удобно записывать в виде:

$$\left. \begin{aligned} \dot{\bar{X}}_t &= a_{1t} \bar{X}_t + a_{0t} + V_1; \\ Z_t &= \dot{Y}_t = b_{1t} \bar{X}_t + V_2; \end{aligned} \right\} \quad (21)$$

$$\left. \begin{aligned} \dot{\hat{X}}_t &= a_{1t} \hat{\bar{X}}_t + a_{0t} + \beta_t (Z_t - b_{1t} \hat{\bar{X}}_t), \quad \beta_t = R_t b_{1t}^T \nu_{2t}^{-1}; \\ \dot{R}_t &= a_{1t} R_t + R_t a_{1t}^T + \nu_{1t} - \beta_t \nu_{2t} \beta_t^T \end{aligned} \right\} \quad (22)$$

при условии  $\det \nu_{2t} \neq 0$  и соответствующих начальных условиях.

**Утверждение 3.1.** Пусть не разрешенная относительно старшей производной система (17) допускает приведение к линейной регрессионной системе (18), а уравнение наблюдения имеет вид (19). Тогда уравнения нормального фильтра имеют вид (20) при условии  $\det(\psi_1 \nu \psi_1^T) \neq 0$ .

**Утверждение 3.2.** Пусть в условиях утверждения 3.1 система допускает приведение к виду (21). Тогда уравнения нормального фильтра имеют вид (22) при условии  $\det \nu_{2t} \neq 0$ .

**3.3.** Применяя экстраполятор Калмана–Бьюси к уравнениям утверждения 3.2, получим следующие соотношения:

$$\begin{aligned} \hat{\bar{X}}_{t+\Delta/t} &= A_{1t} \hat{\bar{X}}_{t+\Delta/t} = u(t+\Delta, t) \hat{\bar{X}}_{t/t} + \int_t^{t+\Delta} u(t+\Delta, \tau) A_{0\tau}(\tau) d\tau \quad (\Delta > 0), \\ \left[ \hat{\bar{X}}_{t+\Delta/t} \right]_{\Delta=0} &= \hat{\bar{X}}_t; \\ \dot{R}_{t+\Delta/t} &= a_{1t} R_{t+\Delta/t} + R_{t+\Delta/t} a_{1t}^T + \psi \nu \psi^T, \quad [R_{t+\Delta/t}]_{\Delta=0} = R_t. \end{aligned} \quad (23)$$

Здесь  $u_t = u(t, \tau)$  — фундаментальное решение уравнения  $\dot{u}_t = A_{1t} u_t$  при условии  $u(t, t) = I$ . Таким образом, экстраполятор будет представлять собой последовательное соединение усилителя  $u(t+\Delta, t)$ , сумматора и дополнительного интегрального члена  $\int_t^{t+\Delta} u(t+\Delta, \tau) A_{0\tau} d\tau$ .

**Утверждение 3.3.** В условиях утверждения 3.2 нормальный экстраполятор определяется уравнениями (23).

**3.4.** Уравнения утверждений 3.1–3.3, если в составном векторе состояния  $\tilde{X}_t = [\bar{X}_t^T \Theta^T]^T$  положить  $\dot{\Theta} = 0$ , представляют собой идентификационные уравнения. Уравнения калибровки представляют собой группу уравнений для непосредственной оценки  $\Theta$ .

В [7, 8] показано, что при использовании МКР в задачах оценивания, идентификации и калибровки целесообразно пользоваться дискретными версиями уравнений состояния, а также фильтров и экстраполяторов Калмана–Бьюси.

## 4 Пример

**4.1.** Сначала, пользуясь результатами разд. 2, получим алгоритм аналитического моделирования для скалярной системы

$$\varphi(\dot{X}_t, X_t) \equiv \varphi_1(\dot{X}_t) + \varphi_2(X_t) + U_{1t} = 0; \quad (24)$$

$$\dot{U}_{1t} = \alpha_{10} + \alpha_{11}U_{1t} + \beta_1 V_{1t}. \quad (25)$$

Здесь  $\varphi_1$  и  $\varphi_2$  — скалярные, в общем случае нелинейные, функции;  $X_t$  и  $\dot{X}_t$  — переменная состояния и ее производная;  $U_{1t}$  — скалярное ударное возмущение;  $V_{1t}$  — нормальный нестационарный белый шум интенсивности  $\nu_{1t}$ ;  $\alpha_{10}$ ,  $\alpha_{11}$  и  $\beta_1$  — постоянные параметры. Выполним регрессионную линеаризацию функций  $\varphi_1$  и  $\varphi_2$  согласно

$$\varphi_1 \approx \varphi_{10} + k_X^{\varphi_1} \dot{X}_t^0; \quad \varphi_2 \approx \varphi_{20} + k_X^{\varphi_2} X_t^0.$$

При условии  $k_X^{\varphi_1} \neq 0$  придем к уравнениям для математических ожиданий и центрированных составляющих:

$$\varphi_{10} + \varphi_{20} + m_{1t}^U = 0; \quad (26)$$

$$\dot{X}_t^0 = a_t X_t^0 + b_t U_{1t}^0. \quad (27)$$

Здесь введены следующие обозначения:

$$\varphi_{10} = \varphi_{10}(\dot{m}_t^X, D_t^X); \quad \varphi_{20} = \varphi_{20}(m_t^X, D_t^X); \quad (28)$$

$$a_t = a_t(m_t^X, \dot{m}_t^X, D_t^X, \dot{D}_t^X, D_t^{U_1}, D_t^{XU_1}) = -k_X^{\varphi_2} (k_X^{\varphi_1})^{-1}; \quad b_t = - (k_X^{\varphi_1})^{-1}.$$

Представим уравнения (27) и (25) для  $U_{1t}^0 U_{1t} - m_{1t}^{U_1}$  для  $\bar{X}_t = [X_t U_{1t}]^T$  в векторном виде:

$$\begin{aligned}\dot{m}_t^{\bar{X}} &= A_{0t} + A_t m_t^{\bar{X}}; \\ \dot{\bar{X}}_t^0 &= A_t \bar{X}_t^0 + B_t V_{1t},\end{aligned}$$

где

$$A_t = \begin{bmatrix} a_t & b_t \\ 0 & \alpha_1 \end{bmatrix}; \quad B_t = \begin{bmatrix} 0 \\ \beta_1 \end{bmatrix}.$$

При этом ковариационная матрица

$$K_t^{\bar{X}} = \begin{bmatrix} D_t^X & K^{XU_1} \\ K_t^{XU_1} & D_t^{U_1} \end{bmatrix}$$

и матрица ковариационных функций

$$K^{\bar{X}}(t_1, t_2) = \begin{bmatrix} K_{11}^{\bar{X}}(t_1, t_2) & K_{12}^{\bar{X}}(t_1, t_2) \\ K_{21}^{\bar{X}}(t_1, t_2) & K_{22}^{\bar{X}}(t_1, t_2) \end{bmatrix}$$

будут удовлетворять линейным уравнениям корреляционной теории [7, 8]:

$$\dot{K}_t^{\bar{X}} = A_t K_t^{\bar{X}} + K_t^{\bar{X}} A_t^T + B_t \nu_{1t} B_t^T, \quad K_{t_0}^{\bar{X}} = K_0^{\bar{X}}; \quad (29)$$

$$\frac{\partial K^{\bar{X}}(t_1, t_2)}{\partial t_2} = K^{\bar{X}}(t_1, t_2) A_{t_2}^T, \quad K^{\bar{X}}(t_1, t_1) = K_{t_1}^{\bar{X}}. \quad (30)$$

Уравнения (29) в развернутом виде имеют вид:

$$\begin{aligned}\dot{D}_t^X &= 2 \left( a_t D_t^X + b_t K_t^{XU_1} \right); \quad \dot{D}_t^{U_1} = 2\alpha_1 D_t^{U_1} + \beta_1^2 \nu_{1t}; \\ \dot{K}_t^{XU_1} &= a_t K_t^{XU_1} + b_t D_t^{U_1} + \alpha_1 K_t^{XU_1}.\end{aligned} \quad (31)$$

В силу (27) после вычисления дисперсии получаем

$$D_t^{\dot{X}} = a_t^2 D_t^X + b_t^2 D_t^{U_1} + 2a_t b_t K_t^{XU_1}. \quad (32)$$

Таким образом, в основе алгоритма аналитического моделирования системы (24), (25) лежат уравнения (26), (28), (29)–(32).

**4.2.** Теперь предположим, что система (24), (25) наблюдается так, что

$$Z_t = \dot{Y}_t = X_t + V_{2t}.$$

Тогда уравнения фильтра Калмана–Бьюси принимают следующий вид:

$$\dot{\hat{X}}_t = A_t \hat{X}_t + \beta_t \left( Z_t - \hat{X}_t \right), \quad \beta_t = R_t \nu_{2t}^{-1} \quad (\det \nu_{2t} \neq 0), \quad \dot{R}_t = 2A_t R_t + \nu_1 - \nu_2 \beta_t^2.$$

Экстраполятор Калмана–Бьюси определяется уравнениями утверждения 3.3 при  $u(t, \tau) = e^{-a(t-\tau)}$ .

## 5 Заключение

Разработано методическое обеспечение для аналитического моделирования, оценивания, идентификации и калибровки существенно нестационарных, например ударных, нормальных СтС, не разрешенных относительно производных, основанное на регрессионной линеаризации. Дан иллюстративный пример.

Полученные результаты могут быть обобщены на случай дифференциальных интегродифференциальных систем, стохастически не разрешенных относительно старших производных, в условиях многоэшелонированных ударных воздействий. Исследование выбросов и накопление возмущений также имеет важное прикладное значение.

## Литература

1. *Синицын И. Н.* Аналитическое моделирование широкополосных процессов в стохастических системах, не разрешенных относительно производных // Информатика и её применения, 2017. Т. 11. Вып. 1. С. 3–10.
2. *Синицын И. Н.* Параметрическое аналитическое моделирование процессов в стохастических системах, не разрешенных относительно производных // Системы и средства информатики, 2017. Т. 27. № 1. С. 21–45.
3. *Синицын И. Н.* Нормальные субоптимальные фильтры для дифференциальных стохастических систем, не разрешенных относительно производных // Информатика и её применения, 2021. Т. 15. Вып. 1. С. 3–10.
4. *Синицын И. Н.* Аналитическое моделирование и фильтрация нормальных процессов в интегродифференциальных стохастических системах, не разрешенных относительно производных // Системы и средства информатики, 2021. Т. 31. № 1. С. 31–56.
5. *Sinitsyn I. N.* Analytical modeling and estimation of normal processes defined by stochastic differential equations with unsolved derivatives // J. Mathematics Statistics Research, 2021. Vol. 3. Iss. 1. Art. 139. 7 p. doi: 10.36266/JMSR/139.
6. *Синицын И. Н.* Нормализация систем, стохастически не разрешенных относительно производных // Информатика и её применения, 2022. Т. 16. Вып. 1. С. 32–38.
7. *Пугачев В. С., Синицын И. Н.* Стохастические дифференциальные системы. Анализ и фильтрация. — М.: Наука, 1990. 632 с.
8. *Пугачев В. С., Синицын И. Н.* Теория стохастических систем. — М.: Логос, 2000; 2004. 1000 с.
9. *Синицын И. Н.* Канонические представления случайных функций и их применение в задачах компьютерной поддержки научных исследований. — М.: ТОРУС ПРЕСС, 2009. 768 с.
10. *Евланов Л. Г., Константинов В. М.* Системы со случайными параметрами. — М.: Наука, 1976. 568 с.
11. *Синицын И. Н.* Фильтры Калмана и Пугачева. — 2-е изд. — М.: Логос, 2007. 776 с.
12. *Синицын И. Н., Шаламов А. С.* Лекции по теории систем интегрированной логистической поддержки. — 2-е изд. — М.: ТОРУС ПРЕСС, 2019. 1072 с.

Поступила в редакцию 26.04.21

## ANALYTICAL MODELING AND ESTIMATION OF NONSTATIONARY NORMAL PROCESSORS WITH UNSOLVED DERIVATIVES

I. N. Sinitsyn<sup>1,2</sup>

<sup>1</sup>Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

<sup>2</sup>Moscow State Aviation Institute (National Research University), 4 Volokolamskoe Shosse, Moscow 125933, Russian Federation

**Abstract:** Methodological and algorithmic support for analytical modeling, estimation, identification, and calibration for essentially nonstationary (e.g., shock) stochastic systems with unsolved derivatives is proposed. Basic theorems are given. Special attention is paid to shock disturbances. An example is provided. Some possible generalizations are presented.

**Keywords:** calibration; estimation (filtering and extrapolation); identification; Kalman and Bucy filter and estimator; shock system with unsolved derivatives; system with stochastically unsolved derivatives

**DOI:** 10.14357/08696527220206

### References

1. Sinitsyn, I. N. 2017. Analiticheskoe modelirovanie shirokopolosnykh protsessov v stokhasticheskikh sistemakh, ne razreshennykh otnositel'no proizvodnykh [Analytical modeling of wide band processes in stochastic systems with unsolved derivatives]. *Informatika i ee Primeneniya — Inform. Appl.* 11(1):3–10.
2. Sinitsyn, I. N. 2017. Parametricheskoe analiticheskoe modelirovanie protsessov v stokhasticheskikh sistemakh, ne razreshennykh otnositel'no proizvodnykh [Parametric analytical modeling of processes in stochastic systems that are not allowed with respect to derivatives]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 27(1):21–45.
3. Sinitsyn, I. N. 2021. Normal'nye suboptimal'nye fil'try dlya differentsial'nykh stokhasticheskikh sistem, ne razreshennykh otnositel'no proizvodnykh [Normal suboptimal filtering for differential stochastic systems with unsolved derivatives]. *Informatika i ee Primeneniya — Inform. Appl.* 15(1):3–10.
4. Sinitsyn, I. N. 2021. Analiticheskoe modelirovanie i fil'tratsiya normal'nykh protsessov v integrodifferentsial'nykh stokhasticheskikh sistemakh, ne razreshennykh otnositel'no proizvodnykh [Analytical modeling and filtering for integrodifferential systems with unsolved derivatives]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 31(1):31–56.
5. Sinitsyn, I. N. 2021. Analytical modeling and estimation of normal processes defined by stochastic differential equations with unsolved derivatives. *J. Mathematics Statistics Research* 3(1):139. 7 p. doi: 10.36266/JMSR/139.

6. Sinitsyn, I. N. 2022. Normalizatsiya sistem, stokhasticheski ne razreshennykh otnositel'no proizvodnykh [Normalization of systems with stochastically unsolved derivatives]. *Informatika i ee Primeneniya — Inform. Appl.* 16(1):32–38.
7. Pugachev, V. S., and I. N. Sinitsyn. 1987. *Stochastic differential systems. Analysis and filtering*. Chichester, NY: J. Wiley & Sons. 549 p.
8. Pugachev, V. S., and I. N. Sinitsyn. 2001. *Stochastic systems. Theory and applications*. Singapore: World Scientific. 908 p.
9. Sinitsyn, I. N. 2009. *Kanonicheskie predstavleniya sluchaynykh funktsiy i ikh primeneniye v zadachakh komp'yuternoy podderzhki nauchnykh issledovaniy* [Canonical expansions of random functions and their applications in computer-aided support]. Moscow: TORUS PRESS. 768 p.
10. Evlanov, L. G., and V. M. Konstantinov. 1976. *Sistemy so sluchaynymi parametrami* [Systems with random parameters]. Moscow: Nauka. 568 p.
11. Sinitsyn, I. N. 2007. *Fil'try Kalmana i Pugacheva* [Kalman and Pugachev filters]. 2nd ed. Moscow: Logos. 776 p.
12. Sinitsyn, I. N., and A. S. Shalamov. 2019. *Leksii po teorii sistem integrirovannoy logisticheskoy podderzhki* [Lectures on theory of integrated logistic support systems]. 2nd ed. Moscow: TORUS PRESS. 1072 p.

*Received April 26, 2021*

## Contributor

**Sinitsyn Igor N.** (b. 1940) — Doctor of Science in technology, professor, honored scientist of RF, principal scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; professor, Moscow State Aviation Institute (National Research University), 4 Volokolamskoe Shosse, Moscow 125933, Russian Federation; [sinitsin@dol.ru](mailto:sinitsin@dol.ru)

## НЕКОТОРЫЕ ПОДХОДЫ К АНАЛИЗУ СЕТЕВЫХ DLP

*А. А. Грушо<sup>1</sup>, Н. А. Грушо<sup>2</sup>, М. И. Забейхайло<sup>3</sup>, Д. В. Смирнов<sup>4</sup>,  
Е. Е. Тимонина<sup>5</sup>, С. Я. Шоргин<sup>6</sup>*

**Аннотация:** Рассматриваются главным образом сетевые системы защиты DLP (Data Lost/Leakage Prevention). Одна из основных задач DLP-систем — выявление ценной информации (ЦИ), которая может появиться в сети на выходе из компьютерной системы. Сетевые DLP-системы приобретают большую важность в связи с широким распространением удаленной работы сотрудников различных организаций. Правильное функционирование DLP определяется распознаванием коротких сообщений как случайных последовательностей, а утечек ЦИ, например персональных данных, — как коротких последовательностей, удовлетворяющих структурным ограничениям. Основная проблема методов защиты информации — ложные тревоги. В работе на простейших моделях оценивается объем данных, в которых возможен поиск ЦИ при малых вероятностях ложных тревог. Построены примеры простых, но быстрых методов выявления утечек ЦИ. Методы разделены по сложности и возможностям их применения. Такие методы могут использоваться не только для анализа DLP-систем, но и для поиска служебных вставок и сигналов в процессе передачи информации, при поиске и анализе плохо защищенных или незащищенных персональных данных, где по правилам требуется обезличивание. Построены оценки вероятностей появления ложных тревог.

**Ключевые слова:** информационная безопасность; DLP-системы; оценки вероятностей появления ложных тревог

**DOI:** 10.14357/08696527220207

### 1 Введение

В последнее время все большее распространение в обеспечении информационной безопасности приобретают системы защиты DLP. Различают три класса DLP-систем: сетевые, агентские и гибридные [1]. В данной работе рассмат-

---

<sup>1</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук, grusho@yandex.ru

<sup>2</sup>Федеральный исследовательский «Информатика и управление» Российской академии наук, info@itake.ru

<sup>3</sup>Федеральный исследовательский «Информатика и управление» Российской академии наук, m.zabekhailo@yandex.ru

<sup>4</sup>ПАО Сбербанк России, dvlsmirnov@sberbank.ru

<sup>5</sup>Федеральный исследовательский «Информатика и управление» Российской академии наук, eltimon@yandex.ru

<sup>6</sup>Федеральный исследовательский «Информатика и управление» Российской академии наук, sshorgin@ipiran.ru

риваются главным образом сетевые DLP-системы. Одна из основных задач DLP-систем — выявление ЦИ, которая может появиться в сети на выходе из компьютерной системы. Сетевые DLP-системы приобретают большую важность в связи с широким распространением удаленной работы сотрудников различных организаций (далее — «удаленка») [2].

Считается, что безопасность «удаленки» должна быть основана на отдельном изолированном компьютере и снабжена связью на основе VPN (Virtual Private Network). Такая же архитектура должна использоваться в облачных системах. Однако указанных правил безопасности недостаточно. Возможны пути внедрения вредоносного кода. Тогда противник может встраивать в выходную последовательность незашифрованные короткие сообщения с ЦИ. Перехватив передачу в глобальной сети, противник получает доступ к украденной ЦИ. В то же время при получении сообщения на легальном приемном конце в процессе расшифрования произойдет зашифрование незашифрованных вставок, что будет рассматриваться как сбой связи и при повторном запросе ликвидируется.

Большой объем трафика не позволяет использовать в DLP-системах сложные алгоритмы и методы. Однако существуют простые методы выявления структур ЦИ, отличных от шифртекста среди последовательности коротких сообщений, или вставок в последовательности [3].

Основная проблема этих методов — ложные тревоги [1–5]. В настоящей работе на простейших моделях оценивается объем данных, в которых возможен поиск ЦИ при малых вероятностях ложных тревог.

Ценная информация должна появляться только в защищенном (зашифрованном) виде. Тогда в качестве модели такой информации можно рассматривать равновероятные последовательности. В качестве альтернативы в этом случае рассматриваются последовательности, удовлетворяющие некоторым ограничениям. Правильное функционирование DLP определяется распознаванием коротких сообщений как случайных последовательностей, а утечек ЦИ, например персональных данных, — как коротких последовательностей, удовлетворяющих структурным ограничениям.

## 2 Структурные модели выявления ценной информации

Определим основные характеристики модели. Система защиты получает на вход большое число  $N$  сообщений малой длины, в которых содержится ЦИ или случайный равновероятный текст. Пусть  $A$  — алфавит, в котором записаны сообщения, мощность множества  $|A| = m$ . В равновероятном случае вероятности появления букв равны  $1/m$  и буквы появляются независимо. Структурные ограничения проще всего задавать в ЦИ в виде простой цепи Маркова, которая описывается матрицей переходных вероятностей цепи Маркова, или ориентированным графом переходов с ненулевыми вероятностями от буквы к букве. Предполагаем, что граф состоит из одной компоненты сильной связности. По определению, последовательность знаков, полностью совпадающая с ориентиро-

ванными дугами в графе, является незащищенной ЦИ, т. е. в рассматриваемом случае структурные особенности ЦИ полностью описываются зависимостью появления букв в тексте ЦИ. Таким образом, ошибки в идентификации ЦИ быть не может. В рассматриваемом случае ошибки могут возникать, когда случайная последовательность удовлетворяет структурным условиям ЦИ, т. е. появляются ложные тревоги.

Оценим вероятность ложной тревоги. Пусть в коротком сообщении имеется последовательность букв

$$x_1, x_2, \dots, x_k. \quad (1)$$

Для простоты оценок предположим, что все исходящие степени вершин графа цепи Маркова (число исходящих дуг из каждой вершины) равны  $s$ , а исходное распределение на буквах в ЦИ имеет положительные вероятности на всех буквах. Тогда число дуг  $n$  в графе связано с параметром  $s$  следующим соотношением:

$$s = \frac{n}{m}.$$

Отсюда можно построить алгоритм вычисления вероятности выявления случайной последовательности (1). На первом шаге алгоритма рассмотрим  $(x_1, x_2)$ . Вероятность выявления случайности последовательности (1) равна

$$P(1) = \frac{m - s}{m}.$$

Соответственно, вероятность не выявить случайную последовательность (1) на первом шаге равна

$$1 - P(1) = \frac{s}{m}.$$

Рассмотрим второй шаг алгоритма при условии, что на первом шаге случайная последовательность (1) не выявлена, т. е. второй шаг алгоритма начинается с вершины  $x_2$  графа. Тогда вероятность попасть из вершины  $x_2$  в вершину  $x_3$  по дуге графа равна  $s/m$ , а вероятность выявления случайной последовательности (1) на этом шаге алгоритма составляет  $(m - s)/m$ . Тогда вероятность выявления случайной последовательности (1) за первые два шага равна

$$P(2) = \frac{m - s}{m} + \frac{s}{m} \frac{m - s}{m}.$$

На третьем шаге алгоритма аналогичным образом получаем формулу:

$$P(3) = \frac{m - s}{m} + \frac{s}{m} \frac{m - s}{m} + \left(\frac{s}{m}\right)^2 \frac{m - s}{m}.$$

Таким образом доказывается, что число шагов до выявления случайности подчиняется геометрическому закону распределения, если граф считать бесконечным.

Полученные формулы позволяют оценить вероятность ошибки, если остановиться на шаге  $k$ . Таким образом, вероятность ложной тревоги оценивается величиной  $(s/m)^{k+1}$ .

По предположению, все сообщения являются короткими. Поэтому  $k$  не может быть большим и при большом числе сообщений  $N$  число ложных тревог будет большим. Среднее число ложных тревог равно  $N(s/m)^{k+1}$ .

Рассмотрим возможность существования нескольких типов описания структуры ЦИ, т. е. предположим, что имеются  $K$  цепей Маркова, описывающих возможные представления ЦИ. Следовательно, имеем  $K$  графов переходов из состояния в состояние различных цепей. Один из подходов к выявлению ложных тревог состоит в распараллеливании оценки случайности сообщения по каждому графу в отдельности. Сложность этого подхода в том, что выявленная случайная дуга в одном графе может оказаться неслучайной дугой в другом графе.

Наиболее продуктивным представляется подход, основанный на предварительной обработке графов. Пусть полученное сообщение начинается с буквы  $x_1$ . Выделим вершину  $x_1$  во всех графах. Множество всех графов разделим на две непересекающиеся группы, а именно: в первую группу поместим все графы, которые имеют дугу из  $x_1$  в вершину  $x_2$  этого графа, где  $x_2$  — вторая буква сообщения. Вторая группа графов будет содержать остальные графы. В графах этой группы случайность можно считать доказанной, так как дуги  $(x_1, x_2)$  нет ни в одном графе. На следующем шаге этого алгоритма разделим графы первой группы на два подмножества по принципу существования дуги из  $x_2$  в вершину  $x_3$  этого графа. Если таких дуг в графе нет, то этот граф относится ко второй группе, где случайность доказана. Если все графы за  $k$  шагов попадают во вторую группу, то случайность последовательности доказана. Если же остались графы в первой группе, то принимается решение о ЦИ. Математическое ожидание вероятности ошибки оценивается величиной  $K(s/m)^k$ . Среднее число ложных тревог оценивается как  $KN(s/m)^k$ . Из этих оценок видно, что в больших данных число ошибок ложных тревог может быть большим.

Рассмотрим следующий уровень структурности — это разбиение на слова. В методах обезличивания (см. токенизацию [6]) сохраняются длины слов. Отсюда становится известной структура разбиения выражения на слова, поэтому можно рассмотреть влияние этой структуры на выявление ложных тревог. Наиболее простой способ — это обобщить на эту структуризацию предыдущие результаты.

Пусть по-прежнему  $A$  — алфавит, появление букв в случайном тексте равновероятно. Пусть последовательность слов в данных  $y_1, y_2, \dots, y_k$ . Обозначим через  $B$  словарь используемых в ЦИ слов. Структурные ограничения опять будем задавать в ЦИ в виде простой цепи Маркова, которая описывается матрицей переходных вероятностей цепи Маркова или ориентированным графом переходов с ненулевыми вероятностями от слова к слову.

Сразу надо отметить, что графов выражений в ЦИ много. Однако для начала рассмотрим один тип выражений, т. е. один граф. Опять для простоты

предположим, что в каждом слове  $b$  букв и вероятность получить случайно данное слово равна  $1/m^b$ . Если  $s$  равно числу слов в ЦИ, в которые может перейти данное слово, то вероятность выявления случайности за  $k$  шагов равна

$$P(k) = \frac{m^b - s}{m^b} + \frac{s}{m^b} \frac{m^b - s}{m^b} + \dots + \left(\frac{s}{m^b}\right)^{k-1} \frac{m^b - s}{m^b}.$$

Отсюда формула для оценки погрешности следующая:  $(s/m^b)^{k+1}$ . Среднее число ложных тревог оценивается величиной  $N(s/m^b)^{k+1}$ . Если число слов, зависимых между собой, велико, то величина  $(s/m^b)^{k+1}$  может иметь значение, при котором число ложных тревог так же быстро растет с ростом  $N$ .

Оценки математического ожидания для  $K$  графов строятся, как было показано выше.

### 3 Неструктурные модели ценной информации

Наиболее удобными для быстрых вычислений неструктурными характеристиками ЦИ служат частотные характеристики. Сюда входят частотные характеристики букв,  $m$ -грамм и частоты встречаемости слов. Даже на небольших участках ЦИ (не превосходящих 100 букв) частоты встречаемости букв легко отличаются от частот равновероятного текста. К сожалению, часто используются простейшие преобразования ЦИ, которые, как считается, делают ЦИ нечитаемой. В задачи этой статьи не входят объяснения того, почему переобозначения букв практически не защищают ЦИ. Объясним, почему после таких преобразований ЦИ легко выявляется на фоне других хорошо защищенных (зашифрованных) сообщений.

Сообщения, которые содержат смысловые характеристики и представлены последовательностями букв, не могут иметь равновероятное распределение вероятностей букв, а значит, замена исходных букв на другие значки является преобразованием, инвариантным относительно вероятностей букв в исходном тексте, т. е. вероятности появления значков в новом тексте совпадают с вероятностями исходных букв. Все эти вероятности можно оценить статистически, и эти оценки будут близки к неравновероятным значениям вероятностей букв в исходном тексте. Такая неравновероятность легко выявляется на любом преобразованном тексте, например с помощью критерия  $\chi^2$  [7]. Использование небуквенных статистик ( $m$ -граммы) требует значительного увеличения длин текстов, что трудно получать на коротких сообщениях.

Рассматривая короткие сообщения как единое целое, можно получить оценку  $N$ , для которой с вероятностью, стремящейся к единице, возникают ложные тревоги. Допустим, что все выражения, в том числе содержащие ЦИ, имеют одну длину  $k$ . Для равновероятного текста вероятность конкретного сообщения равна  $1/m^k$ . Тогда вероятность не появления фиксированного сообщения равна  $(1 - 1/m^k)^N$ .

Среди всех возможных комбинаций из  $k$  букв есть такая, которая по частотным характеристикам букв или слов соответствует ЦИ. Ясно, что появление слов с частотными характеристиками, соответствующими ошибке критерия выявления ЦИ, имеет бóльшую вероятность, чем вероятность выбрать данную последовательность, соответствующую ЦИ. С другой стороны, при фиксированном алфавите и длине коротких сообщений  $k$  имеет место неравенство:

$$\left(1 - \frac{1}{m^k}\right) < 1.$$

Тогда при больших  $N$  величина  $(1 - 1/m^k)^N$  экспоненциально быстро стремится к 0. Это значит, что вероятность появления ложной тревоги для любого частотного метода быстро сходится к 1.

Если типов сообщений  $K$ , то среднее число ложных тревог равно  $K(1 - (1 - 1/m^k)^N)$ .

Возможны другие способы простых преобразований, которые позволяют выявлять слабо преобразованную ЦИ. Приведем следующий пример. Стандартная запись действия в некоторой информационной технологии может быть хорошо зашифрована, но индивидуальные особенности конкретного действия могут передаваться открытыми вставками. И наоборот: в качестве защиты ЦИ могут использоваться «зашумляющие» вставки.

Подходы к выявлению ЦИ при наличии «зашумляющих вставок» исследовались в работе [8].

Важной проблемой в DLP-системах остаются скрытые каналы. Примеры скрытых каналов, никак не связанных со структурой ЦИ и функционирующих в полностью зашифрованном сообщении, можно найти в работе [9].

## 4 Заключение

В работе построены примеры простых, но быстрых методов выявления утечек ЦИ. Методы разделены по сложности и возможностям их применения. Такие методы могут использоваться не только для анализа DLP-систем, но также для поиска служебных вставок и сигналов в процессе передачи информации, при поиске и анализе плохо защищенных или незащищенных персональных данных, где по правилам требуется обезличивание. Объединяющим моментом моделей являются оценки объемов данных, на которых могут работать построенные методы [4–6]. Для формирования матриц переходов (графов) цепей Маркова можно использовать методы машинного обучения на соответствующих открытых компьютерных данных.

В работе не приводятся примеры конкретных расчетов, так как формы сообщений в разных организациях разные, но представленные методы оценок в некотором смысле универсальные.

Построены оценки вероятностей появления ложных тревог.

## Литература

1. Степанов И. Сравнительный обзор средств предотвращения утечек данных (DLP) // Safe-surf, 12.03.2019. [https://safe-surf.ru/specialists/article/5233/609990/?sphrase\\_id=45457](https://safe-surf.ru/specialists/article/5233/609990/?sphrase_id=45457).
2. Блинов А. Обзор «Российский рынок DLP-систем 2021. Проблемы и решения». Часть 1. Особенности современного рынка DLP-систем // ICT-Online, 10.09.2021. <https://ict-online.ru/analytics/a200309>.
3. Королев В. В. Использование методов анализа контента в DLP системах // Проблемы науки, 2016. № 10(11). С. 16–20.
4. Грушо А. А., Грушо Н. А., Забейхайло М. И., Смирнов Д. В., Тимонина Е. Е., Шоргин С. Я. Статистика и кластеры в поисках аномальных вкраплений в условиях больших данных // Информатика и её применения, 2021. Т. 15. Вып. 4. С. 79–86.
5. Грушо А. А., Грушо Н. А., Забейхайло М. И., Смирнов Д. В., Тимонина Е. Е., Шоргин С. Я. Поиск аномалий в больших данных // Системы и средства информатики, 2022. Т. 32. № 1. С. 160–167.
6. Грушо А. А., Смирнов Д. В., Тимонина Е. Е., Шоргин С. Я. Усиленный алгоритм токенизации для защиты персональных данных // Системы и средства информатики, 2021. Т. 31. № 4. С. 135–143.
7. Ширяев А. Н. Вероятность. — 3-е изд. — М.: МЦНМО, 2004. 521 с.
8. Мельников С. Ю. Методы распознавания и идентификации конечных автоматов по статистическим характеристикам выходных и входных последовательностей: Дис. . . . докт. физ.-мат. наук. — М.: РУДН, 2021. 265 с.
9. Тимонина Е. Е. Анализ угроз скрытых каналов и методы построения гарантированно защищенных распределенных автоматизированных систем: Дис. . . . докт. техн. наук. — М.: РГГУ, 2004. 204 с.

*Поступила в редакцию 22.03.22*

---

## SOME APPROACHES TO NETWORK DLP ANALYSIS

*A. A. Grusho<sup>1</sup>, N. A. Grusho<sup>1</sup>, M. I. Zabezhalo<sup>1</sup>, D. V. Smirnov<sup>2</sup>, E. E. Timonina<sup>1</sup>,  
and S. Ya. Shoragin<sup>1</sup>*

<sup>1</sup>Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

<sup>2</sup>Sberbank of Russia, 19 Vavilov Str., Moscow 117999, Russian Federation

**Abstract:** The work deals mainly with network DLP (Data Lost/Leakage Prevention) systems. One of the main tasks of DLP systems is to recognize the appearance of valuable information that can appear on the network at the output of the computer system. Network-based DLP systems are becoming more important due to the development of remote work of employees of various organizations. The proper functioning of the DLP is determined by recognizing short messages as random sequences and leaking valuable information such as

personal data as a short sequence satisfying structural constraints. The main problem of these methods is false alarms. In the work, in the simplest models, the volume of data is estimated in which it is possible to search for valuable information at small probabilities of false alarms. The examples of simple but fast methods for detecting leaks of valuable information are built. The methods are divided by complexity and their applications. Such methods can be used not only to analyze DLP systems, but also to search for service inserts and signals during the transfer of information when searching and analyzing poorly protected or unprotected personal data where depersonalization is required according to the rules. The estimates of the probabilities of false alarms are built.

**Keywords:** information security; DLP systems; estimates of probabilities of false alarms

**DOI:** 10.14357/08696527220207

## References

1. Sravnitel'nyy obzor sredstv predotvrashcheniya utechek dannykh (DLP) [Comparison review of Data Leakage Prevention (DLP) tools]. 12.03.1019. *Safe-surf*. Available at: [https://safe-surf.ru/specialists/article/5233/609990/?sphrase\\_id=45457](https://safe-surf.ru/specialists/article/5233/609990/?sphrase_id=45457) (accessed March 29, 2022).
2. Blinov, A. 10.09.2021. Obzor "Rossiyskiy rynek DLP-sistem 2021. Problemy i resheniya." Chast' 1. Osobennosti sovremennogo rynka DLP-sistem [Overview "Russian market of DLP-systems 2021. Problems and solutions." Part 1. Features of today's DLP market]. *ICT-Online*. Available at: <https://ict-online.ru/analytics/a200309/> (accessed March 29, 2022).
3. Korolev, V. V. 2016. Ispol'zovanie metodov analiza kontenta v DLP sistemakh [Using content analysis methods in DLP systems]. *Problemy nauki* [Problems of Science] 10(11):16–20.
4. Grusho, A. A., N. A. Grusho, M. I. Zabezhailo, D. V. Smirnov, E. E. Timonina, and S. Ya. Shorgin. 2021. Statistika i klasteriy v poiskakh anomal'nykh vkrapleniy v usloviyakh bol'shikh dannykh [Statistics and clusters for detection of anomalous insertions in big data environment]. *Informatika i ee Primeneniya — Inform. Appl.* 15(4):79–86.
5. Grusho, A. A., N. A. Grusho, M. I. Zabezhailo, D. V. Smirnov, E. E. Timonina, and S. Ya. Shorgin. 2022. Poisk anomalii v bol'shikh dannykh [Search of anomalies in big data]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 32(1):160–167.
6. Grusho, A. A., D. V. Smirnov, E. E. Timonina, and S. Ya. Shorgin. 2021. Usilennyi algoritm tokenizatsii dlya zashchity personal'nykh dannykh [Enhanced tokenization algorithm for personal data protection]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 31(4):135–143.
7. Shiryayev, A. N. 2004. *Veroyatnost'* [Probability]. Moscow: MTsNMO. 521 p.
8. Melnikov, S. Yu. 2021. Metody raspoznavaniya i identifikatsii konechnykh avtomatov po statisticheskim kharakteristikam vykhodnykh i vkhodnykh posledovatel'nostey [Methods of recognition and identification of finite automata by statistical characteristics of output and input sequences]. Moscow: Peoples' Friendship University of Russia. D. Sc. Diss. 265 p.

9. Timonina, E. E. 2004. Analiz ugroz skrytykh kanalov i metody postroeniya garantirovanno zashchishchennykh raspredelennykh avtomatizirovannykh sistem [The analysis of threats of covert channels and methods of creation of guaranteed protected distributed automated systems]. Moscow: Russian State University for the Humanities. D. Sc. Diss. 204 p.

*Received March 22, 2022*

## **Contributors**

**Grusho Alexander A.** (b. 1946) — Doctor of Science in physics and mathematics, professor, principal scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; grusho@yandex.ru

**Grusho Nikolai A.** (b. 1982) — Candidate of Science (PhD) in physics and mathematics, senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; info@itake.ru

**Zabezhalo Michael I.** (b. 1956) — Doctor of Science in physics and mathematics, principal scientist, A. A. Dorodnicyn Computing Center, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 40 Vavilov Str., Moscow 119333, Russian Federation; m.zabezhalo@yandex.ru

**Smirnov Dmitry V.** (b. 1984) — business partner for IT security department, Sberbank of Russia, 19 Vavilov Str., Moscow 117999, Russian Federation; dvlsmirnov@sberbank.ru

**Timonina Elena E.** (b. 1952) — Doctor of Science in technology, professor, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; eltimon@yandex.ru

**Shorgin Sergey Ya.** (b. 1952) — Doctor of Science in physics and mathematics, professor, principal scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; sshorgin@ipiran.ru

## АРХИТЕКТУРЫ ИНТЕЛЛЕКТУАЛЬНЫХ АГЕНТОВ СПЛОЧЕННЫХ ГИБРИДНЫХ ИНТЕЛЛЕКТУАЛЬНЫХ МНОГОАГЕНТНЫХ СИСТЕМ\*

*С. В. Листонад<sup>1</sup>, И. А. Кириков<sup>2</sup>*

**Аннотация:** Работа развивает исследования в области моделирования сплоченности коллектива специалистов, решающих проблемы за круглым столом, гибридными интеллектуальными многоагентными системами (ГиИМАС). Формирование сплоченности агентов этих систем обеспечивается такими механизмами, как обмен информацией об интересах, знаниях, опыте с другими агентами, выработка и согласование единых для системы целей, ценностей, норм, средств и способов деятельности, что релевантно процессам, происходящим в длительно существующих коллективах специалистов, совместно решающих практические проблемы. Реализация подобных механизмов позволит существенно упростить процесс интеграции разнородных интеллектуальных агентов в единую систему, что особенно актуально при их разработке независимыми командами разработчиков. В статье предлагаются архитектуры агентов, необходимые для реализации перечисленных механизмов.

**Ключевые слова:** сплоченность; гибридная интеллектуальная многоагентная система; коллектив специалистов

**DOI:** 10.14357/08696527220208

### 1 Введение

Одна из критических точек в построении искусственных интеллектуальных систем, способных решать практические проблемы без их существенного упрощения, — компьютерное моделирование совместной работы специалистов в коллективах. Актуальность такого моделирования продиктована как неотъемлемыми особенностями проблемы (непрозрачность, неоднородность, полителичность и т. д.) [1, 2], так и невозможностью построения всезнающей централизованной интеллектуальной системы, обладающей всеми необходимыми ресурсами, знаниями и инструментами для решения разнообразных проблем [3].

С целью компьютерного моделирования работы коллектива специалистов по решению проблем за круглым столом в [4] предложен подход на основе ГиИМАС. Элементы таких систем — агенты, т. е. относительно автономные программные

---

\* Исследование выполнено при финансовой поддержке РФФИ (проект 20-07-00104а).

<sup>1</sup> Калининградский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, ser-list-post@yandex.ru

<sup>2</sup> Калининградский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, baltbipiran@mail.ru

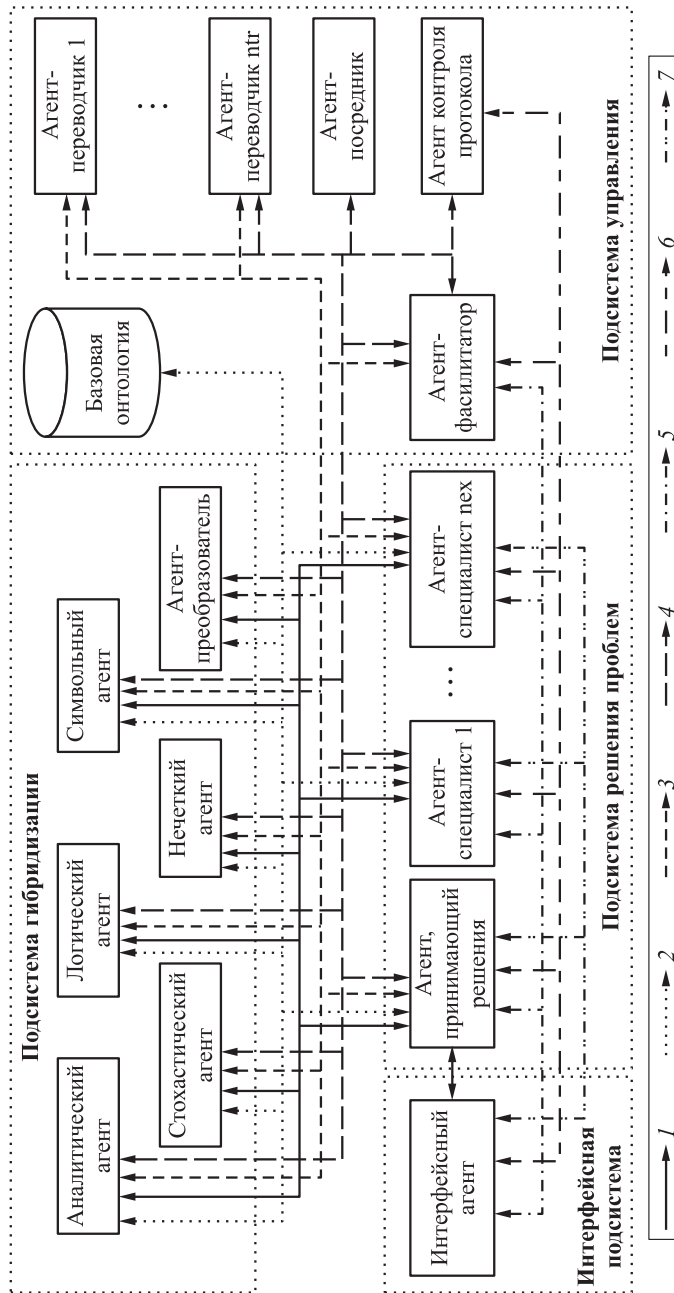
сущности, характеризующиеся проактивностью, реактивностью и социальным поведением [5–8], наличием механизмов целеполагания и методов различных интеллектуальных технологий, а также развитыми моделями предметной области, необходимыми для имитации интеллектуальной деятельности специалистов по решению проблем и их частей. Недостаток предложенного в [4] подхода — необходимость определения на этапе проектирования общей для всех агентов модели предметной области и протокола взаимодействия, а также спецификации целевых функций всех агентов. Такой подход релевантен ситуации, когда ГиИМАС и все ее агенты проектируются одной командой разработчиков. Если проблемы, для решения которых создается ГиИМАС, настолько сложны, что в ее составе должны присутствовать агенты, построенные разными разработчиками, требуются значительные трудозатраты для их объединения в целостную систему. Для упрощения этой процедуры в [9] предложена концепция сплоченных ГиИМАС (СГИМАС), агенты которых самостоятельно согласовывают свои цели, модели предметной области и версии протокола решения проблемы. Цель настоящей работы — разработка архитектур агентов СГИМАС, обеспечивающих моделирование перечисленных механизмов.

## **2 Функциональная структура сплоченной гибридной интеллектуальной многоагентной системы**

Для компьютерного моделирования сплоченности коллектива специалистов в [9] предложена функциональная структура СГИМАС, представленная на рис. 1. В ней выделяются четыре подсистемы: интерфейсная, гибридизации, решения проблем и управления. Интерфейсная подсистема состоит из единственного агента, отвечающего за взаимодействие с пользователем. Он запрашивает входные данные, выдает результат и визуализирует процессы в СГИМАС.

Подсистема управления содержит базовую онтологию, множество агентов-переводчиков, агента-посредника, агента-фасилитатора и агента контроля протокола. Базовая онтология — семантическая сеть, основа взаимодействия агентов, обеспечивающая понимание ими смысла передаваемых сообщений по согласованию собственных моделей предметной области и целей и формированию протокола сплоченного взаимодействия. Агенты-переводчики предназначены для трансляции передаваемых агентами сообщений с одного языка на другой [6]. Агент-посредник отслеживает имена, модели и возможности зарегистрированных агентов. Агент-фасилитатор отвечает за организацию эффективной коллективной работы агентов и запуск релевантных ситуаций процессов переговоров о целях, моделях предметной области или протоколах решения проблемы. Агент контроля протокола фиксирует изменения протокола решения проблемы, принятые агентами подсистемы решения проблемы, а также контролирует его соблюдение.

Агенты подсистемы гибридизации не имеют развитой модели предметной области или целеполагания. Они выполняют поручения других агентов, комби-



**Рис. 1** Функциональная структура СТИМАС: 1 — отношения агентов: запросы информации, запросы помощи в решении задач, передача результатов их решения; 2 — взаимодействие (получение сведений из модели, обновление модели) агентов с базовой онтологией; 3 — взаимодействие агентов: запросы перевода сообщений с одного языка кодирования сообщений на другой; 4 — отношения агентов: запросы имен и адресов агентов с заданными возможностями; 5 — отношения фасилитации, инициализирующие процессы согласования целей, моделей предметной области и протокола взаимодействия агентами подсистемы решения проблем; 6 — отношения выработки и контроля соблюдения решения проблемы; 7 — отношения согласования целей и моделей предметной области агентами подсистемы решения проблем

нируя разнородные знания с использованием следующих моделей и алгоритмов: алгебраических уравнений для описания причинно-следственных связей концептов предметной области; метода Монте-Карло; продукционной экспертной системы с рассуждениями в прямом направлении; алгоритма нечеткого вывода Мамдани и др.

Агенты подсистемы решения проблем моделируют рассуждения коллектива специалистов под руководством лица, принимающего решения, за круглым столом по решению поставленной проблемы. Агент, принимающий решения, ставит задачи агентам-специалистам, собирает результаты работы, определяет, достигнут ли критерий останова, и либо принимает итоговое решение, либо запускает новую итерацию процесса решения проблемы. Агенты-специалисты генерируют решения частей проблемы или альтернативные решения проблемы в целом в зависимости от поставленной им задачи.

Рассмотрим архитектуры агентов, участвующих в процессах формирования сплоченности: агента-фасилитатора, агента контроля протокола, агента, принимающего решения, и агентов-специалистов.

### 3 Архитектуры интеллектуальных агентов, обеспечивающих моделирование сплоченности

Архитектура агента-фасилитатора представлена на рис. 2. Маршрутизатор сообщений СГИМАС не является частью агента, а представляет собой подсистему обмена сообщениями программной платформы, на которой реализована СГИМАС, и обеспечивает доставку сообщений между агентами. Подсистема приема/отправки сообщений позволяет агентам взаимодействовать друг с другом, формируя сообщения на языке передачи сообщений, например QXML [10] или FIPA ACL [11]. Интерпретатор базовой онтологии обеспечивает гене-

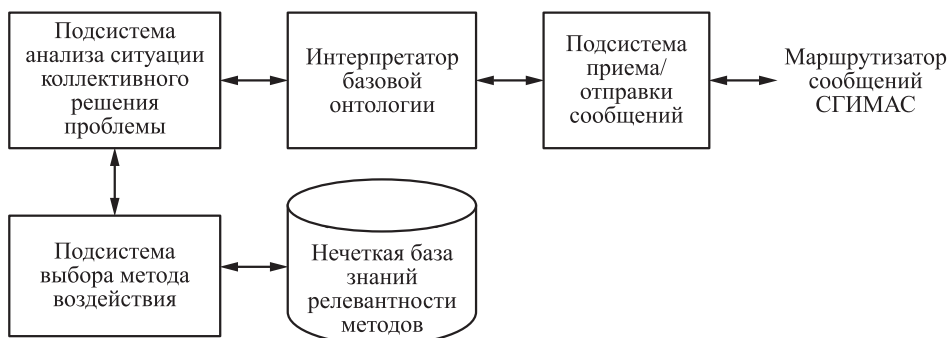
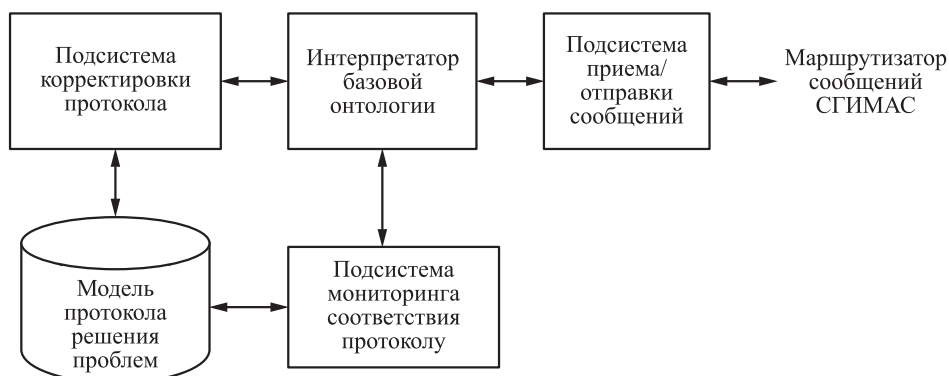


Рис. 2 Архитектура агента-фасилитатора

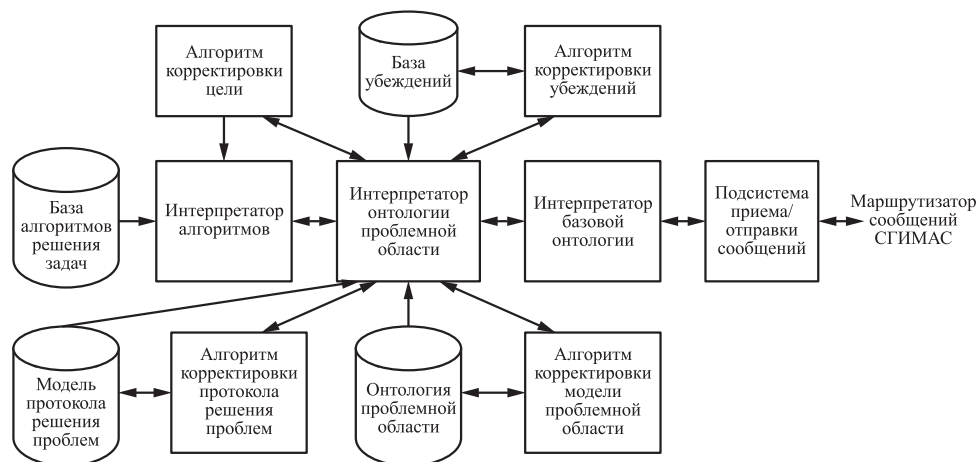
рацию тела сообщения с использованием базовой онтологии СГИМАС (см. рис. 1), анализирует семантику тела разобранных сообщений из подсистемы приема/отправки сообщений, генерирует программные объекты на основе тела сообщения и базовой онтологии СГИМАС и направляет их в соответствующую подсистему согласно типу и содержанию. Это типовые подсистемы для агентов СГИМАС, и при рассмотрении архитектур других агентов они опускаются.

Если базовый интерпретатор онтологии не может определить семантику полученного сообщения, оно передается для анализа интерпретатору онтологии предметной области. Подсистема анализа ситуации коллективного решения проблемы идентифицирует текущий этап процесса решения проблемы с учетом степени согласованности целей и онтологии предметной области агентов, а также наличия согласованного протокола решения проблемы. Подсистема выбора метода воздействия с использованием нечеткой базы знаний имитирует работу фасилитатора по выбору средств активизации командной работы и разрешения конфликтов группы специалистов, релевантных ситуации коллективного решения проблем. Нечеткая база знаний релевантности методов описывает правила выбора метода воздействия на агентов подсистемы решения проблем в зависимости от ситуаций принятия решений в СГИМАС, а также различных особенностей проблемы. Для формирования такой базы необходимо провести серию вычислительных экспериментов и установить соответствие между классом проблемы и методами ее решения.

Архитектура агента контроля протокола представлена на рис. 3. Модель протокола решения проблем фиксирует договоренности агентов о согласованном протоколе решения проблем на текущий момент. Подсистема корректировки протокола отслеживает сообщения, связанные с разработкой и модификацией согласованного протокола решения проблем, и фиксирует эти изменения в мо-



**Рис. 3** Архитектура агента контроля протокола



**Рис. 4** Архитектура агента-специалиста

дели протокола. Подсистема мониторинга соответствия протоколу отслеживает все сообщения, которые агенты-специалисты и агент, принимающий решения, отправляют во время решения проблемы, и проверяет их соответствие протоколу. Если она обнаруживает нарушение протокола, инициируется отправка соответствующего сообщения агенту-посреднику и агенту интерфейса через базовый интерпретатор онтологии и подсистему приема/отправки сообщений.

Архитектура агента-специалиста представлена на рис. 4. Интерпретатор онтологии проблемной области в зависимости от содержания полученного сообщения запускает один из следующих алгоритмов: решения проблемы с использованием интерпретатора алгоритмов; корректировки цели; корректировки убеждений; корректировки протокола решения проблем; корректировки модели проблемной области. В качестве аргументов запускаемому алгоритму могут быть переданы следующие данные: содержимое полученного сообщения в структурированном виде; информация о цели агента; его убеждения; модель проблемы; модель протокола решения проблемы.

В базе алгоритмов решения задач хранится набор алгоритмов, которые агент-специалист выполняет с помощью интерпретатора при решении задачи с учетом своей цели. Они могут содержать запросы к агентам интеллектуальных технологий для выполнения определенных функций. Чтобы найти агента интеллектуальной технологии, способного выполнять такие функции, интерпретатор алгоритмов отправляет запрос агенту-посреднику через интерпретатор онтологии предметной области, базовый интерпретатор онтологии и подсистему приема/отправки сообщений. Алгоритм корректировки цели позволяет корректировать целевую функцию агента в результате взаимодействия с другими агентами. База убеждений содержит убеждения агента относительно онтологии, которые могут

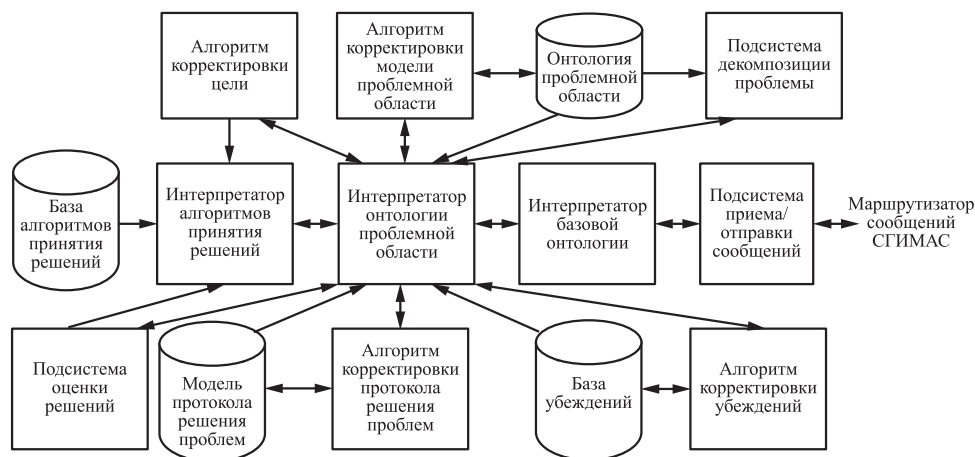


Рис. 5 Архитектура агента, принимающего решения

быть дополнены и скорректированы алгоритмом корректировки убеждений в результате взаимодействия с другими агентами и моделью предметной области. Модель протокола решения проблем хранит текущее состояние модели протокола, разработанное агентами подсистемы решения проблем. Она может изменяться с помощью алгоритма корректировки протокола решения проблем в результате переговоров агентов. Онтология проблемной области — это семантическая сеть, основанная на концептуальной модели решаемой проблемы. Данная онтология разрабатывается для каждого агента-специалиста индивидуально в соответствии с профилем специалиста, знания которого моделируются агентом. В общем случае разработку агентов-специалистов и их онтологий могут осуществлять разные независимые команды, поэтому онтологии могут различаться и даже противоречить друг другу. Онтология проблемной области не статична, а изменяется алгоритмом корректировки проблемной области в процессе решения проблемы в результате переговоров между агентами подсистемы решения проблем. При этом, хотя онтологии проблемных областей различных агентов сближаются в ходе их переговоров, они не становятся идентичными.

Архитектура агента, принимающего решения, представлена на рис. 5.

Подсистема декомпозиции проблемы, основанная на анализе онтологии предметной области, распределяет между агентами-специалистами задачи поставленной проблемы и исходные данные, необходимые для их решения. Подсистема оценки решений в соответствии с собственной целью рассчитывает показатель качества решений, представленных агентами-специалистами. Если решения удовлетворительного качества не найдены, эта подсистема запускает новую итерацию решения проблемы СГИМАС. В базе алгоритмов принятия решений хранится набор алгоритмов, используемых агентом при выборе окончательного

решения в зависимости от протокола решения задач. Интерпретатор алгоритмов принятия решений — это подсистема для выполнения алгоритмов принятия решений. Остальные блоки аналогичны блокам архитектуры агента-специалиста (см. рис. 4). Первоначально цели, убеждения, модели проблемной области и протокола решения проблем для агентов-специалистов и агента, принимающего решения, устанавливаются их разработчики, однако, работая над проблемами, эти агенты обмениваются данными и знаниями, подтверждающими предлагаемые решения проблемы, и при необходимости изменяют свои цели, убеждения и модели. Таким образом, имея возможность запросить помощь в решении проблемы у любого из агентов интеллектуальных технологий, агенты-специалисты каждый раз, решая новую проблему, формируют новый релевантный ей интегрированный метод решения, демонстрируя признаки «сильной» самоорганизации, возникающей в результате распределенного взаимодействия агентов без явного централизованного управления [12], что релевантно коллективам специалистов, решающим проблемы за круглым столом.

## 4 Заключение

В работе показана необходимость компьютерного моделирования механизмов формирования сплоченности с целью снижения трудозатрат при интеграции агентов, созданных разными коллективами разработчиков. Рассмотрена функциональная структура СГИМАС. Предложены архитектуры интеллектуальных агентов, обеспечивающих моделирование сплоченности. Наличие блоков согласования целей, убеждений, моделей проблемной области и протокола решения проблем у агентов-специалистов и агента, принимающего решения, позволяет им вырабатывать решения с учетом не только собственного «видения проблемы», предопределенного разработчиками при моделировании знаний и поведения соответствующего специалиста, но и «точки зрения» агентов, моделирующих других специалистов. Это позволяет снизить вероятность досрочного завершения работы над проблемой и принятия никого не устраивающего решения.

## Литература

1. Спиридонов В. Ф. Психология мышления: решение задач и проблем. — М.: Генезис, 2006. 319 с.
2. Seel N. M. Problems: Definition, types, and evidence // Encyclopedia of the sciences of learning / Ed. N. M. Seel. — Boston, MA, USA: Springer, 2012. P. 2690–2693.
3. Тарасов В. Б. Агенты, многоагентные системы, виртуальные сообщества: стратегическое направление в информатике и искусственном интеллекте // Новости искусственного интеллекта, 1998. № 2. С. 5–63.
4. Колесников А. В., Кириков И. А., Листопад С. В. Гибридные интеллектуальные системы с самоорганизацией: координация, согласованность, спор. — М.: ИПИ РАН, 2014. 189 с.

5. Wooldridge M., Jennings N. Intelligent agents: Theory and practice // Knowl. Eng. Rev., 1995. Vol. 10. Iss. 2. P. 115–152.
6. Городецкий В. И., Грушинский М. С., Хабалов А. В. Многоагентные системы (обзор) // Новости искусственного интеллекта, 1998. № 2. С. 64–116.
7. Хорошевский В. Ф. Поведение интеллектуальных агентов: модели и методы реализации // 4-й Междунар. семинар по прикладной семиотике, семиотическому и интеллектуальному управлению: Сб. научных трудов. — Переславль-Залесский: РАИИ, 1999. С. 5–20.
8. Hajduk M., Sukop M., Haun M. Multi-agent systems — terminology and definitions // Cognitive multi-agent systems. — Studies in systems, decision and control ser. — Cham: Springer, 2019. Vol. 138. P. 1–9.
9. Лисопад С. В. Функциональная структура сплоченной гибридной интеллектуальной многоагентной системы // Системы и средства информатики, 2020. Т. 30. № 2. С. 78–88. doi: 10.14357/08696527200208.
10. Finin T., Fritzson R., McKay D., McEntire R. KQML as an agent communication language // 3rd Conference (International) on Information and Knowledge Management Proceedings. — New York, NY, USA: Association for Computing Machinery, 1994. P. 456–463.
11. FIPA ACL Message Structure Specification. <http://www.fipa.org/specs/fipa00061/index.html>.
12. Serugendo G. D. M., Gleizes M.-P., Karageorgos A. Self-organization in multiagent systems // Knowl. Eng. Rev., 2005. Vol. 20. Iss. 2. P. 165–189.

Поступила в редакцию 04.03.22

---

---

## ARCHITECTURES OF INTELLIGENT AGENTS OF COHESIVE HYBRID INTELLIGENT MULTIAGENT SYSTEMS

*S. V. Listopad and I. A. Kirikov*

Kaliningrad Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 5 Gostinaya Str., Kaliningrad 236000, Russian Federation

**Abstract:** The paper develops research in the field of modeling the cohesion of a team of specialists who solve problems “at a round table” with hybrid intelligent multiagent systems that combine the dynamic synthesis of heterogeneous models of intelligent agents and the methods of group dynamics. The formation of the cohesion of agents of these systems is provided by such mechanisms as exchange of information about interests, knowledge, experience with other agents, development and coordination of goals, values, norms, means, and methods of activity common for the system which is relevant to the processes taking place in long-term teams of specialists who jointly solve practical problems.

The implementation of such mechanisms will significantly simplify the process of integrating heterogeneous intelligent agents into a single system which is especially important when they are created by independent development teams. The paper proposes the architectures of agents necessary for the implementation of the listed mechanisms.

**Keywords:** cohesion; hybrid intelligent multiagent system; team of specialists

**DOI:** 10.14357/08696527220208

## Acknowledgments

The reported study was funded by RFBR, project number 20-07-00104a.

## References

1. Spiridonov, V. F. 2006. *Psikhologiya myshleniya: reshenie zadach i problem* [The psychology of mind: Solving of tasks and problems]. Moscow: Genesis. 319 p.
2. Seel, N. M. 2012. Problems: Definition, types, and evidence. *Encyclopedia of the sciences of learning*. Ed. N. M. Seel. Boston, MA: Springer. 2690–2693.
3. Tarasov, V. B. 1998. Agenty, mnogoagentnye sistemy, virtual'nye soobshchestva: strategicheskoe napravlenie v informatike i iskusstvennom intellekte [The agents, multiagent system, virtual communities: Strategic direction in computer science and artificial intelligence]. *Novosti iskusstvennogo intellekta* [News of Artificial Intelligence] 2:5–63.
4. Kolesnikov, A. V., I. A. Kirikov, and S. V. Listopad. 2014. *Gibridnye intellektual'nye sistemy s samoorganizatsiyey: koordinatsiya, soglasovannost', spor* [Hybrid intelligent systems with self-organization: Coordination, consistency, and dispute]. Moscow: IPI RAN. 189 p.
5. Wooldridge, M., and N. Jennings. 1995. Intelligent agents: Theory and practice. *Knowl. Eng. Rev.* 10(2):115–152.
6. Gorodetskiy, V. I., M. S. Grushinskiy, and A. V. Khabalov. 1998. Mnogoagentnye sistemy (obzor) [Multiagent systems (review)]. *Novosti iskusstvennogo intellekta* [News of Artificial Intelligence] 2:64–116.
7. Horoshevskij, V. F. 1999. Povedenie intellektual'nykh agentov: Modeli i metody realizatsii [The behavior of intelligent agents: Models and methods of implementation]. *4th Workshop (International) on Applied Semiotics, Semiotics, and Intelligent Management Proceedings*. Pereslavl'-Zalesskiy: RAAI. 5–20.
8. Hajduk, M., M. Sukop, and M. Haun. 2019. Multi-agent systems — terminology and definitions. *Cognitive multi-agent systems*. Studies in systems, decision and control ser. Cham: Springer. 138:1–9.
9. Listopad, S. V. 2020. Funktsional'naya struktura splochennoy gibridnoy intellektual'noy mnogoagentnoy sistemy [Functional structure of the cohesive hybrid intelligent multiagent system]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(2):78–88. doi: 10.14357/086965272200208.
10. Finin, T., R. Fritzson, D. McKay, and R. McEntire. 1994. KQML as an agent communication language. *3rd Conference (International) on Information and Knowledge Management Proceedings*. New York, NY: ACM. 456–463.

11. FIPA ACL Message Structure Specification. Available at: <http://www.fipa.org/specs/fipa00061/index.html> (accessed March 15, 2022).
12. Serugendo, G. D. M., M.-P. Gleizes, and A. Karageorgos. 2005. Self-organization in multiagent systems. *Knowl. Eng. Rev.* 20(2):165–189.

*Received March 4, 2022*

## Contributors

**Listopad Sergey V.** (b. 1984) — Candidate of Science (PhD) in technology, senior scientist, Kaliningrad Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 5 Gostinaya Str., Kaliningrad 236000, Russian Federation; ser-list-post@yandex.ru

**Kirikov Igor A.** (b. 1955) — Candidate of Science (PhD) in technology, director, Kaliningrad Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 5 Gostinaya Str., Kaliningrad 236000, Russian Federation; baltbipiran@mail.ru

## ИССЛЕДОВАНИЕ МЕТОДОВ ДЕТЕКТИРОВАНИЯ КЛЮЧЕВЫХ ТОЧЕК ПРИ СОЗДАНИИ ПАНОРАМНЫХ ИЗОБРАЖЕНИЙ

*П. О. Архипов<sup>1</sup>, А. К. Трофименков<sup>2</sup>, М. В. Цуканов<sup>3</sup>, Н. Ю. Носова<sup>4</sup>*

**Аннотация:** Статья посвящена проблеме создания панорам в условиях отличия цветовой палитры и уровня освещенности сшиваемых снимков, вызванных как особенностями оптической системы, так и погодой или иными факторами. Предложена процедура коррекции яркости и цвета для минимизации отличий между сшиваемыми снимками и уменьшения негативных воздействий. Приведены особенности детектирования ключевых точек различными методами во время сшивания панорамных изображений в условиях разнорельефной съемки. Описано влияние дополнительной процедуры коррекции яркости и цвета на число найденных ключевых точек. Оценено качество итоговых панорамных изображений и время их сшивания. В ходе работы получены статистические данные и проведен их сравнительный анализ.

**Ключевые слова:** коррекция яркости; коррекция цвета; детектирование; ключевые точки; панорамные изображения; обработка изображений

**DOI:** 10.14357/08696527220209

### 1 Введение

Технология обработки данных аэрофотосъемки беспилотными летательными аппаратами (БПЛА) наземных объектов, включающая построение фотопланов и панорам, сегментацию, обнаружение аномалий, коррекцию яркости и цвета при создании бесшовных панорамных изображений востребована для решения многих фундаментальных и прикладных задач.

Одна из основных проблем обработки панорам, полученных с БПЛА, состоит в повышении достоверности и правильной интерпретации изображений на основе визуального анализа изображений квалифицированным специалистом, поддержки принятия решений за ограниченное время и оперативного контроля (мониторинга) территорий.

В настоящее время дистанционный мониторинг объектов земной поверхности необходимо осуществлять с требуемым качеством (достоверностью, своевремен-

<sup>1</sup>Орловский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, [apaul@mail.ru](mailto:apaul@mail.ru)

<sup>2</sup>Орловский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, [andretrf@yandex.ru](mailto:andretrf@yandex.ru)

<sup>3</sup>Орловский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, [tsukanov.m.v@yandex.ru](mailto:tsukanov.m.v@yandex.ru)

<sup>4</sup>Орловский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, [ofipiran.lab01@yandex.ru](mailto:ofipiran.lab01@yandex.ru)

ностью, точностью), используя современные методы представления и обработки больших массивов графических панорамных изображений.

Предлагается автоматизировать процесс создания панорамных изображений с использованием специального программного обеспечения (СПО), предназначенного для оценки влияния процедуры коррекции яркости и цвета на различные методы детектирования ключевых точек с фиксацией результатов.

Ранее авторами в рамках научных исследований предлагалась информационная модель технологии коррекции яркости и цвета при создании панорамных изображений [1]. Для уточнения существующей модели предлагается дополнительно провести оценку влияния процедуры коррекции яркости и цвета на особенности детектирования ключевых точек при создании панорамных изображений.

Результатом исследования должны стать статистические данные, полученные в ходе сшивания панорам и позволяющие провести их сравнительный анализ, наглядно демонстрирующий влияние модели на качество итоговых изображений.

## 2 Методы детектирования ключевых точек на изображениях

Процедура сшивания изображений в панораму реализуется на основе работы с наборами ключевых точек сшиваемых изображений. Вычисление ключевых точек наиболее часто выполняется путем применения общеизвестных детекторов ключевых точек и их дескрипторов, таких как SIFT, SURF, FAST, BRIEF, BRISK, AKAZE, MSER, ORB и др. [2–13]. Однако часть названных методов являются конвергенцией, модификацией и объединением других методов либо они защищены авторским правом, что может вызывать затруднения в их использовании на практике. Поэтому предлагается исследование популярных методов, зарекомендовавших себя наилучшим образом и которые могут быть использованы без ограничений. Также одним из решающих факторов является реализация метода в библиотеке компьютерного зрения OpenCV, так как это значительно упрощает проведение исследования.

Согласно вышеизложенным условиям отбора наиболее подходящими методами в рамках исследования оказались SIFT, BRISK и ORB.

## 3 Проведение экспериментальных исследований

Для проведения исследования использовались наборы аэрофотоснимков, полученных с камеры БПЛА DJI Mavic 2 при высоте полета 150 и 250 м в количестве семи и четырех изображений соответственно для выбранных участков территории. Каждый из снимков имеет разрешение  $1920 \times 1080$  пикселей. Также было задействовано СПО [14], предназначенное для создания панорамных изображений с возможностью использования различных методов детектирова-

Перечисление шагов алгоритма программы экспериментальных исследований

| № сшивания | Метод детектирования | Применение коррекции яркости и цвета | Обрезание краев изображений |
|------------|----------------------|--------------------------------------|-----------------------------|
| 1          | SIFT                 | Нет                                  | 0%                          |
| 2          |                      | Да                                   |                             |
| 3          |                      | Нет                                  | 10%                         |
| 4          |                      | Да                                   |                             |
| 5          |                      | Нет                                  | 20%                         |
| 6          |                      | Да                                   |                             |
| 7          | ORB                  | Нет                                  | 0%                          |
| 8          |                      | Да                                   |                             |
| 9          |                      | Нет                                  | 10%                         |
| 10         |                      | Да                                   |                             |
| 11         |                      | Нет                                  | 20%                         |
| 12         |                      | Да                                   |                             |
| 13         | BRISK                | Нет                                  | 0%                          |
| 14         |                      | Да                                   |                             |
| 15         |                      | Нет                                  | 10%                         |
| 16         |                      | Да                                   |                             |
| 17         |                      | Нет                                  | 20%                         |
| 18         |                      | Да                                   |                             |

ния ключевых точек. Специальное программное обеспечение представляет собой набор программных модулей, разработанных на языке программирования Python 3.6 и функционирующее на компьютере со следующей конфигурацией: процессор AMD FX-8320, оперативная память 8 ГБ с частотой 1600 МГц и 64-разрядная операционная система Windows 10.

Программа экспериментальных исследований включает в себя серию сшиваний панорамных изображений в различных режимах для каждого набора аэрофотоснимков, представленных в таблице.

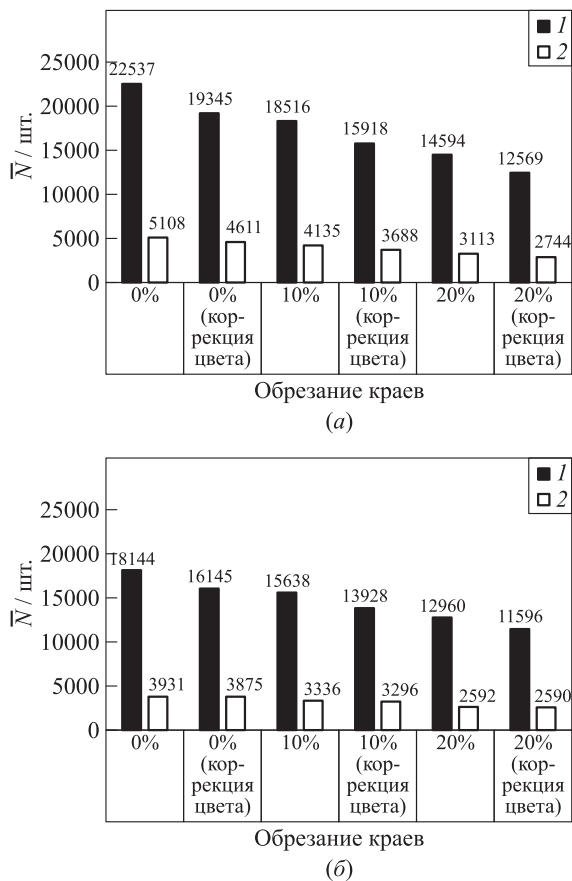
Обрезание краев изображений приводит к исключению искажений, вызванных дисторсией оптической системы камеры БПЛА при разновысотной съемке [15].

Для оценки влияния процедуры коррекции яркости и цвета на методы детектирования точек были выбраны следующие показатели:

- общее число ключевых точек  $N_{\text{общ}}$ ;
- число совпавших ключевых точек между сшиваемыми изображениями  $N_{\text{совп}}$ ;
- время, затраченное на сшивание панорамного изображения  $T_{\text{сш}}$ ;
- точность сшивания  $Q_{\text{сш}}$  изображений на итоговой панораме, выраженной в процентном отношении числа корректно сформированных объектов на панораме к общему числу сшиваемых объектов.

#### 4 Результаты экспериментальных исследований

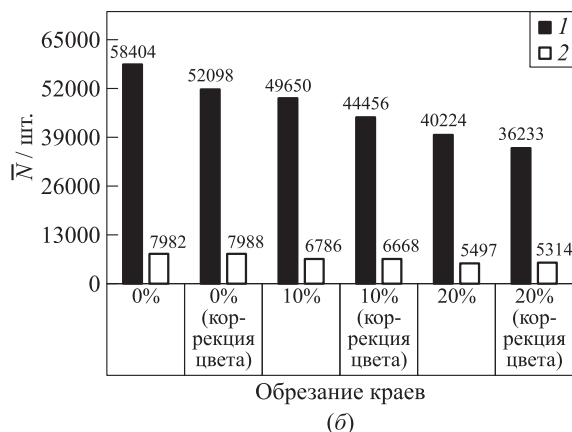
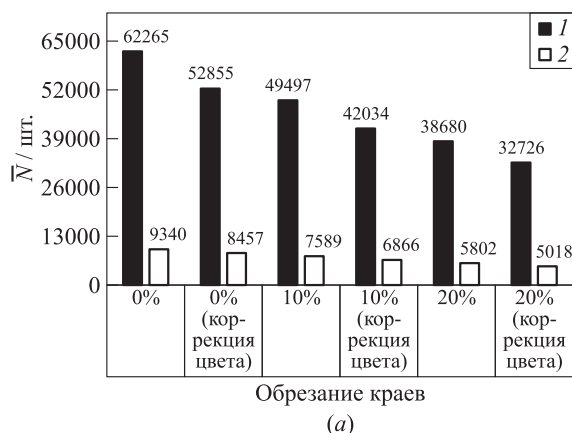
**Метод SIFT.** Первым методом в соответствии с таблицей для исследования был выбран метод SIFT, в ходе проведения испытаний которого были получены данные, приведенные на рис. 1. Общее число ключевых точек для каждого из испытаний метода SIFT рассчитывалось как усредненное значение  $\bar{N}_{\text{общ}}$ . Аналогичный расчет проводился и с числом совпавших ключевых точек, однако здесь вычислялось среднее значение совпавших точек между всеми парами сшиваемых снимков  $\bar{N}_{\text{совп}}$ . Исходя из информации, представленной на графиках, прослеживается закономерность влияния процедуры коррекции яркости и цвета



**Рис. 1** Число найденных ключевых точек методом SIFT при высоте съемки 150 (а) и 250 м (б): 1 —  $\bar{N}_{\text{общ}}$ ; 2 —  $\bar{N}_{\text{совп}}$

на результат работы метода SIFT. Применение коррекции значительно снижает общее число ключевых точек, найденных на фотоснимках, но число совпавших точек между парами сшиваемых изображений остается практически неизменным. Также закономерным является тот факт, что обрезание краев изображений влияет как на исходное число ключевых точек, так и на число совпавших точек.

**Метод ORB.** Следующий реализованный метод — ORB. Его испытания при тех же условиях представлены на рис. 2. Заметное отличие, выделяющее его на фоне других методов, — это большее чем в 3 раза число найденных ключевых точек. Аналогично предыдущему методу процедура коррекции яркости и цвета заметно уменьшает общее число ключевых точек и незначительно

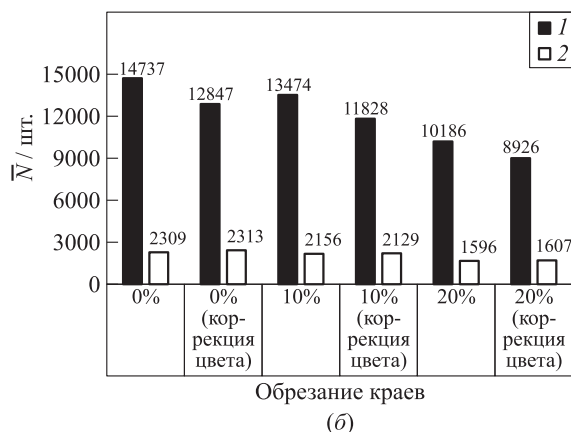
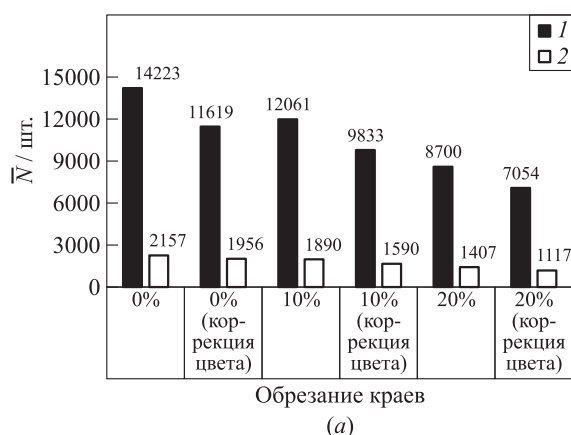


**Рис. 2** Число найденных ключевых точек методом ORB при высоте съемки 150 (а) и 250 м (б): 1 —  $\bar{N}_{\text{общ}}$ ; 2 —  $\bar{N}_{\text{совп}}$

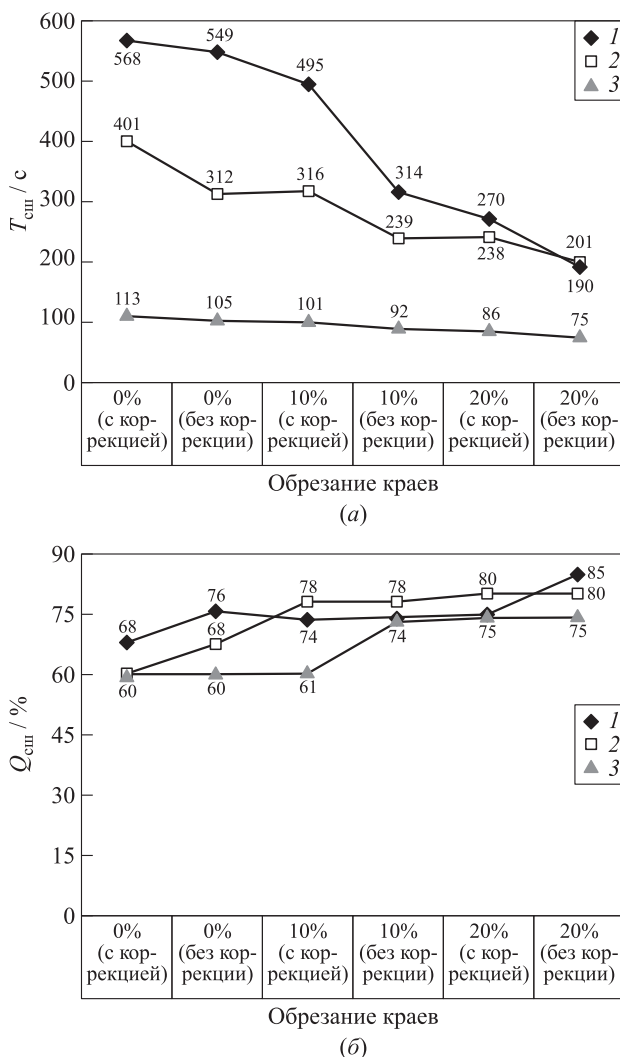
число совпавших точек. В случае обрезания краев изображений результаты также идентичны методу SIFT.

**Метод BRISK.** Испытания метода BRISK при тех же исходных данных показали результаты, приведенные на рис. 3. По результатам испытания метода BRISK тенденция, прослеживаемая в предыдущих испытаниях, не изменилась.

Далее рассмотрим влияние изменения числа ключевых точек из-за воздействия коррекции и обрезания на скорость создания и точность сшивания панорамного изображения. В случае сшивания снимков, полученных с высоты 150 м, заметно увеличение скорости получения панорам для методов SIFT и ORB,



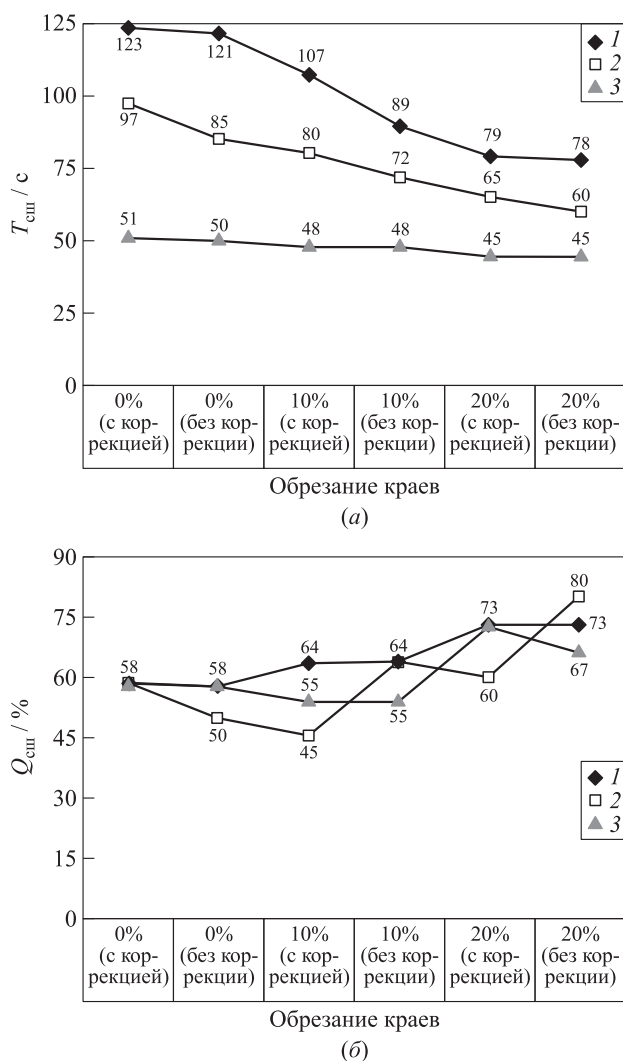
**Рис. 3** Число найденных ключевых точек методом BRISK при высоте съемки 150 (a) и 250 м (б): 1 —  $\bar{N}_{\text{общ}}$ ; 2 —  $\bar{N}_{\text{совп}}$



**Рис. 4** Время  $T_{\text{сш}}$  (а) и точность  $Q_{\text{сш}}$  (б) сшивания панорамных изображений методами SIFT (1), ORB (2) и BRISK (3) для высоты 150 м

при этом точность сшивания также увеличивается, но уже для всех методов. Описанные результаты представлены на рис. 4.

Результаты исследования скорости и точности сшивания для высоты 250 м показаны на рис. 5. Графики аналогичны зависимостям, представленным на рис. 4, но в случае метода ORB заметно небольшое ухудшение в точности



**Рис. 5** Время  $T_{\text{сш}}$  (а) и точность  $Q_{\text{сш}}$  (б) сшивания панорамных изображений методами SIFT (1), ORB (2) и BRISK (3) для высоты 250 м

сшитых панорам. Причина такого результата может быть вызвана большой высотой съемки, так как детализация объектов становится ниже, что влияет на число ключевых точек. Ранее отмечалось, что коррекция незначительно уменьшает долю совпавших ключевых точек от общего числа точек, поэтому это могло повлиять на итоговую панораму. Вдобавок к этому можно заметить общее

ухудшение в точности сшивания панорамы для высоты 250 м. Вызвано это также уменьшением детализации из-за увеличения высоты съемки.

## 5 Заключение

В результате проведенного исследования был сделан вывод о том, что для создания панорамных изображений важными показателями являются: общее число найденных ключевых точек  $N_{\text{общ}}$ , число совпавших ключевых точек  $N_{\text{совп}}$ , точность сшивания изображений  $Q_{\text{сш}}$  и время  $T_{\text{сш}}$ , затраченное на сшивание панорамного изображения. Благодаря применению процедуры коррекции яркости и цвета уменьшилось общее число ключевых точек для каждого из методов детектирования SIFT, ORB и BRISK, но при этом число совпавших точек изменилось незначительно. Примеры полученных панорамных изображений для разных высот показаны на рис. 6.

Доля совпавших ключевых точек от общего числа найденных увеличилась в среднем на 18% для указанных методов, при этом время на сшивание панорам уменьшилось в среднем на 164 с за счет обрезания краев изображений на 20%.

Дополнительно было отмечено, что увеличение высоты съемки оказывает отрицательное воздействие на точность сшивания итоговой панорамы, вызванное



**Рис. 6** Примеры панорамных изображений при высоте съемки 150 (а) и 250 м (б)

уменьшением числа совпадающих ключевых точек между парами сшиваемых снимков. Для использованных методов SIFT, ORB и BRISK оно составило 23%, 14,9% и 17,9% соответственно.

## Литература

1. *Архипов П. О., Цуканов М. В.* Информационная модель технологии коррекции яркости и цвета при создании панорамных изображений // Системы высокой доступности, 2020. Т. 16. № 3. С. 46–51. doi: 10.18127/j20729472-202003-04.
2. *Harris C., Stephens M.* A combined corner and edge detector // 4th Alvey Vision Conference Proceedings. — Manchester, U.K.: Alvey Vision Club, 1988. P. 147–151.
3. *Rublee E., Rabaud V., Konolige K., Bradski G.* ORB: An efficient alternative to SIFT or SURF // Conference (International) on Computer Vision. — Piscataway, NJ, USA: IEEE, 2011. P. 2564–2571. doi: 10.1109/ICCV.2011.6126544.
4. *Гончаренко М. О.* Сравнительный анализ методов формирования дескрипторов изображений в контексте задачи сегментации видеопотока // Бионика интеллекта, 2015. № 2(85). С. 90–94.
5. *Супрун Д. Е.* Алгоритм сопоставления изображений по ключевым точкам при масштабировании и вращении объектов // Вестник МГТУ им. Н. Э. Баумана. Сер. Приборостроение, 2016. № 5. С. 86–98. doi: 10.18698/0236-3933-2016-5-86-98.
6. *Прозоров А. В.* Детекторы особых точек в задаче одновременной локализации и картирования // DSPA: вопросы применения цифровой обработки сигналов, 2016. Т. 6. № 4. С. 804–808.
7. *Залесский Б. А.* Детекторы экстремальных особых точек на изображениях // Докл. НАН Беларуси, 2017. Т. 61. № 5. С. 37–41.
8. *Макаров А. С., Болсуновская М. В.* Сравнительный анализ методов обнаружения особых точек на изображениях при различных уровнях освещения // Научно-технические ведомости Санкт-Петербургского гос. политехнического ун-та. Информатика. Телекоммуникации. Управление, 2018. Т. 11. № 2. С. 7–18. doi: 10.18721/JCSTCS.11201.
9. *Нечаев А. П.* Исследование алгоритмов детектирования особых точек для решения задачи сравнения изображений // Перспективные информационные технологии: Труды Междунар. научн.-технич. конф. — Самара: Самарский научный центр РАН, 2018. С. 506–508.
10. *Сафронова О. Р., Трубаков Е. О.* Анализ методов детектирования ключевых точек изображения для визуальной одометрии // Информационные технологии. Проблемы и решения, 2019. № 3(8). С. 125–130.
11. *Бондаренко В. А., Каплинский Г. Э., Павлова В. А., Тушиков В. А.* Метод поиска и сопоставления ключевых особенностей изображений для распознавания образов и сопровождения объектов // Известия ЮФУ. Технические науки, 2019. № 1(203). С. 281–293. doi: 10.23683/2311-3103-2019-1-281-293.
12. *Краснобаев Е. А., Чистобаев Д. В., Мальшев А. Л.* Сравнение бинарных дескрипторов особых точек изображений в условиях искажений // Компьютерная оптика, 2019. Т. 43. № 3. С. 434–445. doi: 10.18287/2412-6179-2019-43-3-434-445.
13. *Трубаков Е. О., Трубаков А. О., Трубакова О. Р.* Исследование результатов работы алгоритмов сопоставления ключевых точек для анализа космоснимков //

Информационные системы и технологии: Сб. мат-лов XXVI Междунар. научн.-технич. конф. — Нижний Новгород: Нижегородский ГТУ им. Р.Е. Алексеева, 2020. С. 1066–1073.

14. *Архипов О. П., Архипов П. О., Цуканов М. В.* Программная система автоматизированного создания и сравнения фотопланов с функцией диагностирования аномалий (С-19). Свидетельство о государственной регистрации программы для ЭВМ № 2019663720.
15. *Архипов П. О., Цуканов М. В.* Модификация алгоритма минимизации искажений при сшивании аэрофотоснимков в панораму // Системы высокой доступности, 2020. Т. 16. № 3. С. 52–57. doi: 10.18127/j20729472-202003-05.

*Поступила в редакцию 27.08.21*

---

## STUDY OF METHODS FOR DETECTING KEY POINTS WHEN CREATING PANORAMIC IMAGES

*P. O. Arkhipov, A. K. Trofimenkov, M. V. Tsukanov, and N. Yu. Nosova*

Orel Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 137 Moskovskoe Shosse, Orel 302025, Russian Federation

**Abstract:** The article is devoted to the problem of creating panoramas in conditions of differences in the color palette and illumination level of stitched images caused both by the features of the optical system and by the weather and other factors. A procedure for correcting brightness and color is proposed to minimize differences between stitched images and to reduce negative impacts. The features of detection of key points by various methods during the stitching of panoramic images under conditions of shooting at different heights are presented. The effect of an additional procedure for correcting brightness and color on the number of found key points is described. The quality of the final panoramic images and the time of their stitching were evaluated. In the course of the work, statistical data were obtained and their comparative analysis was carried out.

**Keywords:** brightness correction; color correction; detection; key points; panoramic images; image processing

**DOI:** 10.14357/08696527220209

## References

1. Arkhipov, P. O., and M. V. Tsukanov. 2020. Informatsionnaya model' tekhnologii korrektsii yarkosti i tsвета pri sozdanii panoramnykh izobrazheniy [Information model of brightness and color correction technology for creating panoramic images]. *Sistemy vysokoy dostupnosti* [Highly Available Systems] 16(3):46–51. doi: 10.18127/j20729472-202003-04.

2. Harris, C., and M. Stephens. 1988. A combined corner and edge detector. *4th Alvey Vision Conference Proceedings*. Manchester, U.K.: Alvey Vision Club. 147–151.
3. Rublee, E., V. Rabaud, K. Konolige, and G. Bradski. 2011. ORB: An efficient alternative to SIFT or SURF. *Conference (International) on Computer Vision Proceedings*. Piscataway, NJ: IEEE. 2564–2571. doi: 10.1109/ICCV.2011.6126544.
4. Goncharenko, M. O. 2015. Sravnitel'nyy analiz metodov formirovaniya deskriptorov izobrazheniy v kontekste zadachi segmentatsii videopotoka [Comparative analysis of key points detection methods as a part of video segmentation problem]. *Bionika intellekta* [Bionics of Intelligence] 2(85):90–94.
5. Suprun, D. E. 2016. Algoritm sopostavleniya izobrazheniy po klyuchevym tochkam pri masshtabiruemosti i vrashchenii ob"ektov [Image-matching algorithm using key points with scalability and rotation of objects]. *Vestnik MGTU im. N. E. Baumana. Ser. Priborostroenie* [Herald of the Bauman Moscow State Technical University. Ser. Instrument engineering] 5:86–98. doi: 10.18698/0236-3933-2016-5-86-98.
6. Prozorov, A. V. 2016. Detektory osobykh toчек v zadache odnovremennoy lokalizatsii i kartirovaniya [Special point detectors in the task of simultaneous localization and mapping]. *DSPA: Voprosy primeneniya tsifrovoy obrabotki signalov* [DSPA: Digital Signal Processing Applications] 6(4):804–808.
7. Zalesky, B. A. 2017. Detektory ekstremal'nykh osobykh toчек na izobrazheniyakh [Detectors of extremal key points on images]. *Dokl. NAN Belarusi* [Doklady of the National Academy of Sciences of Belarus] 61(5):37–41.
8. Makarov, A. S., and M. V. Bolsunovskaya. 2018. Sravnitel'nyy analiz metodov obnaruzheniya osobykh toчек na izobrazheniyakh pri razlichnykh urovnyakh osveshcheniya [Comparative analysis of methods for detecting special points in images at different levels of illumination]. *Nauchno-tehnicheskie vedomosti SPbGPU. Informatika. Telekommunikatsii. Upravlenie* [St. Petersburg State Polytechnical University J. Computer Science. Telecommunications and Control Systems] 11(2):7–18. doi: 10.18721/JCSTCS.11201.
9. Nechaev, A. P. 2018. Issledovanie algoritmov detektirovaniya osobykh toчек dlya resheniya zadachi sravneniya izobrazheniy [Investigation of algorithms for detecting singular points for solving the problem of comparing images]. *Scientific and Technical Conference (International) "Advanced Information Technologies" Proceedings*. Samara: Samara Scientific Center RAS. 506–508.
10. Safronova, O. R., and E. O. Trubakov. 2019. Analiz metodov detektirovaniya klyuchevykh toчек izobrazheniya dlya vizual'noy odometrii [Analysis of methods for detecting the key image points for the visual odometry]. *Informatsionnye tekhnologii. Problemy i resheniya* [Information Technology] 3(8):125–130.
11. Bondarenko, V. A., G. E. Kaplinskiy, V. A. Pavlova, and V. A. Tupikov. 2019. Metod poiska i sopostavleniya klyuchevykh osobennostey izobrazheniy dlya raspoznavaniya obrazov i soprovozhdeniya ob"ektov [A method for detection and matching of image keypoints for object recognition and tracking]. *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFEDU. Engineering Sciences] 1(203):281–293. doi: 10.23683/2311-3103-2019-1-281-293.
12. Krasnobaev, E. A., D. V. Chistobaev, and A. L. Malyshev. 2019. Sravnenie binarnykh deskriptorov osobykh toчек izobrazheniy v usloviyakh iskazheniy [Comparison of binary feature points descriptors of images under distortion conditions]. *Komp'yuternaya*

- optika* [Computer Optics] 43(3):434–445. doi: 10.18287/2412-6179-2019-43-3-434-445.
13. Trubakov, E. O., A. O. Trubakov, and O. R. Trubakova. 2020. Issledovanie rezul'tatov raboty algoritmov sopostavleniya klyuchevykh toчек dlya analiza kosmosnimkov [Investigation of the results of the operation of algorithms for matching key points for the analysis of space images]. *26th Scientific and Technical Conference (International) "Information Systems and Technologies" Proceedings*. Nizhniy Novgorod: Nizhegorodskiy GTU im. R. E. Alekseeva. 1066–1073.
  14. Arkhipov, O. P., P. O. Arkhipov, and M. V. Tsukanov. 2019. Programmnaya sistema avtomatizirovannogo sozdaniya i sravneniya fotoplanov s funktsiey diagnostirovaniya anomaliiy (C\_19) [Software system for automated creation and comparison of photographic plans with anomaly diagnostics function (C\_19)]. Certificate of State Registration No. 2019663720.
  15. Arkhipov, P. O., and M. V. Tsukanov. 2020. Modifikatsiya algoritma minimizatsii iskazheniy pri sshivanii aerofotosnimkov v panoramu [Modification of the algorithm for minimizing distortion when stitching aerial photographs into a panorama]. *Sistemy vysokoy dostupnosti* [Highly Available Systems] 16(3):52–57. doi: 10.18127/j20729472-202003-05.

*Received August 27, 2021*

## **Contributors**

**Arkhipov Pavel O.** (b. 1979) — Candidate of Science (PhD) in technology, director, Orel Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 137 Moskovskoe Shosse, Orel 302025, Russian Federation; arpaul@mail.ru

**Trofimenkov Andrei K.** (b. 1965) — Candidate of Science (PhD) in technology, associate professor, senior scientist, Orel Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 137 Moskovskoe Shosse, Orel 302025, Russian Federation; andretrf@yandex.ru

**Tsukanov Maxim V.** (b. 1995) — junior scientist, Orel Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 137 Moskovskoe Shosse, Orel 302025, Russian Federation; tsukanov.m.v@yandex.ru

**Nosova Natalia Yu.** (b. 1964) — leading engineer, Orel Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 137 Moskovskoe Shosse, Orel 302025, Russian Federation; ofipiran.lab01@yandex.ru

## МОДЕЛЬ ОПРЕДЕЛЕНИЯ НЕОБХОДИМОГО ЧИСЛА ОПТИЧЕСКИХ ДИСКОВ, ОРГАНИЗОВАННЫХ В ГИБРИДНЫЕ СТРУКТУРЫ, ДЛЯ СОЗДАНИЯ И ПОДДЕРЖАНИЯ ДОЛГОВРЕМЕННЫХ ЭЛЕКТРОННЫХ АРХИВОВ

*А. В. Чернышов<sup>1</sup>*

**Аннотация:** Гибридные структуры оптических дисков представляют собой метод архивного хранения информации с повышенной надежностью, аналогичный RAID-массивам дисков. Но в отличие от классических массивов RAID в гибридных структурах одна из копий информации хранится на одинокых дисках, что обеспечивает быстрый и простой доступ к информации с помощью всего одного стандартного привода. При длительном хранении оптические диски с записью способны деградировать. Знание числа оптических дисков, которое необходимо ежегодно приобретать для создания архива, а затем для поддержания его целостности, необходимо для планирования жизненного цикла архива и затрат на его содержание еще на этапе проектирования. В работе построены аналитические модели, позволяющие определить необходимое ежегодно число оптических дисков для четырех возможных вариантов (стратегий) заполнения архива, а также моменты времени, характеризующие процесс заполнения архива и позволяющие оценить принципиальную достижимость плановой емкости архива. Приведен пример реального расчета.

**Ключевые слова:** долговременное электронное хранилище информации; оптические диски однократной записи (WORM); гибридные структуры оптических дисков; замена отказавших оптических дисков; ежегодно необходимое число оптических дисков

**DOI:** 10.14357/08696527220210

### 1 Введение

Современные электронные документы, пройдя естественный цикл своего существования, должны в конечном итоге быть сохранены в архиве. Для любых носителей информации существует проблема порчи записанной на них информации. Поэтому в процессе хранения необходим регулярный контроль целостности носителей и записанной на них информации [1, 2].

Согласно современным стандартам в качестве архивных носителей информации должны использоваться устройства типа WORM (write once, read many) [3], что автоматически означает использование только оптических дисков однократной записи (типа CD-R, DVD-R или BD-R). Поэтому далее будем рассматривать только архивы, построенные на таких носителях.

---

<sup>1</sup>Мытищинский филиал МГТУ им. Н. Э. Баумана, sch-ru@yandex.ru

Оптические диски с записью с течением времени способны деградировать [4, 5], поэтому в архиве они могут быть продублированы [6, 7] либо объединены в структуры типа RAID (с использованием роботизированных библиотек [8]). Но наиболее перспективным представляется их организация в гибридные структуры [9].

В гибридной структуре весь записываемый объем информации делится на блоки, каждый из которых включает  $k$  одиночных дисков и  $L(n - 1)$  массивов RAID, где  $n$  — число запасных копий информации. Каждый массив включает  $k^* = k + m$  дисков, где  $k$  — условное число дисков с информацией;  $m$  — условное число дисков с контрольными суммами (например, для RAID6  $k > 1$ ,  $m = 2$ ), причем  $(n - 1)$  запасных копий массивов RAID при  $n > 2$  фактически вместо массивов типа RAID5 или RAID6 образуют массивы типа RAID15 или RAID16. Требуется определить число оптических дисков, приобретаемых ежегодно для заполнения и поддержания архива, а также время (в годах), в течение которого при заданной интенсивности поступления информации архив будет заполнен до своего планового объема.

Опубликовано немало статей, посвященных дисковым массивам типа RAID различных видов. Но практически во всех этих работах рассматриваются массивы RAID, составленные из жестких дисков, где рассматриваемые структуры не применяются. В связи с этим необходимо построение новой модели.

## 2 Построение модели

Рассмотрим 4 варианта (стратегии) заполнения архива.

1. Архив заполняется в первый год существования. Все последующие годы требуется только выявлять отказавшие диски и заменять их новыми.
2. Аппаратура архива не накладывает ограничений на интенсивность поступления информации, но сама интенсивность невелика, что приводит к заполнению архива в течение нескольких лет.
3. Пропускная способность аппаратуры архива ограничена, что приводит к вынужденному заполнению архива в течение нескольких лет, причем из-за необходимости тратить пропускную способность аппаратуры на контроль целостности ранее записанных дисков пропускная способность аппаратуры для записи новой информации постепенно снижается.
4. Комбинация вариантов 2 и 3.

Необходимо построить математические модели для каждого варианта.

Будем использовать следующие обозначения:  $W$  — максимальная (плановая) емкость архива;  $t$  — время существования архива, год;  $T$  — время, выделяемое в течение года на технические работы в архиве;  $V$  — емкость одного оптического диска;  $q$  — вероятность отказа оптического диска с записью при хранении в течение года;  $q_w$  — вероятность отказа оптического диска в момент записи

(брак);  $n$  — число запасных копий каждого блока информации (при этом одна запасная копия находится на одиночном диске и  $(n - 1)$  копия — на массивах RAID);  $l = \lceil W/V \rceil$  — число одиночных дисков емкости  $V$ , на которые может быть записан объем информации  $W$ ;  $L = \lceil l/k \rceil$  — число массивов RAID с параметрами  $k$  и  $m$ , на которые может быть записан объем информации  $W$ ;  $k^* = k + m$  — общее число дисков в одном RAID;  $l^* = Lk^*$  — общее число дисков в  $L$  массивах RAID;  $N(t)$  — число оптических дисков, которые необходимо приобрести в год с номером  $t$  для создания или поддержания архива.

**Вариант 1.** В первый год диски расходуются на заполнение архива:

$$N(1) = \left\lceil \frac{Lk^*(n-1) + l}{1 - q_w} \right\rceil = \left\lceil \frac{l^*(n-1) + l}{1 - q_w} \right\rceil.$$

В любой последующий год требуются только диски для замены вышедших из строя:

$$N(t) = \left\lceil \frac{\lceil ((n-1)l^* + l)q \rceil}{1 - q_w} \right\rceil. \quad (1)$$

Здесь важно, что  $l$  необязательно кратно  $l^*$ , поскольку в последнем блоке может быть записано меньше чем  $k$  дисков, но массивы RAID все равно должны быть собраны в полном объеме.

**Вариант 2.** Введем дополнительные обозначения. Для одиночных дисков:  $t_w$  — время, необходимое для записи с верификацией одного оптического диска;  $t_{r0}$  — время, необходимое для планового контроля одного диска;  $t_r$  — время, необходимое для считывания копии одного диска. Для дисков, входящих в массивы RAID, будем использовать обозначения  $t_{wL}$ ,  $t_{r0L}$  и  $t_{rL}$ , учитывающие возможность увеличения скорости обработки за счет одновременной работы с дисками массива (если это позволяет аппаратура).

В год будет записываться  $W_y = \text{const}$  информации; для записи этой информации потребуется  $l_y = \lceil W_y/V \rceil$  одиночных дисков либо  $L_y = \lceil l_y/k \rceil$  массивов RAID (без учета запасных копий). Тогда можно записать:

$$\begin{aligned} T \geq & \left\lceil \frac{l_y}{1 - q_w} \right\rceil t_w + \left\lceil \frac{(n-1)L_y k^*}{1 - q_w} \right\rceil t_{wL} + (t-1)l_y t_{r0} + \\ & + (t-1)(n-1)L_y k^* t_{r0L} + \lceil (t-1)l_y q \rceil t_r + \lceil (t-1)(n-1)L_y k^* q \rceil t_{rL} + \\ & + \left\lceil \frac{\lceil (t-1)l_y q \rceil}{1 - q_w} \right\rceil t_w + \left\lceil \frac{\lceil (t-1)(n-1)L_y k^* q \rceil}{1 - q_w} \right\rceil t_{wL}. \quad (2) \end{aligned}$$

Чаще всего для гибридных структур используется всего один привод. При этом  $t_r = t_{rL}$ ,  $t_w = t_{wL}$ ,  $t_{r0} = t_{r0L}$  и выражение сильно упрощается. Однако рассмотрим ситуацию в общем случае.

Оценим максимально возможное число дисков  $l_y$ , которые могут быть записаны в год и определяют максимально допустимое значение  $W_y$ . В действительности функция  $L_y(l_y)$  принимает ступенчатые значения, поэтому аналитически для значения  $l_y$  можно лишь получить две граничные оценки, для чего рассмотрим оценки для  $L_y = l_y/k$  и для  $L_y = l_y/k + 1$ .

Действуя по методике [7], получим:

$$\left\lfloor \frac{T - K^*(A_L + (t-1)B_L)}{A_G + (t-1)B_G} \right\rfloor \leq l_y \leq \left\lfloor \frac{T}{A_G + (t-1)B_G} \right\rfloor.$$

Здесь

$$A_L = \frac{t_w L}{1 - q_w}; \quad B_L = t_{r0L} + qt_{rL} + \frac{qt_w L}{1 - q_w}; \\ K^* = (n-1)k^*; \quad A_G = A + K A_L; \quad B_G = B + K B_L,$$

где

$$A = \frac{t_w}{1 - q_w}; \quad B = t_{r0} + qt_r + \frac{qt_w}{1 - q_w}; \quad K = (n-1)\frac{k^*}{k}.$$

Из-за ограничений объема статьи далее везде будем выполнять оценку только для максимального значения  $l_y$  (наиболее быстрое заполнение).

Год заполнения архива  $t_{\text{full}}$  вычислим, исходя из соотношения  $l_y t_{\text{full}} \geq l$ :

$$t_{\text{full}} = \left\lceil \frac{l(A + K A_L - (B + K B_L))}{T - l(B + K B_L)} \right\rceil = \left\lceil \frac{l(A_G - B_G)}{T - l B_G} \right\rceil.$$

В произвольный год  $t$  ( $t < t_{\text{full}}$ ) необходимо будет приобрести для записи

$$N(t) = \left\lceil \frac{l_y + (n-1)L_y k^*}{1 - q_w} \right\rceil + \left\lceil \frac{[(t-1)(l_y + (n-1)L_y k^*)q]}{1 - q_w} \right\rceil = \\ = \left\lceil \frac{l_y + (n-1)l_y^*}{1 - q_w} \right\rceil + \left\lceil \frac{[(t-1)(l_y + (n-1)l_y^*)q]}{1 - q_w} \right\rceil \quad (3)$$

оптических дисков. Здесь  $l_y^* = L_y k^*$ .

Модель варианта 2 построена.

**Вариант 3.** В отличие от варианта 2 теперь  $l_y = l_f(t) \neq \text{const}$ ,  $L_y = L_f(t) = \lceil l_f(t)/k \rceil$ . При этом в рамках RAID будут записываться  $l_f^*(t) = L_f(t)k^*$  дисков.

В первый год ( $t = 1$ ) можно записать максимальное число дисков:

$$\left\lfloor \frac{T - K^* A_L}{A_G} \right\rfloor \leq l_f(1) \leq \left\lfloor \frac{T}{A_G} \right\rfloor.$$

Для любого последующего года:

$$\begin{aligned}
 T \geq & \left\lceil \frac{l_f(t)}{1 - q_w} \right\rceil t_w + t_{r0} \sum_{i=1}^{t-1} (l_f(i)) + \left\lceil q \sum_{i=1}^{t-1} (l_f(i)) \right\rceil t_r + \left\lceil \frac{q \sum_{i=1}^{t-1} (l_f(i))}{1 - q_w} \right\rceil t_w + \\
 & + \left\lceil \frac{(n-1) \lceil l_f(t)/k \rceil k^*}{1 - q_w} \right\rceil t_{wL} + (n-1) t_{r0L} \sum_{i=1}^{t-1} \left( \left\lceil \frac{l_f(i)}{k} \right\rceil k^* \right) + \\
 & + \left\lceil (n-1) q \sum_{i=1}^{t-1} \left( \left\lceil \frac{l_f(i)}{k} \right\rceil k^* \right) \right\rceil t_{rL} + \left\lceil \frac{(n-1) q \sum_{i=1}^{t-1} (\lceil l_f(i)/k \rceil k^*)}{1 - q_w} \right\rceil t_{wL}. \quad (4)
 \end{aligned}$$

Откуда по той же методике получим:

$$\left\lfloor \frac{T - K^*(A_L + (t-1)B_L)}{A_G} E^{t-1} \right\rfloor \leq l_f(t) \leq \left\lfloor \frac{T}{A_G} E^{t-1} \right\rfloor, \quad E = 1 - \frac{B_G}{A_G}; \quad (5)$$

$$\frac{T - K^*(A_L + (t-1)B_L)}{B_G} (1 - E^{t-1}) \leq \sum_{i=1}^{t-1} (l_f(i)) \leq \frac{T}{B_G} (1 - E^{t-1}). \quad (6)$$

С учетом восстанавливаемых дисков в год с номером  $t$  необходимо будет записывать

$$\begin{aligned}
 N(t) = & \left\lceil \frac{l_f(t) + (n-1)L_f(t)k^*}{1 - q_w} \right\rceil + \\
 & + \left\lceil \frac{q \left( \sum_{i=1}^{t-1} l_f(i) + (n-1) \sum_{i=1}^{t-1} (L_f(i)k^*) \right)}{1 - q_w} \right\rceil.
 \end{aligned}$$

Поскольку  $L_f(i) = \lceil l_f(i)/k \rceil$ , раскрывая выражение под знаками сумм, проведем оценку  $N(t)$  по верхнему краю, т. е.  $L_f(i) = (l_f(i)/k) + 1$ , а при раскрытии значения  $L_f(t)$  будем использовать максимальную оценку значения  $l_f(t)$ :

$$\begin{aligned}
 N(t) = & \left\lceil \frac{1}{1 - q_w} \left( \left\lfloor \frac{T}{A_G} E^{t-1} \right\rfloor + \left\lfloor \frac{(T/A_G)E^{t-1}}{k} \right\rfloor K^* \right) \right\rceil + \\
 & + \left\lceil \frac{1}{1 - q_w} \left[ q \left( (K+1) \frac{T}{B_G} (1 - E^{t-1}) + (t-1)K^* \right) \right] \right\rceil.
 \end{aligned}$$

Год  $t_{\text{full}}$  определим, исходя из соотношения  $l \leq \sum_{i=1}^{t_{\text{full}}} l_f(i)$ .

В итоге для правой части выражения (6) (оценка нижней границы  $t_{\text{full}}$  при максимальной скорости заполнения) получим  $t_{\text{full}} = \lceil \log_E(1 - lB_G/T) \rceil$ .

При  $t > t_{\text{full}}$  для поддержания архива необходимо будет ежегодно приобретать  $N(t)$  дисков, где  $N(t)$  определяется выражением (1).

Определим также номер года  $t_f$ , начиная с которого пропускная способность аппаратуры архива не позволит помещать в него новую информацию. Оценим его из соотношения  $l_f(t_f) < 1$ .

Получим для правой части (5) (нижняя граница оценки  $t_f$ ):

$$t_f = \left\lceil \log_E \left( \frac{A_G}{T} \right) \right\rceil.$$

Для достижения плановой емкости архива  $W$  должно выполняться  $t_{\text{full}} < t_f$ .

**Вариант 4.** До года  $\tilde{t}$  архив будет заполняться по методу варианта 2. Далее до полного заполнения — по методу варианта 3.

Параметр  $\tilde{t}$  оценим, исходя из выражения (2). Получим:

$$\tilde{t} = \left\lfloor \frac{T - (A_G l_y + K^* A_L)}{B_G l_y + K^* B_L} \right\rfloor + 1.$$

В течение этого периода времени в год потребуются записывать  $N(t)$  дисков, где  $N(t)$  определяется выражением (3), а за весь период в архив будет помещено  $\tilde{t}(l_y + (n-1)l_y^*)$  дисков.

Дальнейшее заполнение архива начнется с года  $t = \tilde{t} + 1$ .

Введем новую шкалу времени  $\bar{t} = t - \tilde{t}$ . Подставив  $\bar{t}$  в выражение (4) и выполнив преобразования, получим:

$$l_f(\bar{t}) = \left\lfloor D_G^* \left( 1 - \frac{B_G}{A_G} \right)^{\bar{t}-1} \right\rfloor = \left\lfloor D_G^* E^{\bar{t}-1} \right\rfloor; \quad \sum_{i=1}^{\bar{t}-1} l_f(i) = \frac{D_G^* A_G}{B_G} (1 - E^{\bar{t}-1}),$$

где

$$D_G^* = \frac{T}{A_G} - \frac{\tilde{t}}{A_G} (B l_y + (n-1) B_L l_y^*).$$

Необходимое число записываемых дисков в год с номером  $t$  ( $t = \bar{t} + \tilde{t}$ ):

$$\begin{aligned} N(t) = & \left\lceil \frac{1}{1 - q_w} \left( \left\lfloor D_G^* E^{t-\tilde{t}-1} \right\rfloor + (n-1) \left\lfloor \frac{D_G^* E^{t-\tilde{t}-1}}{k} \right\rfloor k^* \right) \right\rceil + \\ & + \left\lceil \frac{1}{1 - q_w} \left[ q \left( \tilde{t} \left( l_y + (n-1) \left\lfloor \frac{l_y}{k} \right\rfloor k^* \right) + \right. \right. \right. \\ & \left. \left. \left. + (K+1) \frac{D_G^* A_G}{B_G} (1 - E^{t-\tilde{t}-1}) + (n-1) (t - \tilde{t} - 1) k^* \right) \right] \right\rceil. \end{aligned}$$

Плановой емкости архив должен достигнуть не позднее года  $t_{\text{full}}$ . Исходя из соотношения

$$l \leq \tilde{t}l_y + \sum_{i=1}^{t_{\text{full}}-\tilde{t}} l_f(i),$$

получим

$$t_{\text{full}} = \left\lceil \log_E \left( 1 - \frac{B_G}{D_G^* A_G} (1 - \tilde{t}l_y) \right) \right\rceil + \tilde{t}.$$

При  $t > t_{\text{full}}$  для поддержания архива ежегодно потребуется приобретать (1) дисков.

Параметр  $t_f$  оценим, исходя из соотношения  $l_f(t_f - \tilde{t}) < 1$ . В результате получим

$$t_f = \left\lceil \log_E \left( \frac{1}{D_G^*} \right) \right\rceil + \tilde{t} + 1.$$

Модель построена.

### 3 Пример использования модели

Необходимо оценить параметры создаваемого архива с максимальной емкостью  $W = 32$  ТБ. Архив будет создаваться на оптических дисках типа BD-R емкостью  $V = 25$  ГБ, организованных в гибридные структуры с использованием RAID16 ( $k = 2$ ,  $m = 2$ ). Требуемая проектом надежность хранения при  $q = 0,1$  и  $q_w = 0,02$  (значения параметров получены экспериментально) достигается при  $n = 3$ . Максимальный объем поступающей информации в год  $W_y = 4$  ТБ. Весь архив будет обслуживаться минимальным составом аппаратных средств (один привод BD-R) и экспериментально получены следующие значения параметров:  $t_w = t_{wL} = 33$  мин;  $t_r = t_{rL} = 18$  мин;  $t_{r0} = t_{r0L} = 15$  мин. На технические работы в архиве в год выделено  $T = 131\,400$  мин (4 часа в сутки).

Архив будет реализован по варианту 4. Расчеты по модели показывают:  $\tilde{t} = 7$  лет;  $t_{\text{full}} = 10$  лет;  $t_f = 14$  лет (т.е. проектная емкость архива будет достигнута). Весь архив разместится на 6400 дисках. В процессе заполнения архива максимальное число закупаемых дисков придется на год  $t = 8$ ;  $N(8) = 2009$ . После заполнения архива на поддержание его в состоянии целостности ежегодно потребуется приобрести  $N(t) = 654$  диска.

### 4 Заключение

В результате исследования построена аналитическая модель функционирования долговременного электронного архива, построенного на оптических дисках, объединенных в гибридные структуры. Модель рассматривает 4 варианта заполнения архива в зависимости от пропускной способности аппаратуры и позволяет

определить число ежегодно приобретаемых дисков для создания архива и поддержания его целостности, а также основные параметры времени, характеризующие процесс заполнения архива. Построенная модель может быть полезна на практике при анализе эксплуатационных характеристик проектируемых архивов длительного хранения электронной информации.

## Литература

1. *Corrado E. M., Moulaison H. L.* Digital preservation for libraries, archives, and museums. — Lanham, MD, USA: Rowman & Littlefield Pubs., 2014. 294 p.
2. *Залаев Г. З., Каленов Н. Е., Цветкова В. А.* Некоторые вопросы длительного хранения электронных документов // Научно-техническая информация. Серия 1: Организация и методика информационной работы, 2016. № 12. С. 22–28.
3. ISO/TR 18492:2005 Long-term preservation of electronic document-based information. [http://www.iso.org/iso/catalogue\\_detail.htm?csnumber=38716](http://www.iso.org/iso/catalogue_detail.htm?csnumber=38716).
4. ISO/IEC 10995:2011(E). Information technology — digitally recorded media for information interchange and storage — test method for the estimation of the archival lifetime of optical media. 2nd ed. June 15, 2011. 28 p.
5. *Smith E.* When discs die // Tedium, February 2, 2017.
6. *Zheng J., Slaterry O. T.* NIST/Library of Congress (LC) optical disc longevity study. NIST, 2007. Final Report. 32 p. [https://www.loc.gov/preservation/resources/rt/NIST\\_LC\\_OpticalDiscLongevity.pdf](https://www.loc.gov/preservation/resources/rt/NIST_LC_OpticalDiscLongevity.pdf).
7. *Чернышов А. В.* Модель планирования количества оптических дисков, необходимых для создания и поддержания долговременного электронного архива // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления, 2019. Т. 15. Вып. 4. С. 578–591. doi: 10.21638/11702/srbu10.2019.413.
8. Data Archiver LB-DH8 series. <http://panasonic.net/avc/archiver/lb-dh8>.
9. *Чернышов А. В.* Модель надежности долговременного хранения информации на оптических дисках, объединенных в гибридные структуры // Известия высших учебных заведений. Поволжский регион. Технические науки, 2020. № 2(54). С. 25–33. doi: 10.21685/2072-3059-2020-2-3.

*Поступила в редакцию 26.08.21*

## DETERMINATION OF THE REQUIRED NUMBER OF OPTICAL DISKS ORGANIZED INTO HYBRID STRUCTURES FOR THE PURPOSE OF DEVELOPMENT AND MAINTENANCE OF LONG-TERM ELECTRONIC ARCHIVES

*A. V. Chernyshov*

Mytischki Branch of Bauman Moscow State Technical University, 1, 1st Institutskaya Str., Mytischki 141005, Moscow Region, Russian Federation

**Abstract:** Hybrid structures of optical disks are a method of archival storage of information with increased reliability, similar to RAID arrays of disks. Unlike classic RAID arrays in hybrid structures, one of the copies of information is stored on single disks that provides quick and easy access to information using just one standard drive. During long-term storage, optical disks with recording can degrade. Knowing the number of optical disks that will need to be purchased yearly to create an archive and then to maintain its integrity is necessary for planning the life cycle of the archive and the costs of its maintenance at the design stage. The paper presents analytical models that allow one to determine the required number of optical disks yearly for four possible options (strategies) for filling the archive as well as time points that characterize the process of filling the archive and to assess the principal achievability of the planned archive capacity. An example of a real calculation is given.

**Keywords:** long-term electronic archives of information; write once optical disks (WORM); hybrid structures of optical disks; replacement of failed optical disks; yearly required number of optical disks

**DOI:** 10.14357/08696527220210

### References

1. Corrado, E. M., and H. L. Moulaison. 2014. *Digital preservation for libraries, archives, and museums*. Lanham, MD: Rowman & Littlefield Publs. 294 p.
2. Zalaev, G. Z., N. E. Kalenov, and V. A. Tsvetkova. 2016. Some issues of long-term storage of electronic documents. *Scientific Technical Information Processing* 43(4):268–274.
3. ISO/TR 18492:2005. 2005. Long-term preservation of electronic document-based information. Available at: <https://www.iso.org/standard/38716.html> (accessed March 22, 2022).
4. ISO/IEC 10995:2011(E). June 15, 2011. Information technology — digitally recorded media for information interchange and storage — test method for the estimation of the archival lifetime of optical media. 2nd ed. Available at: [https://standards.iso.org/ittf/PubliclyAvailableStandards/c056910\\_ISO\\_IEC\\_10995\\_2011\(E\).zip](https://standards.iso.org/ittf/PubliclyAvailableStandards/c056910_ISO_IEC_10995_2011(E).zip) (accessed March 22, 2022).

5. Smith, E. February 2, 2017. When discs die. *Tedium*. Available at: <https://tedium.co/2017/02/02/disc-rot-phenomenon/> (accessed March 22, 2022).
6. Zheng, J., and O. T. Slattery. 2007. NIST/Library of Congress (LC) optical disc longevity study. NIST. Final Report. 32 p. Available at: [https://www.loc.gov/preservation/resources/rt/NIST\\_LC\\_OpticalDiscLongevity.pdf](https://www.loc.gov/preservation/resources/rt/NIST_LC_OpticalDiscLongevity.pdf) (accessed March 22, 2022).
7. Chernyshov, A. V. 2019. Model' planirovaniya kolichestva opticheskikh diskov, neobkhodimyykh dlya sozdaniya i podderzhaniya dolgovremennogo elektronnoy arkhiva [Model for planning the number of optical disks needed to create and maintain a long-term electronic archive]. *Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes* 15(4):578–591. doi: 10.21638/11702/spbu10.2019.413.
8. Data Archiver LB-DH8 series. Available at: <http://panasonic.net/avc/archiver/lb-dh8/> (accessed March 22, 2022).
9. Chernyshov, A. V. 2020. Model' nadezhnosti dolgovremennogo khraneniya informatsii na opticheskikh diskakh, ob"yedinennykh v gibridnyye struktury [Model of reliability of long-term storage of information on optical disks united in hybrid structures]. *University Proceedings. Volga Region. Technical Sciences* 2(54):25–33.

*Received August 26, 2021*

## Contributor

**Chernyshov Alexandr V.** (b. 1970) — Candidate of Science (PhD) in technology, associate professor, Department of Applied Mathematics, Computer Science and Engineering, Mytischki Branch of Bauman Moscow State Technical University, 1, 1st Institutskaya Str., Mytischki 141005, Moscow Region, Russian Federation; sch-ru@yandex.ru

## МОДЕЛЬ ПРОЦЕССА ИЗВЛЕЧЕНИЯ НОВЫХ ТЕРМИНОВ И ТОНАЛЬНЫХ СЛОВ ИЗ ТЕКСТОВ\*

*И. М. Зацман*<sup>1</sup>

**Аннотация:** Рассматривается модель, ориентированная на проектирование информационных технологий и автоматизированных систем, обеспечивающих извлечение новых терминов в процессе программного и экспертного анализа текстов (далее — информационно-технологически ориентированная модель, или модель ИТО). Предлагается модификация этой модели для описания извлечения новых терминов из научных документов и тональных слов из сообщений пользователей социальных сетей. Рассмотрение этих двух моделей представлено в статье как поэтапная трансформация спиральной модели генерации нового знания. Основная цель статьи состоит в описании этапов трансформации, которая рассматривается на примере извлечения новых терминов из текстов научных документов и тональных слов из сообщений пользователей социальных сетей.

**Ключевые слова:** спиральная модель; новые термины; модель ИТО; тональные слова; модификация модели ИТО; трансформация спиральной модели

**DOI:** 10.14357/08696527220211

### 1 Введение

Цель статьи — трансформировать спиральную модель генерации знания (= модель SECI) [1–3] так, чтобы в результате получить модель для описания процесса извлечения из текстов новых терминов и тональных слов (sentiments), которые иногда называют терминами-маркерами. Они говорят об оценочном содержании (например, отрицательные, нейтральные, положительные оценки) и/или эмоциях авторов текстов [4, 5]. Трансформация включает два этапа: сначала на основе модели SECI создается модель ИТО, а затем она модифицируется так, чтобы с ее помощью можно было описать процесс одновременного извлечения из текстов новых терминов и тональных слов. Первый этап трансформации был описан ранее в работах [6–8]. Применение модели ИТО как теоретической основы разработки информационных систем (ИС) обнаружения и заполнения лагун в системах лингвистических знаний рассмотрено в работах [9–12]. На первом этапе создания модели ИТО остановимся кратко, только на тех его стадиях, которые необходимы для описания второго этапа трансформации модели SECI.

---

\*Исследование выполнено с использованием ЦКП «Информатика» ФИЦ ИУ РАН при финансовой поддержке РФФИ (проект 20-04-60185).

<sup>1</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук, izatsman@yandex.ru

В настоящее время модель SECI стала одной из самых популярных в сфере экономики. Она носит объяснительный характер, является апостериорной и описывает только качественно процессы генерации знания, которые уже состоялись. В ней определены две категории знания: индивидуальные (созданные одним экспертом) и коллективные (согласованные в коллективе экспертов). Каждая из них делится на две подкатегории: выраженное (явное) и невыраженное (неявное) знание. Таким образом, эта модель включает в себя следующие четыре вида знания:

- (1) индивидуальное неявное знание (например, личностное понимание нового термина одним экспертом, сформированное им в процессе семантического анализа текстов с этим термином);
- (2) коллективное неявное знание (согласованное понимание нового термина в коллективе экспертов);
- (3) коллективное явное знание (дефиниция нового термина, согласованная в коллективе экспертов);
- (4) индивидуальное явное знание (персональная дефиниция нового термина, не согласованная в коллективе экспертов).

В модели SECI помимо этих четырех видов знания определены следующие четыре процесса перехода (их преобразования): социализация, экстернализация, синтез и интернализация (Socialisation, Externalisation, Combination, Internalisation — SECI). Описание этих процессов дано в работах [1–3]. Их смысловое содержание будет описано ниже в контексте извлечения из текстов новых терминов и тональных слов. По определению модели SECI, каждый виток ее спирали включает последовательность процессов: интернализация → социализация → экстернализация → синтез, после которого с процесса интернализации начинается следующий виток генерации нового знания (рис. 1).

Рассматриваемые в статье модели, создаваемые на основе модели SECI, применяются при проектировании информационной технологии извлечения из текстов новых терминов и тональных слов. Ее проектирование выполняется в рамках проекта РФФИ № 20-04-60185 по терминологическому сопоставлению научных документов о пандемии нового коронавируса и сообщений пользователей социальных сетей, обсуждающих пандемию с использованием тональных слов. Появление новых терминов обусловлено пополнением и изменением во времени научного знания о пандемии.

В рамках этого проекта модель ИТО, создание которой рассматривается в разд. 2, используется только для описания процесса извлечения из текстов новых терминов. Ее модифицированный вариант, который рассматривается в разд. 3 и который будем называть модель ИТО-Sent, применяется для описания процесса одновременного извлечения из текстов новых терминов и тональных слов.



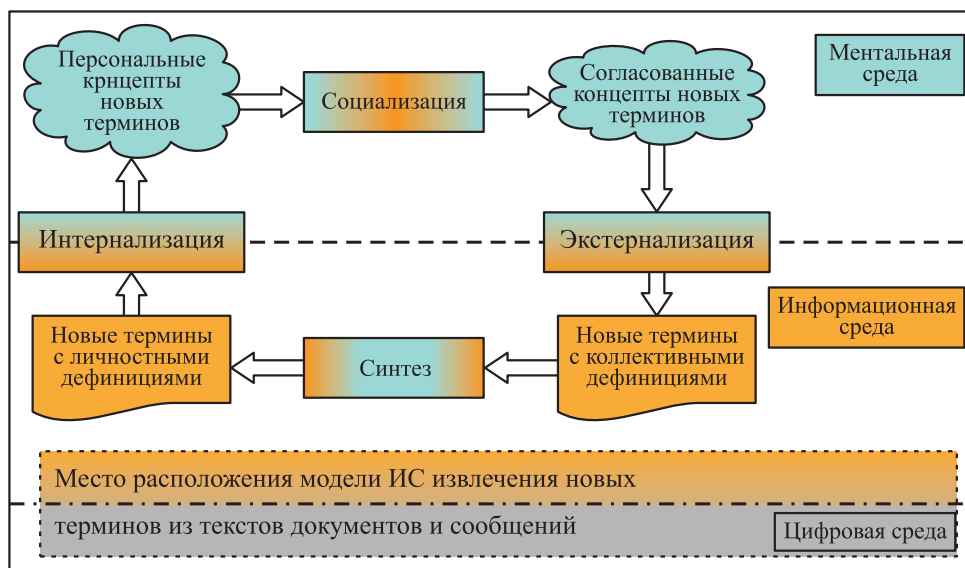
**Рис. 1** Спиральная модель генерации знаний [13, с. 69]

## 2 Первый этап трансформации

В процессе трансформации модели SECI с целью создания моделей ИТО и ИТО-Sent информатика позиционируется как комплекс отраслей науки, изучающих информационные трансформации в искусственных, живых и социальных системах, а не только в компьютерах и сетях, при этом информационные трансформации позиционируются как основной объект исследований в ментальной, информационной, цифровой и других средах предметной области информатики, а также на границах между средами различной природы [14–17].

Сначала повернем рис. 1 на  $90^\circ$  по часовой стрелке (рис. 2) и выделим в модели SECI ментальную среду *невыраженного знания* (в проекте РФФИ — это персональные и согласованные концепты новых терминов, которые эксперты обнаружили, анализируя документы и сообщения) и информационную среду *выраженного знания* (это аннотации новых терминов с их личностными и коллективными дефинициями).

Проведем границу (штриховая линия на рис. 2) между этими двумя средами, распределив по ним подкатегории каждой из двух категорий знания (как индивидуального, так и коллективного). Таким образом, модель SECI описывает четыре процесса перехода между четырьмя видами знания в двух средах: ментальной и информационной. Затем добавим цифровую среду (штрихпунктирная линия) и обозначим с помощью прямоугольника модель той информационной системы, которую после ее создания объединим с моделью SECI.



**Рис. 2** Модель SECI и место расположения модели ИС

Теоретической основой создания модели ИТО служит парадигма использования сред предметной области информатики и границ между ними [17] при описании в проекте РФФИ процесса извлечения экспертами новых терминов из текстов, в котором:

- ментальная среда содержит концепты терминов как понятий человеческого знания в умах (сознании) людей, которые формируются в рамках той или иной вербальной знаковой системой (об обусловленности концептов знаковой системой см. работу [18]);
- информационная среда включает в себя перцептивные формы представления концептов знания человека (тексты, диаграммы, формулы, изображения и др.);
- цифровая среда охватывает компьютерные коды документов и сообщений в компьютерах, сетях, базах данных и т. д.

В соответствии с парадигмой деления предметной области информатики на среды разной природы в проекте четыре вида знания модели SECI распределены между двумя средами: понятия невыраженного (неявного) знания находятся в ментальной среде, а их перцептивные формы представления (т. е. выраженное (явное) знание) находятся в информационной среде. Процессы социализации и синтеза охватывают виды знания обеих сред, но для упрощения рисунков

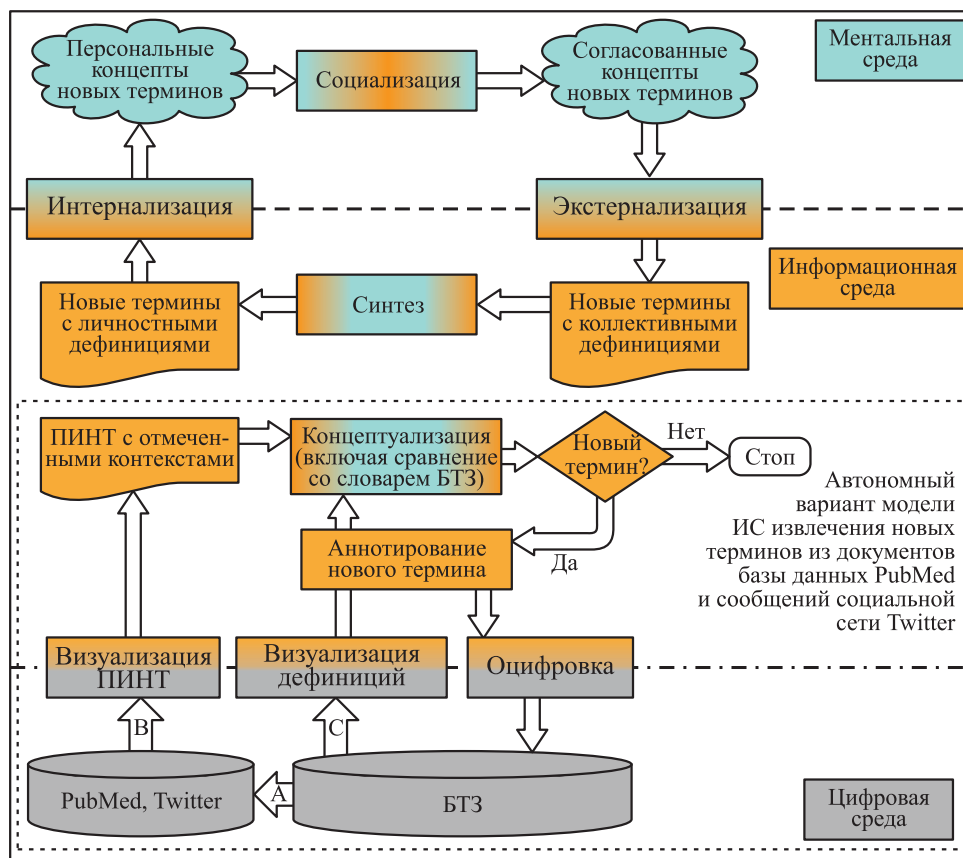


Рис. 3 Модель SECI и автономный вариант модели ИС

статьи разместим эти два процесса в разных средах. Такое их размещение обусловлено начальной и конечной фазами каждого из них: начальная и конечная фазы процесса социализации относятся к ментальной среде, а процесса синтеза — к информационной. Процессы интернализации и экстернализации начинаются и заканчиваются в разных средах, поэтому они размещены на границе между ними.

В проекте эксперты используют коллекцию документов PubMed и массив сообщений о пандемии нового коронавируса, выгруженный из социальной сети Twitter (рис. 3). При этом они анализируют только те документы и сообщения, которые потенциально могут содержать новые термины и/или тональные слова. Такие документы и сообщения будем называть потенциальными источниками новых терминов и/или тональных слов (ПИИТ). Они отбираются програм-

мно с проверкой по эталону новизны для последующего выполнения процесса их концептуализации<sup>1</sup>. В каждом ПИНТ программно выделяются контексты потенциально новых терминов (ПНТ)<sup>2</sup> и тональных слов. В процессе концептуализации эксперты сопоставляют смысловое содержание отмеченных контекстов ПНТ с эталоном новизны, в качестве которого выступает словарь уже известных терминов, включающий их дефиниции, отобранный экспертами и входящий в состав базы терминологических знаний (БТЗ)<sup>3</sup>.

Необходимость в экспертном сопоставлении контекстов ПНТ и словарных дефиниций вызвана тем, что программа может принять за новый термин синоним другого термина, уже существующего в словаре. Эксперт сопоставляет каждый отмеченный контекст ПНТ с дефинициями известных терминов, задавая ключевые слова контекста и проводя по ним поиск дефиниций в словаре терминов. Сопоставление найденных дефиниций с контекстом служит основой или для отклонения термина, предложенного программой, или для присвоения ему статуса нового термина. В последнем случае эксперт создает личностную дефиницию<sup>4</sup> и формирует личностную аннотацию термина (см. процесс «Аннотирование нового термина» на рис. 3). Она включает кроме дефиниции контекст, из которого был извлечен новый термин, его адрес и адрес контекста термина в документе или сообщении.

Отметим, что модель SECI не предполагает использование ПИНТ, но ее интеграция с моделью ИС открывает эту возможность. Использование ПИНТ становится возможным после включения в модель ИС четырех процессов: визуализации (отдельно для ПИНТ и дефиниций терминов словаря), концептуализации, аннотирования и оцифровки. На рис. 3 представлен автономный вариант модели ИС до ее объединения с моделью SECI. Этот вариант описывает работу одного эксперта без согласования личностных дефиниций обнаруженных новых терминов.

Таким образом, чтобы найти новые термины, сначала программно формируется массив ПИНТ с использованием словаря из БТЗ (стрелка с буквой «А» на рис. 3), отмечаются контексты ПНТ (стрелка с буквой «В»), которые затем визуализируются. Потом выполняется их концептуализация, когда эксперт сравнивает контекст каждого ПНТ с найденными дефинициями словаря, извлекаемыми из БТЗ (стрелка с буквой «С»). Если термин признается экспертом новым, то его аннотация с личностной дефиницией вводится в БТЗ, включающую словарь известных терминов и аннотации новых терминов с их дефинициями. В противном случае обработка этого контекста ПНТ прекращается.

---

<sup>1</sup>Использование словаря как эталона новизны условно обозначено стрелкой с буквой «А» на рис. 3.

<sup>2</sup>Поиск документов, сообщений и выделение в них границ контекстов условно обозначено стрелкой с буквой «В» на рис. 3.

<sup>3</sup>Поиск и извлечение дефиниций из БТЗ условно обозначено стрелкой с буквой «С» на рис. 3.

<sup>4</sup>В модели ИТО личностная дефиниция нового термина согласовывается в коллективе экспертов проекта (см. процесс «Социализация» на рис. 4).

Рисунок 3 содержит два независимых сегмента. Верхний сегмент соответствует модели SECI, адаптированной для описания процесса извлечения экспертами новых терминов из текстов, с ее четырьмя переходными процессами и разделением знаний на четыре вида. Нижний сегмент соответствует автономному варианту модели ИС, которая содержит коллекцию документов PubMed и массив сообщений Twitter, четыре новых процесса перехода, которые связывают их с БТЗ.

Создание модели ИТО завершается объединением этих двух сегментов. Это объединение осуществляется путем разрыва связи между процессами экстернализации и синтеза с последующим изменением последовательности переходных процессов, их входа, выхода и, частично, их содержания (рис. 4). В мо-

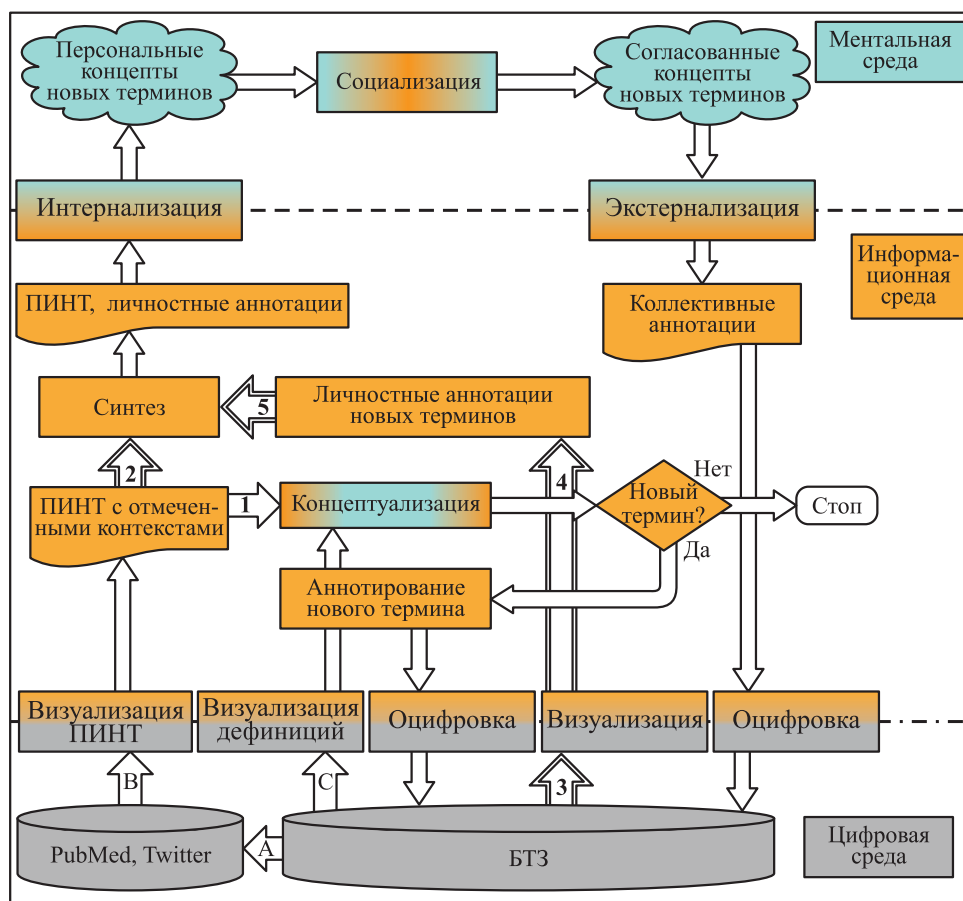


Рис. 4 Модель ИТО

дели SECI новые термины с коллективными дефинициями подаются на вход процесса синтеза, при выполнении которого формируются новые термины с личностными дефинициями, в качестве источника которых модель SECI указывает коллективное знание экспертов, не конкретизируя его первоисточники. Модель ИТО содержит описание ПИНТ и процессов извлечения из них новых терминов.

В предлагаемой модели ИТО, чтобы сформировать коллективную аннотацию нового термина, из БТЗ извлекается его личная аннотация, она визуализируется (стрелки 3 и 4) и затем, в процессе синтеза, объединяется с тем документом или сообщением (стрелки 2 и 5), из которого был извлечен новый термин в процессе концептуализации (стрелка 1), для последующего согласования его дефиниции в коллективе экспертов согласно модели SECI. Но есть одно существенное отличие от этой модели: согласованная аннотация вводится в БТЗ.

Модель ИТО описывает и ту ситуацию, когда в процессе концептуализации одного и того же контекста разные эксперты могут сформировать разные аннотации. Если имеются несколько личностных аннотаций одного термина, сформированных разными экспертами на основе одного и того же контекста ПИНТ, то все эти аннотации извлекаются из БТЗ, визуализируются и объединяются с тем документом или сообщением, из которого был извлечен новый термин. В этих случаях в процессе социализации используются личностные аннотации, согласовывается понимание этого термина между экспертами и ими формируется коллективная аннотация, которая оцифровывается и загружается в БТЗ.

### 3 Второй этап трансформации

Модель ИТО дает описание процесса извлечения новых терминов из текстов без поиска тональных слов. Как отмечалось выше, цель проекта РФФИ — терминологическое сравнение научной системы знаний о пандемии коронавируса и знаний пользователей социальных сетей о ней с учетом тональных слов (= термины-маркеры, на рис. 5 они обозначены как «Т-М»). В рамках проекта учитывается использование тональных слов в сообщениях социальных сетей при выполнении второго этапа трансформации модели SECI и создания модели ИТО-Sent. На этом этапе добавляется словарь терминов-маркеров (= словарь Т-М) и процесс их аннотирования, который выполняется в тех случаях, если документ (сообщение) содержит новое тональное слово. Согласно модели ИТО-Sent, программа при простановке отметки «ПИНТ» использует два словаря: терминов и тональных слов. Если в первом словаре отсутствует хотя бы один термин, входящий в документ (сообщение), то его программа помечает как ПИНТ. Если во втором словаре отсутствует хотя бы одно тональное слово, входящее в сообщение, то его программа помечает как ПИНТ.

В моделях ИТО и ИТО-Sent процесс *концептуализации* выполняется для всех ПИНТ, что показано стрелкой с цифрой 1 (см. рис. 4 и 5). Однако

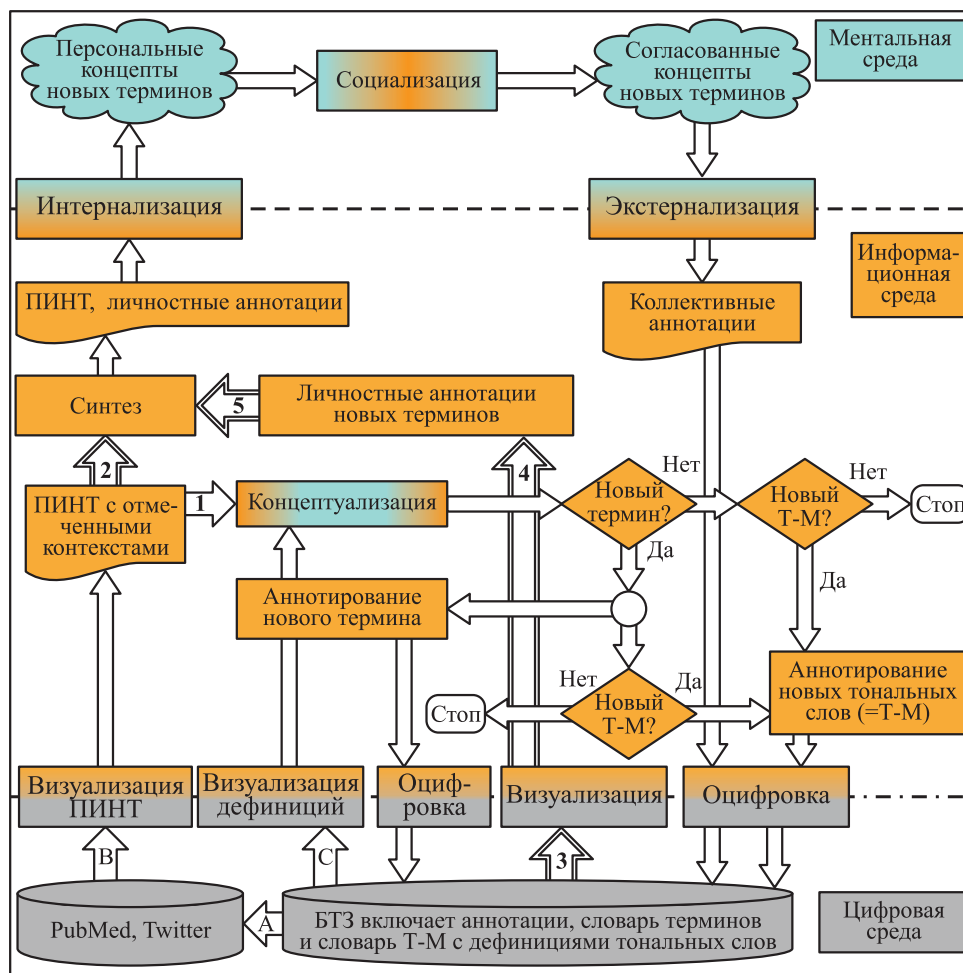


Рис. 5 Модель ITO-Sent

на вход процесса *синтеза* подаются только те тексты, которые действительно оказались источниками нового термина и/или тонального слова, что показано стрелкой с двойным контуром и цифрой 2. Их новизна выясняется только после завершения *концептуализации* контекстов ПИНТ, и только их аннотации сохраняются в БТЗ. После извлечения из БТЗ эти аннотации становятся входом операции *синтеза*. Это показано стрелками с двойным контуром и цифрами 3, 4 и 5. Двойной контур стрелок на рис. 4 и 5 подчеркивает одновременность использования в операции *синтеза* и ПИНТ, и аннотаций.

## 4 Заключение

На основе модели ИТО-Sent в рамках проекта РФФИ разрабатывается информационная технология, которая обеспечивает экспертное сравнение двух типов информации: документов, полученных из базы данных PubMed, и сообщений пользователей Twitter, обсуждающих различные аспекты пандемии SARS-CoV-2. Поэтому модели ИТО и ИТО-Sent и были названы ориентированными на проектирование информационных технологий. В заключение отметим, что модель ИТО в настоящее время используется для описания процесса извлечения нового лингвистического знания о значениях немецких модальных глаголов, а также при проектировании информационной технологии извлечения нового терминологического знания из текстов медицинских документов [6, 7, 19].

## Литература

1. Nonaka I. The knowledge-creating company // Harvard Bus. Rev., 1991. Vol. 69. No. 6. P. 96–104.
2. Nonaka I. A dynamic theory of organizational knowledge creation // Organ. Sci., 1994. Vol. 5. No. 1. P. 14–37.
3. Нонака И., Такеучи Х. Компания — создатель знания / Пер. с англ. — М.: Олимп-бизнес, 2003. 384 с. (Nonaka I., Takeuchi H. The knowledge-creating company. — Oxford, NY, USA: Oxford University Press, 1995. 284 p.).
4. Alamoodi A. H., Zaidan B. B., Zaidan A. A., et al. Sentiment analysis and its applications in fighting COVID-19 and infectious diseases: A systematic review // Expert Syst. Appl., 2021. Vol. 167. P. 1–13. <https://www.sciencedirect.com/science/article/pii/S0957417420308988>.
5. Nemes L., Kiss A. Information extraction and named entity recognition supported social media sentiment analysis during the COVID-19 pandemic // Appl. Sci. — Basel, 2021. Vol. 11. No. 22. P. 1–33. <https://www.mdpi.com/2076-3417/11/22/11017>.
6. Zatsman I., Khakimova A. New knowledge discovery for creating terminological profiles of diseases // 22nd European Conference on Knowledge Management Proceedings. — Reading, U.K.: Academic Publishing International Ltd., 2021. P. 837–846.
7. Zatsman I. A model of goal-oriented knowledge discovery based on human-computer symbiosis // 16th Forum (International) on Knowledge Asset Dynamics Proceedings. — Matera, Italy: Arts for Business Institute, 2021. P. 297–312.
8. Зацман И. М. Компьютерная и экономическая модели генерации нового знания: сопоставительный анализ // Системы и средства информатики, 2021. Т. 31. № 4. С. 84–96.
9. Zatsman I. Finding and filling lacunas in knowledge systems // 20th European Conference on Knowledge Management Proceedings. — Reading, U.K.: Academic Publishing International Ltd., 2019. Vol. 2. P. 1143–1151.
10. Zatsman I. Finding and filling lacunas in linguistic typologies // 15th Forum (International) on Knowledge Asset Dynamics Proceedings. — Matera, Italy: Arts for Business Institute, 2020. P. 780–793.
11. Zatsman I. Three-dimensional encoding of emerging meanings in AI-systems // 21st European Conference on Knowledge Management Proceedings. — Reading, U.K.: Academic Publishing International Ltd., 2020. P. 878–887.

12. Зацман И. М. Формы представления нового знания, извлеченного из текстов // Информатика и её применения, 2021. Т. 15. Вып. 3. С. 83–90.
13. Wierzbicki A. P., Nakamori Y. Basic dimensions of creative space // Creative space: Models of creative processes for knowledge civilization age / Eds. A. P. Wierzbicki, Y. Nakamori. — Berlin–Heidelberg: Springer Verlag, 2006. P. 59–90.
14. Denning P., Rosenbloom P. Computing: The fourth great domain of science // Commun. ACM, 2009. Vol. 52. No. 9. P. 27–29.
15. Rosenbloom P. S. On computing: The fourth great scientific domain. — Cambridge, MA, USA: MIT Press, 2013. 308 p.
16. Зацман И. М. Таблица интерфейсов информатики как информационно-компьютерной науки // Научно-техническая информация. Серия 1: Организация и методика информационной работы, 2014. № 11. С. 1–15.
17. Зацман И. М. Интерфейсы третьего порядка в информатике // Информатика и её применения, 2019. Т. 13. Вып. 3. С. 82–89.
18. Зацман И. М. Кодирование концептов в цифровой среде // Информатика и её применения, 2019. Т. 13. Вып. 4. С. 97–106.
19. Зацман И. М. Проблемно-ориентированная актуализация словарных статей двуязычных словарей и медицинской терминологии: сопоставительный анализ // Информатика и её применения, 2021. Т. 15. Вып. 1. С. 94–101.

*Поступила в редакцию 20.03.22*

---

---

## A MODEL OF DISCOVERING NOVEL TERMS AND SENTIMENTS IN TEXTS

*I. M. Zatsman*

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The paper considers a model focused on the design of information technology and systems that ensure discovering novel terms in the process of computer–human analysis of texts (hereinafter referred to as the information technology-oriented model, or the ITO model). A modification of the ITO model is proposed to describe discovering novel terms in scientific documents and sentiments in the messages of social network users. Consideration of these two models is presented in the paper as a step-by-step transformation of the spiral model of new knowledge generation. The main purpose of the paper is to describe the step-by-step transformation which is considered by the example of discovering novel terms in the texts of scientific documents and sentiments in the messages of social network users.

**Keywords:** spiral model; novel terms; ITO model; sentiments; ITO model modification; spiral model transformation

**DOI:** 10.14357/08696527220211

## Acknowledgments

The research was carried out using the infrastructure of the Shared Research Facilities “High Performance Computing and Big Data” (CKP “Informatics”) of FRC CSC RAS (Moscow) with the financial support of the Russian Foundation for Basic Research (project 20-04-60185).

## References

1. Nonaka, I. 1991. The knowledge-creating company. *Harvard Bus. Rev.* 69(6):96–104.
2. Nonaka, I. 1994. A dynamic theory of organizational knowledge creation. *Organ. Sci.* 5(1):14–37.
3. Nonaka, I., and H. Takeuchi. 1995. *The knowledge-creating company*. Oxford, NY: Oxford University Press. 284 p.
4. Alamoodi, A. H., B. B. Zaidan, A. A. Zaidan, *et al.* 2021. Sentiment analysis and its applications in fighting COVID-19 and infectious diseases: A systematic review. *Expert Syst. Appl.* 167:1–13. Available at: <https://www.sciencedirect.com/science/article/pii/S0957417420308988> (accessed May 6, 2022).
5. Nemes, L., and A. Kiss. 2021. Information extraction and named entity recognition supported social media sentiment analysis during the COVID-19 pandemic. *Appl. Sci. — Basel* 11(22):1–33. Available at: <https://www.mdpi.com/2076-3417/11/22/11017> (accessed May 6, 2022).
6. Zatsman, I., and A. Khakimova. 2021. New knowledge discovery for creating terminological profiles of diseases. *22nd European Conference on Knowledge Management Proceedings*. Reading, U.K.: Academic Publishing International Ltd. 837–846.
7. Zatsman, I. 2021. A model of goal-oriented knowledge discovery based on human-computer symbiosis. *16th Forum (International) on Knowledge Asset Dynamics Proceedings*. Matera, Italy: Arts for Business Institute. 297–312.
8. Zatsman, I. 2021. Komp’yuternaya i ekonomicheskaya modeli generatsii novogo znaniya: sopostavitel’nyy analiz [Computer and economic models of new knowledge generation: A comparative analysis]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 31(4):84–96.
9. Zatsman, I. 2019. Finding and filling lacunas in knowledge systems. *20th European Conference on Knowledge Management Proceedings*. Reading, U.K.: Academic Publishing International Ltd. 2:1143–1151.
10. Zatsman, I. 2020. Finding and filling lacunas in linguistic typologies. *15th Forum (International) on Knowledge Asset Dynamics Proceedings*. Matera, Italy: Arts for Business Institute. 780–793.
11. Zatsman, I. 2020. Three-dimensional encoding of emerging meanings in AI-systems. *21st European Conference on Knowledge Management Proceedings*. Reading, U.K.: Academic Publishing International Ltd. 878–887.
12. Zatsman, I. M. 2021. Formy predstavleniya novogo znaniya, izvlechenno iz tekстов [Forms representing new knowledge discovered in texts]. *Informatika i ee Primeniye — Inform. Appl.* 15(3):83–90.
13. Wierzbicki, A. P., and Y. Nakamori. 2006. Basic dimensions of creative space. *Creative space: Models of creative processes for knowledge civilization age*. Eds. A. P. Wierzbicki and Y. Nakamori. Berlin–Heidelberg: Springer Verlag. 59–90.

14. Denning, P., and P. Rosenbloom. 2009. Computing: The fourth great domain of science. *Commun. ACM* 52(9):27–29.
15. Rosenbloom, P. 2013. *On computing: The fourth great scientific domain*. Cambridge, MA: MIT Press. 308 p.
16. Zatsman, I. 2014. A table of interfaces of informatics as computer and information science. *Sci. Tech. Inf. Proc.* 41(4):230–243.
17. Zatsman, I. M. 2019. Interfeysy tret'ego poryadka v informatike [Third-order interfaces in informatics]. *Informatika i ee Primeneniya — Inform. Appl.* 13(3):82–89.
18. Zatsman, I. M. 2019. Kodirovanie kontseptov v tsifrovoy srede [Digital encoding of concepts]. *Informatika i ee Primeneniya — Inform. Appl.* 13(4):97–106.
19. Zatsman, I. 2021. Problemno-orientirovannaya aktualizatsiya slovarnykh statey dvuyazychnykh slovarey i meditsinskoy terminologii: sopostavitel'nyy analiz [Problem-oriented updating of dictionary entries of bilingual dictionaries and medical terminology: Comparative analysis]. *Informatika i ee Primeneniya — Inform. Appl.* 15(1):94–101.

*Received March 20, 2022*

## Contributor

**Zatsman Igor M.** (b. 1952) — Doctor of Science in technology, Head of Department, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; izatsman@yandex.ru

## ПОДХОД К РАЗРАБОТКЕ КОНЦЕПТУАЛЬНОЙ СХЕМЫ БАЗЫ ГЕОДАННЫХ С ПОДДЕРЖКОЙ МНОВАРИАНТНОГО ПРЕДСТАВЛЕНИЯ ГЕООБЪЕКТОВ

*Д. А. Никишин*<sup>1</sup>

**Аннотация:** Представлена структура концептуальной схемы и основные проектные решения для построения варианта базы геоданных (БГД), обеспечивающей поддержку согласованного мультимасштабного и многовариантного представления геообъектов. Рассмотрена сущность вариантности геомodelей, выделена типология взаимосвязей между вариантными представлениями геоданных. Объединение вариантов представления геоданных в контексте общей БГД позволит обеспечить прямой доступ к ним и на этой основе реализовать новые методы решения задач анализа и обработки данных, а также обеспечить синхронизацию и согласование изменений геоданных.

**Ключевые слова:** геоданные; базы геоданных; концептуальная схема базы геоданных; моделирование геообъектов; вариантность геомodelей

**DOI:** 10.14357/08696527220212

### 1 Введение

Ранее, в публикации [1], были рассмотрены предпосылки для создания единой концептуальной схемы<sup>2</sup> классификации топографической информации, нацеленной в числе прочего на унификацию и интеграцию частных концептуальных схем как для отдельных уровней детализации (масштабов), так и для различных прикладных задач. Затем, в работе [2], был представлен вариант концептуальной схемы БГД с поддержкой обратимой генерализационной связанности моделей геообъектов.

Настоящая работа продолжает развитие данного направления исследований и посвящена особенностям реализации многовариантного представления геоданных в единой БГД. В разд. 2 рассматривается сущность вариантности и основные виды преобразований между вариантными представлениями геоданных. В разд. 3 представлены проектные решения и структуры данных, предназначенные для построения образца БГД, поддерживающей как генерализационную, так и вариантную связанность моделей геообъектов. Такой подход призван обеспечить потенциал для совершенствования функциональности БГД, в частности мето-

<sup>1</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, dmnikishin@mail.ru

<sup>2</sup>Концептуальная схема здесь понимается как система взаимосвязанных понятий, необходимая и достаточная для описания требуемых аспектов моделируемого объекта и включающая модели геообъектов, сопряженные с ними ограничения и методы их обработки.

дов анализа и отображения информации о местности [3–5]. При этом БГД можно рассматривать как один из важнейших компонентов информационных, управляющих и телекоммуникационных систем.

## 2 Вариантность геоданных

Каждый объект местности в общем случае может быть представлен несколькими версиями его модельного описания (моделями геообъекта, *геомоделями*). Эти модели могут представлять геообъект с различной степенью подробности (детализации и/или точности) описания тех или иных его свойств. При этом различия могут относиться как к пространственным, так и к семантическим характеристикам геообъекта [1].

Хотя различные геомодели могут применяться в контексте разных уровней обобщения геоданных (генерализации) и некоторых других аспектов (например, альтернативности, темпоральности [6]), под *вариантностью* далее будем понимать только одновременное существование в контексте одного уровня детализации разных представлений геообъекта, обусловленных предназначением этих моделей для использования в разных прикладных областях. При этом различия вариантов геоданных обусловлены специфическими особенностями представления ситуации на местности в конкретных прикладных областях или для решения отдельных специфических задач.

Аспект вариантности данных сам по себе не нов (см., например, [7]), но его применение к БГД представляет собой элемент новизны.

Для того чтобы охарактеризовать различия между вариантами, можно выделить следующие возможные *вариантные преобразования* некоторого исходного набора геоданных в производный набор.

1. Трансформирование геоданных, которое не вызывает ни формальных, ни содержательных изменений, а заключается, например, в преобразовании метрики геомодели в другую систему координат или проекцию.
2. Изменение формы представления данных без изменения их содержания; обычно это делается с целью адаптации данных для их специфического использования (анализа и т. п.), например перекодирование семантических характеристик модели геообъекта в альтернативную систему классификации.
3. Изменение детальности и/или точности геоданных (как отдельных моделей геообъектов, так и модели местности в целом)<sup>1</sup>, т. е. изменение содержания геомодели, в том числе:

- (а) снижение информационной нагрузки за счет исключения элементов содержания или излишней детализации, ненужных для решения специфи-

<sup>1</sup>В контексте п. 3 условно можно выделить базовые варианты, содержащие обобщенные данные (например, общетопографические), и варианты, дополняющие базовые модели данными, специфичными для различных прикладных областей.

ческих задач; это преобразование носит характер генерализации (отбор, упрощение и т. п.);

- (б) повышение информативности модели путем сбора дополнительных данных или уточнения, детализации уже имеющихся (например, при создании производной специфической модели).

Анализируя особенности некоторых межвариантных преобразований, связанные как со сбором дополнительных данных, так и с достаточно трудоемкими в некоторых случаях процедурами самого преобразования, можно отметить, что создание вариантных представлений геоданных по мере надобности (особенно в реальном времени), например на основе некоторого базового набора данных о местности, затруднительно.

Практичным подходом к обеспечению вариантности представляется заблаговременное формирование и сохранение геомodelей различной прикладной направленности. Эти сохраненные модели предполагается повторно использовать до тех пор, пока не изменится ситуация — в этом случае изменения должны вноситься в тот из вариантов, который представляет изменившееся свойство наиболее подробно, а затем распространять изменения на все остальные варианты методами вариантных преобразований.

Интеграция в контексте общей, единой многовариантной БГД различных вариантов геомodelей позволит осуществлять непосредственный переход между разными представлениями ситуации на местности или использовать их совместно. Это позволит реализовать решение некоторых задач анализа и обработки данных, находящихся на стыке разных предметных областей. Вариантные связи также обеспечат синхронизацию и согласование возможных изменений геоданных.

При отсутствии в БГД таких явных связей потребуются сложные процедуры поиска и извлечения геомodelей в наборах данных необходимой полноты и детальности.

*Применение вариантной связанности* в контексте единой БГД может обеспечить прямой доступ к более детальным вариантам геомodelей, присущим специфическим прикладным областям. При этом появляется возможность расширить функциональность такой БГД или оптимизировать некоторые методы решения задач анализа и обработки данных.

Например, «для одного масштабного уровня может понадобиться как площадная метрика дороги... так и ее осевая линия...» [8]. В первом случае речь может идти, например, об оценке объемов дорожных работ, а во втором — о построении маршрута (в контексте графа дорожной сети) для решения навигационной задачи [9, 10].

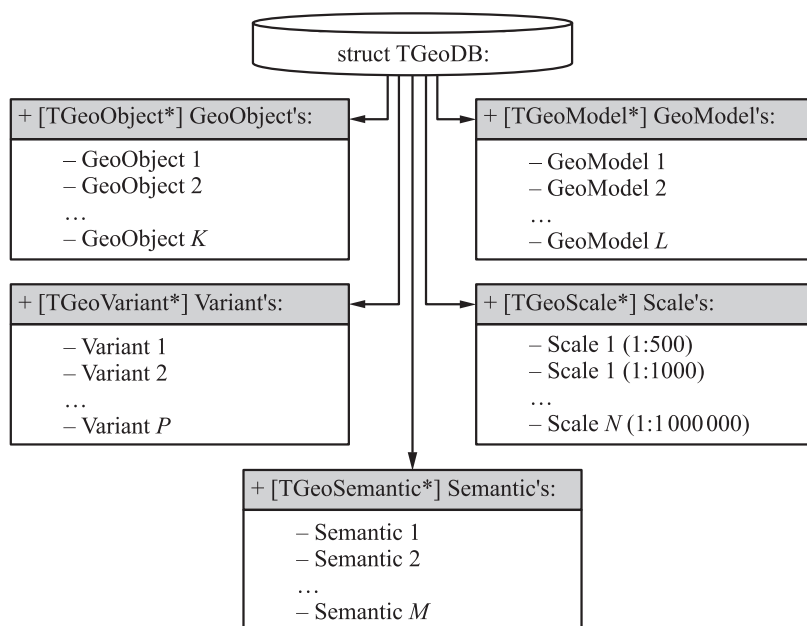
Еще один пример возможного применения вариантности: лесной массив в топографическом варианте описывается слишком обобщенно, без детализации состояния древостоя. Отсутствие такой дифференциации затрудняет оценку этого массива в качестве источника материала, но такая дифференциация может

иметь место в специфической лесотаксационной модели. Возможны и другие задачи из данной предметной области [11].

Своеобразной разновидностью вариантности можно также считать наборы геоданных, различающиеся по уровню *публичности* (открытые и содержащие специфические сведения о местности).

### 3 Модель многовариантной базы геоданных

На рис. 1 и 2 представлена<sup>1</sup> структура концептуальной схемы БГД как попытка практической реализации БГД с поддержкой связанности вариантов геомоделей, представляющих геообъект в различных прикладных контекстах. Данная структура является развитием рассмотренной в [2] мультимасштабной БГД (МБГД), в которую добавлена поддержка вариантного аспекта геоданных, поэтому такую БГД далее условно будем именовать мультимасштабной вариантной (МВБГД). Исходным прототипом служит рассмотренная в [1] общая концептуальная схема БГД с поддержкой вариантного и темпорального аспектов геоданных.



**Рис. 1** Общая структура МВБГД

<sup>1</sup> В целях более четкого выделения структур данных и сокращения объема на рис. 1 и 2 некоторые элементы структур опущены или оформлены в упрощенном виде, с применением псевдокода.

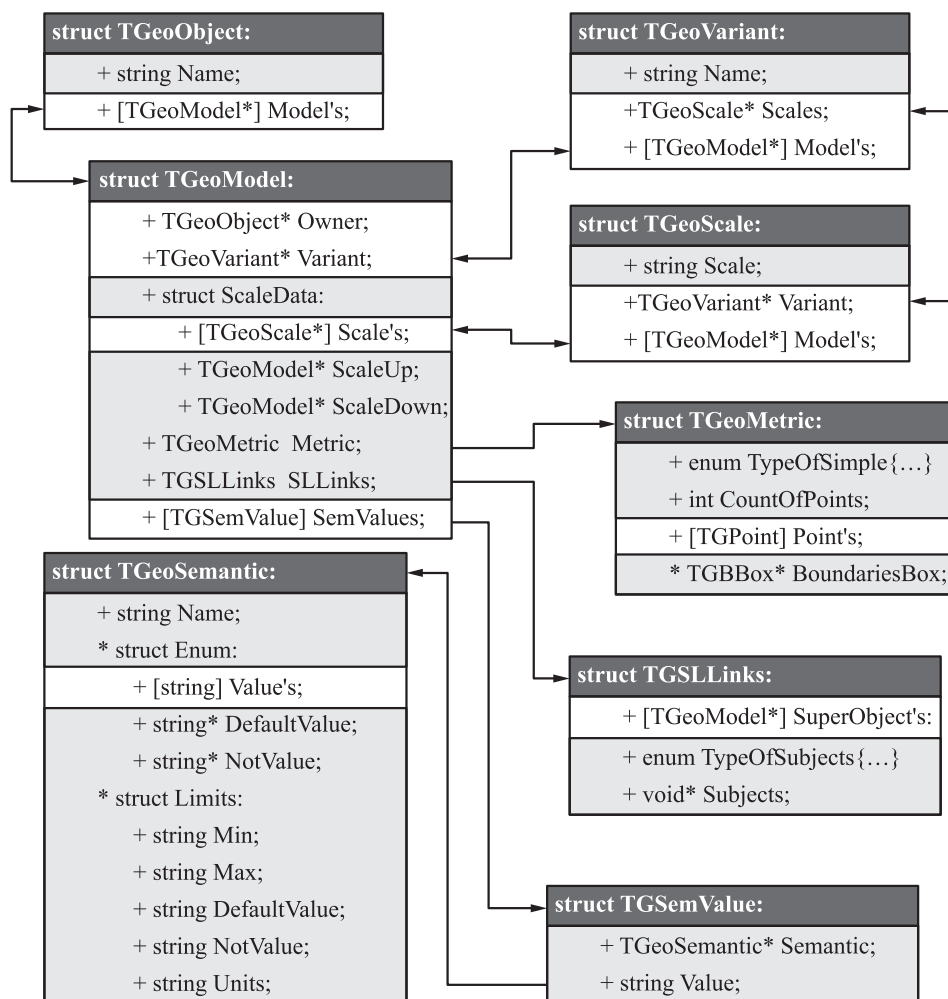


Рис. 2 Основные структуры данных МВБГД

На рис. 1 представлена общая структура МВБГД, обозначенная как «TGeoDB». Она включает в себя пять основных таблиц, четыре из которых были рассмотрены в [2] и по своей структуре соответствуют основным концептам МБГД: «геообъект» (TGeoObject), «геомодель» (TGeoModel), «семантическое свойство» (TGeoSemantic) и «масштабный уровень» (TGeoScale); новой является таблица, соответствующая концепту «вариант» (TGeoVariant). Эти таблицы содержат основные данные БГД и могут служить для обращения к соответствующим информационным объектам.

На рис. 2 представлена структура основных концептов МВБГД и связей между ними. При этом концепты рассматриваются как структуры данных соответствующих информационных объектов БГД.

Большинство представленных на рис. 2 концептов были рассмотрены в [2, 12], поэтому ниже остановимся только на вновь введенных и измененных.

1. Класс TGeoVariant является новой структурой данных, предназначенной для объединения в себе всех геомodelей, относящихся к конкретному варианту моделирования местности; при этом в контексте каждого такого варианта предусматривается разделение геоконтента по уровням детализации («масштабным уровням»). Эта структура включает в себя поля Name, Scale и Geomode's: поле Name служит идентификатором конкретного варианта геоданных, массив Scale содержит ряд указателей на доступные для этого варианта уровни детализации, а массив Model's содержит совокупность указателей на все геомodelи, релевантные данному варианту геоданных.
2. В структуру TGeoScale, представляющую в МБГД данные конкретного уровня детализации («масштабного уровня»), добавлено поле Variant, которое позволяет разделить данные одного уровня детализации на разные варианты их представления. Соответственно, массив Model's содержит совокупность указателей на геомodelи, релевантные не только данному уровню детализации, но и конкретному варианту представления геоконтента.
3. В структуру TGeoModel, предназначенную для описания модели геообъекта, добавлено поле для указателя на объект типа TGeoVariant, соответствующий конкретному варианту представления геоконтента.

Надо отметить, что структурам TGeoScale и TGeoVariant соответствуют разные множества геомodelей. Объект типа TGeoScale, ранее отвечавший за конкретный масштабный уровень, здесь фактически трансформировался в концепт масштабно-вариантного среза БГД, а множество геомodelей в объекте TGeoVariant, по сути, позволяет выделить из общей БГД конкретную прикладную составляющую (как многомасштабную структуру геоданных).

Практическая реализация МВБГД предусматривается в виде создания специальной программной надстройки над стандартной реляционной системой управления базами данных.

## 4 Заключение

В работе представлена структура концептуальной схемы и основные проектные решения для построения образца БГД, обеспечивающей поддержку согласованного мультимасштабного и многовариантного представления геообъектов. В контексте данной задачи также рассмотрены сущность вариантности и особенности вариантных преобразований.

Интеграция в контексте общей, единой многовариантной БГД различных вариантов геоданных позволит обеспечить непосредственный переход между разными представлениями ситуации на местности или осуществлять совместное их использование при решении некоторых задач, находящихся на стыке разных предметных областей, а также обеспечить синхронизацию и согласование изменений геоданных.

Поддержка вариантности геоданных может дать БГД новые возможности для анализа (в том числе выявление неявных зависимостей в геоданных) и решения задач (как типовых, так и новых) за счет привлечения специализированных геоомделей, а также способствовать оперативной актуализации данных. В перспективе предполагается добавить в модель БГД поддержку темпорального аспекта.

## Литература

1. Дулин С. К., Никишин Д. А. Подходы к интеграции прикладных концептуальных схем в составе унифицированной геоонтологии // Системы и средства информатики, 2020. Т. 30. № 2. С. 68–77.
2. Никишин Д. А. Вариант концептуальной схемы базы геоданных с поддержкой обратимой генерализационной связанности моделей геообъектов // Системы и средства информатики, 2021. Т. 31. № 2. С. 119–128.
3. Pandey V., Kipf A., Vorona D., Mühlbauer T., Neumann T., Kemper A. High-performance geospatial analytics in HyPerSpace // Conference (International) on Management of Data Proceedings. — New York, NY, USA: ACM, 2016. P. 2145–2148. doi: 10.1145/2882903.2899412.
4. Wang S., Zhong Y., Lu H., Wang E., Yun W., Cai W. Geospatial big data analytics engine for spark // 6th ACM SIGSPATIAL Workshop on Analytics for Big Geospatial Data Proceedings. — New York, NY, USA: ACM, 2017. P. 42–45. doi: 10.1145/3150919.3150923.
5. Vollmer J. O., Trapp M., Schumann H., Döllner J. Hierarchical spatial aggregation for level-of-detail visualization of 3D thematic data // ACM Trans. Spatial Algorithms Systems, 2018. Vol. 4. Iss. 3. Art. 9. 23 p. doi: 10.1145/3234506.
6. Никишин Д. А. Структура и особенности генерализации в контексте функционирования темпоральной базы геоданных // Системы и средства информатики, 2020. Т. 30. № 1. С. 147–159.
7. Codd E. F. Extending the database relational model to capture more meaning // ACM T. Database Syst., 1979. Vol. 4. Iss. 4. P. 395–434.
8. Важенин И. А., Гусачев М. С., Карманов Д. В. Технология хранения мультимасштабных пространственных данных в объектно-ориентированной базе геоданных ArcGIS // Информация и космос, 2013. № 1. С. 48–52.
9. Jilani M., Corcoran P., Bertolotto M. Multi-granular street network representation towards quality assessment of OpenStreetMap Data // 6th ACM SIGSPATIAL Workshop (International) on Computational Transportation Science Proceedings. — New York, NY, USA: ACM, 2013. P. 19–24. doi: 10.1145/2533828.2533833.
10. Jilani M., Corcoran P., Bertolotto M. Automated highway tag assessment of OpenStreetMap road networks // 22nd ACM SIGSPATIAL Conference (Inter-

- national) on Advances in Geographic Information Systems Proceedings. — New York, NY, USA: ACM, 2014. P. 449–452. doi: 10.1145/2666310.2666476.
11. Wang H., Chen X., Polys N., Sforza P. A Web3D forest geo-visualization and user interface evaluation // 22nd International Conference on 3D Web Technology Proceedings. — New York, NY, USA: ACM, 2017. Art. No. 9. 9 p. doi: 10.1145/3055624.3075956.
  12. Никитин Д. А. Подход к совершенствованию концептуальных схем баз геоданных посредством моделей для пространственно-логического связывания геообъектов // Системы и средства информатики, 2021. Т. 31. № 1. С. 157–167.

Поступила в редакцию 21.03.22

---

## AN APPROACH TO THE DEVELOPMENT OF A CONCEPTUAL SCHEME OF A GEODATA DATABASE WITH SUPPORT FOR A MULTIVARIATE REPRESENTATION OF GEOOBJECTS

D. A. Nikishin

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

**Abstract:** The structure of the conceptual scheme and the main design solutions for constructing a variant of the geodata database that provides support for a consistent multiscale and multivariate representation of geo objects are presented. The essence of the variation of geomodels is considered and the typology of relationships between variant representations of geodata is highlighted. Combining the options for representing geodata in the context of a common geodatabase will allow for direct access to them and, on this basis, implement new methods for solving problems of data analysis and processing as well as ensure synchronization and coordination of changes in geodata.

**Keywords:** geodata; geodata databases; conceptual scheme of the geodata database; modeling of geo objects; variation of geomodels

**DOI:** 10.14357/08696527220212

### References

1. Dulin, S. K., and D. A. Nikishin. 2020. Podkhody k integratsii prikladnykh kontseptual'nykh skhem v sostave unifitsirovannoy geoontologii [Approaches to the integration of the application conceptual schemas in the unified geoontology]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(2):68–77.
2. Nikishin, D. A. 2020. Variant kontseptual'noy skhemy bazy geodannykh s podderzhkoy obratimoy generalizatsionnoy svyazannosti modeley geoob"ektov [A variant of the conceptual schema of the geodata database with support for reversible generalization connectivity of geo object models]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 31(2):119–128.

3. Pandey, V., A. Kipf, D. Vorona, T. Mühlbauer, T. Neumann, and A. Kemper. 2016. High-performance geospatial analytics in HyPerSpace. *Conference (International) on Management of Data Proceedings*. New York, NY: ACM. 2145–2148. doi: 10.1145/2882903.2899412.
4. Wang, S., Y. Zhong, H. Lu, E. Wang, W. Yun, and W. Cai. 2017. Geospatial big data analytics engine for spark. *6th ACM SIGSPATIAL Workshop on Analytics for Big Geospatial Data Proceedings*. New York, NY: ACM. 42–45. doi: 10.1145/3150919.3150923.
5. Vollmer, J. O., M. Trapp, H. Schumann, and J. Döllner. 2018. Hierarchical spatial aggregation for level-of-detail visualization of 3D thematic data. *ACM Trans. Spatial Algorithms Systems* 4(3):9. 23 p. doi: 10.1145/3234506.
6. Nikishin, D. A. 2020. Struktura i osobennosti generalizatsii v kontekste funktsionirovaniya temporal'noy bazy geodannykh [Structure and features of generalization in the context of functioning of a temporal geodatabase]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(1):147–159.
7. Codd, E. F. 1979. Extending the database relational model to capture more meaning. *ACM T. Database Syst.* 4(4):395–434.
8. Vazhenin, I. A., M. S. Gusachev, and D. V. Karmanov. 2013. Tekhnologiya khra-neniya mul'timasshtabnykh prostranstvennykh dannykh v ob"ektno-orientirovannoy baze geodannykh ArcGIS [Technology for storing multiscale spatial data in an object-oriented ArcGIS geodatabase]. *Informatsiya i kosmos* [Information and Space] 1:48–52.
9. Jilani, M., P. Corcoran, and M. Bertolotto. 2013. Multi-granular street network representation towards quality assessment of OpenStreetMap Data. *6th Workshop (International) on Computational Transportation Science Proceedings*. New York, NY: ACM. 19–24. doi: 10.1145/2533828.2533833.
10. Jilani, M., P. Corcoran, and M. Bertolotto. 2014. Automated highway tag assessment of OpenStreetMap road networks. *22nd Conference (International) on Advances in Geographic Information Systems Proceedings*. New York, NY: ACM. 449–452. doi: 10.1145/2666310.2666476.
11. Wang, H., X. Chen, N. Polys, and P. Sforza. 2017. A Web3D forest geo-visualization and user interface evaluation. *22nd Conference (International) on 3D Web Technology Proceedings*. New York, NY: ACM. Art. No. 9. 9 p. doi: 10.1145/3055624.3075956.
12. Nikishin, D. A. 2021. Podkhod k sovershenstvovaniyu kontseptual'nykh skhem baz geodannykh posredstvom modeley dlya prostranstvenno-logicheskogo svyazyvaniya geo-ob"ektov [An approach to improving the conceptual schemes of geodata by means of models for spatial and logical linking of geoobjects]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 31(1):157–167.

Received March 21, 2022

## Contributor

**Nikishin Dmitry A.** (b. 1976)—Candidate of Science (PhD) in technology, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; dmnikishin@mail.ru

## ПОДХОД К СВЯЗЫВАНИЮ ЗАПИСЕЙ В ТЕХНОЛОГИИ ПОДДЕРЖКИ КОНКРЕТНО-ИСТОРИЧЕСКИХ ИССЛЕДОВАНИЙ, ОСНОВАННЫЙ НА НЕЧЕТКИХ МНОЖЕСТВАХ

*И. М. Адамович<sup>1</sup>, О. И. Волков<sup>2</sup>*

**Аннотация:** Статья посвящена дальнейшему развитию распределенной технологии поддержки конкретно-исторических исследований (ПКИИ), основанной на принципах краудсорсинга и ориентированной на широкий круг не относящихся к профессиональным историкам и биографам пользователей. Развитие осуществляется за счет модификации алгоритма автоматического связывания записей при обработке номинативных документов семейной структуры, опирающегося на метод восстановления истории семей, успешно применяемый в исторической демографии, и структуру графов родственных связей в семье. Модификация заключается в опоре на теорию нечетких множеств и использовании мер сходства для принятия решения о связывании записей в случае, когда алгоритмы метода восстановления истории семей не позволяют сделать однозначный вывод о связанности. Описывается методика построения функций принадлежности и обосновывается использование меры сходства, основанной на мерах сходства по Лукасевичу и по Заде.

**Ключевые слова:** конкретно-историческое исследование; распределенная технология; связывание записей; источники семейной структуры; нечеткие множества

**DOI:** 10.14357/08696527220213

### 1 Введение

Поддержка конкретно-исторических исследований стала одной из актуальных задач современности, что обусловлено вовлечением в исследовательский процесс не только членов профессионального исторического сообщества, но и самых широких слоев непрофессионалов в связи со все возрастающим интересом к частной, семейной истории [1].

В [2, 3] описана разработанная в ФИЦ ИУ РАН распределенная технология ПКИИ, основанная на принципах краудсорсинга (мобилизации ресурсов широкого круга добровольцев посредством информационных технологий), и создан ряд программных средств ее поддержки.

---

<sup>1</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Adam@amsd.com

<sup>2</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Volkov@amsd.com

Вся совокупность исторических источников, содержащих основной объем исходных данных для конкретно-исторического исследования, делится на три основные группы: номинативные (содержащие персональную информацию с указанием имени собственного), статистические и нарративные (повествовательные) источники [4]. Наибольший интерес для исследователя представляют номинативные источники, подразделяющиеся на источники персональной структуры и источники семейной структуры [5].

При обработке номинативных источников одной из важнейших и сложнейших задач является поиск соответствий и взаимоувязывание данных. Связывание записей (Record Linkage) — классическая проблема в исторической информатике. Несмотря на наличие значительных негативных особенностей номинативных исторических источников, задача автоматизированного связывания записей для информации, полученной из таких источников, довольно успешно решается в рамках метода восстановления истории семей (ВИС), используемого в исторической демографии [4].

В задачах связывания записей используются детерминированные и вероятностные алгоритмы. Детерминированные алгоритмы основаны на полном совпадении атрибутов записей, и для целей ВИС они применяться практически не могут. Более предпочтительной альтернативой являются вероятностные алгоритмы. В этом случае для атрибутов задаются весовые коэффициенты, определяющие влияние на итоговую оценку вероятности того, что оцениваемые записи относятся к одному и тому же объекту. Пары записей с вероятностями выше некоторого порога считаются связанными, а пары с вероятностями ниже другого порога считаются несвязанными. Пары, которые попадают между этими двумя пороговыми значениями, считаются кандидатами на связывание и обрабатываются отдельно (например, вручную).

Опыт реальных исследований показывает, что наибольших успехов в области автоматизации связывания записей исторического регистра в рамках ВИС удастся добиться при обработке метрических записей и подобных персональных источников демографических событий, но при обработке источников семейной структуры возникают проблемы [6]. Из этого следует, что имеется потребность в механизмах автоматического связывания для семейных структур в ситуации, когда стандартные алгоритмы ВИС не смогли установить связь ни для одного члена семьи.

## 2 Структурный подход к связыванию семей

Связывание записей в технологии ПКИИ осуществляется сравнением персон, информация о которых уже зафиксирована в семантической сети технологии, с персонами, упоминаемыми в некотором номинативном историческом документе. В [7] описан подход к связыванию записей в ситуации, когда при обработке номинативных источников семейной структуры автоматическими алгоритмами связывания метод ВИС не дает уверенного ответа о связанности, т. е. нет пар

с вероятностью связанности, достаточной, чтобы считать их определенно связанными, но имеются пары с вероятностью связанности выше порога определенной несвязанности. Для этих пар будем говорить о наличии потенциальной связи (Р-связи).

Поскольку родственные связи образуют направленный граф с именованными связями и вершинами, соответствующими персонам, алгоритм опирается на оценку сходства структур таких графов. Для сравнения структур алгоритм учитывает:

- (1) объем потенциальных совпадений в структурах графов (доля узлов с Р-связями);
- (2) объем признанных на основании экспертной оценки естественными (легкообъяснимыми) несовпадений в структурах графов (доля соответствующих узлов);
- (3) объем труднообъяснимых несовпадений в структурах графов (доля прочих узлов).

Формула вычисления вероятности признания графов сходными представляет собой взвешенную разность доли потенциальных совпадений и доли труднообъяснимых несовпадений. Весовые коэффициенты определяются по результатам испытаний алгоритма в реальных конкретно-исторических исследованиях. Рост объема совпадений увеличивает вероятность признания графов сходными, а рост объема труднообъяснимых несовпадений — уменьшает. При превышении с вычисленной по соответствующей формуле вероятностью некоторого порогового значения записи с Р-связями признаются связанными.

Несмотря на то что данный подход существенно дополняет и развивает технологию ПКИИ, он не свободен от ряда недостатков. Основные из них — неудовлетворительная точность некоторых критериев и субъективность в определении ряда параметров алгоритма.

Недостаточная точность вытекает из дихотомического подхода к определению естественности/труднообъяснимости несовпадений в структурах графов. В реальности мнения экспертов часто оцениваются по порядковой шкале [8].

Любое автоматическое принятие решения невозможно без задания некоторых пороговых значений исследователем исходя из его субъективных предпочтений и опыта. Но рассматриваемый алгоритм имеет неоправданно большое число таких задаваемых пользователем параметров:

- параметры алгоритмов ВИС, позволяющие считать записи связанными, потенциально связанными или несвязанными;
- весовые коэффициенты формулы вычисления вероятности признания графов сходными;
- пороговое значение вычисленной вероятности признания графов сходными, превышение которого позволяет признать записи с Р-связями связанными.

Из недостаточной точности и субъективности рассмотренного метода вытекает необходимость в построении нового алгоритма связывания записей в технологии ПККИ при обработке номинативных источников семейной структуры, опирающегося на алгоритмы ВИС и свободного от этих недостатков.

### 3 Подход на основе нечетких множеств

Предлагается новый подход к решению задачи связывания записей при обработке документов семейной структуры в случае, когда метод ВИС не дает однозначного ответа. Подход основан на нечетких множествах.

Под нечетким множеством  $A$  понимается совокупность упорядоченных пар, составленных из элементов  $x$  универсального множества  $X$  и соответствующих степеней принадлежности  $\mu_A(x)$ :  $A = \{(x, \mu_A(x)) | x \in X\}$ , где  $\mu_A(x)$  — функция принадлежности, принимающая, как правило, значения в диапазоне  $[0, 1]$  и указывающая, в какой степени (мере) элемент  $x$  принадлежит нечеткому множеству  $A$  [9].

Для нечетких множеств определено отношение сходства. Среди большого числа мер сходства наиболее широкое распространение получили:

(1) мера сходства по Лукасевичу:

$$m_L(\tilde{A}, \tilde{B}) = 1 - \frac{\sum_{i=1}^n |\mu_{\tilde{A}}(x_i) - \mu_{\tilde{B}}(x_i)|}{n};$$

(2) мера сходства по Заде:

$$m_Z(\tilde{A}, \tilde{B}) = \frac{\sum_{i=1}^n \min(\max(1 - \mu_{\tilde{A}}(x_i), \mu_{\tilde{B}}(x_i)), \max(\mu_{\tilde{A}}(x_i), 1 - \mu_{\tilde{B}}(x_i)))}{n}.$$

В приведенных формулах  $\tilde{A}(x_i)$  и  $\tilde{B}(x_i)$  — нечеткие множества, определенные на базовом множестве признаков  $x_i \in X$ ,  $i = \overline{1, n}$ , которые, в свою очередь, могут принимать как количественные (числовые), так и качественные (лингвистические) значения, обуславливающие применение той или иной формулы. В первом случае целесообразно воспользоваться мерой сходства по Лукасевичу, имеющей смысл расстояния между элементами, а во втором — мерой сходства по Заде, используемой при выполнении операций над логическими переменными и высказываниями.

В случае когда в описании ситуаций одновременно присутствуют и количественные, и качественные параметры, целесообразно применять комбинированную меру сходства

$$C_0(\tilde{A}, \tilde{B}) = \beta m_Z(\tilde{A}, \tilde{B}) + (1 - \beta) m_L(\tilde{A}, \tilde{B}),$$

в которой конкретные значения коэффициента  $\beta \in [0, 1]$  выбираются экспертами исходя из содержательного смысла используемых параметров [10].

#### 4 Представление семейной структуры как нечеткого множества

Имеются два множества: множество персон, относящихся к определенной семейной структуре, информация о которых уже зафиксирована в семантической сети технологии, и множество персон, упоминаемых в некотором номинативном историческом документе. На их основе формируются два нечетких множества с соответствием, определяемым связанностью элементов, определенной в результате выполнения процедур ВИС (Р-связностью), и структурой отношений родства.

Пусть  $A$  — множество узлов семантической сети, соответствующих составу семьи в предыдущий период в соответствии с ранее обработанными документами семейной структуры.

Пусть  $B$  — множество персон, упоминаемых в обрабатываемом документе семейной структуры, для которых уже в результате анализа документа выявлены отношения родства и для некоторых из них установлены Р-связи с соответствующими элементами множества  $A$  средствами ВИС. Подмножество  $B^0$  множества  $B$ , для элементов которого установлены Р-связи с элементами  $A$ , будем называть ядром множества  $B$ . Ему соответствует ядро  $A^0$  множества  $A$ . Элемент  $B^0$ , связанный Р-связью с элементом  $a \in A^0$ , обозначим  $p(a)$ . Вероятность соответствия, определенная методом ВИС для  $a_i \in A^0$ , обозначим как  $v_i$ .

Как показано выше, дихотомическая оценка несовпадений структур слишком груба и субъективна. Целесообразно использовать экспертную оценку по порядковой шкале. Существует ряд методов построения по экспертным оценкам функции принадлежности нечеткого множества. В целях снижения субъективного влияния эксперта на результаты построения функций принадлежности, которое в ряде случаев приводит к сдвиганию оценки объектов в направлении концов оценочной шкалы, в рамках косвенных методов построения функций принадлежности используются: метод статистических данных; метод парных сравнений; метод экспертных оценок; параметрический метод; метод интервальных оценок; метод ранговых оценок [11]. Целесообразно строить алгоритм построения функции принадлежности на базе модифицированного метода Ягера и модифицированного метода анализа иерархий Саати [12]. Значение построенной функции принадлежности для узла  $x_i$  одного графа, не имеющего аналога в другом графе, т. е. не принадлежащего ядру, отражающее оценку степени объяснимости несоответствия для данного узла, обозначим как  $o_i$ .

Обозначим  $\bar{A} = A \cup B \setminus B^0$ ,  $\bar{B} = B \cup A \setminus A^0$ . Между элементами множеств  $\bar{A}$  и  $\bar{B}$  существует взаимно однозначное соответствие (биекция)  $f: \bar{A} \leftrightarrow \bar{B}$ :

$$f(a_i \in \overline{A}) = \begin{cases} p(a_i), & a_i \in A^0; \\ a_i, & a_i \in A \setminus A^0 \cup B \setminus B^0. \end{cases}$$

Это позволяет рассмотреть два нечетких множества  $\tilde{A}$  и  $\tilde{B}$ , определенных на универсальном множестве  $\overline{B}$  и с функциями принадлежности  $\mu_{\tilde{A}}$  и  $\mu_{\tilde{B}}$  соответственно:

$$\mu_{\tilde{A}}(x_i) = \begin{cases} 1 & |f^{-1}(x_i) \in A^0; \\ 1 & |x_i \in A \setminus A^0; \\ o_i & |x_i \in B \setminus B^0, \end{cases} \quad \mu_{\tilde{B}}(x_i) = \begin{cases} v_i & |x_i \in B^0; \\ 1 & |x_i \in B \setminus B^0; \\ o_i & |x_i \in A \setminus A^0. \end{cases}$$

Функция  $\mu_{\tilde{A}}$  отражает тот факт, что множество  $A$  является эталонным, а элементы  $B$ , отсутствующие в  $A$ , относятся к нему с вероятностями, определенными на базе экспертных оценок степени объяснимости несоответствия. Функция  $\mu_{\tilde{B}}$  отражает тот факт, что элементы ядра множества  $B$  соотносятся с элементами множества  $A$  с вероятностями, определенными методом ВИС, прочие элементы множества  $B$  относятся к нему безусловно, а элементы  $A$ , отсутствующие в  $B$ , относятся к нему с вероятностями, определенными на базе экспертных оценок степени объяснимости несоответствия.

## 5 Решение о связывании записей

Решение о связывании или несвязывании записей номинативных источников семейной структуры целесообразно принимать, оценивая значение меры сходства двух нечетких множеств  $\tilde{A}$  и  $\tilde{B}$ , построенных описанным выше способом. Поскольку функции принадлежности строились на основании как экспертных данных, так и результатов расчетов, целесообразно применять комбинированную меру сходства

$$C_0(\tilde{A}, \tilde{B}) = \beta m_Z(\tilde{A}, \tilde{B}) + (1 - \beta)m_L(\tilde{A}, \tilde{B}),$$

где  $m_Z$  — мера сходства по Заде;  $m_L$  — мера сходства по Лукасевичу.

Коэффициент  $\beta$ , отражающий соотношение количественных и качественных параметров, может быть вычислен по формуле:

$$\beta = \frac{|A \setminus A^0| + |B \setminus B^0|}{|A| + |B| + |B \setminus B^0|}.$$

## 6 Выводы

Предложенный алгоритм позволяет значительно улучшить входящий в состав технологии ПКИИ механизм автоматизации проведения такой важной составляющей конкретно-исторического исследования, как идентификация персон в номинативных исторических документах семейной структуры, за счет снижения

влияния субъективного мнения исследователя на результат и увеличения точности результата за счет отказа от дихотомии при классификации несовпадений структур отношений родства.

Дальнейшее развитие предложенного алгоритма возможно за счет использования экспертных систем при формировании оценок несовпадений этих структур.

## Литература

1. *Помникова А. Ю.* Семейная история в дискурсивном пространстве // Вестник Мининского университета, 2019. Т. 7. № 1. Ст. 9. 22 с.
2. *Адамович И. М., Волков О. И.* Технология распределенного автоматизированного анализа исторических текстов // Системы и средства информатики, 2016. Т. 26. № 3. С. 148–161.
3. *Адамович И. М., Волков О. И.* Единая технология поддержки конкретно-исторических исследований // Системы и средства информатики, 2019. Т. 29. № 1. С. 194–205.
4. *Антонов Д. Н.* Восстановление истории семей: метод, источники, анализ: Дис. . . . канд. ист. наук. — М.: РГГУ, 2000. 290 с.
5. *Торвальдсен Г.* Номинативные источники в контексте всемирной истории переписей: Россия и Запад // Известия Уральского федерального университета. Сер. 2: Гуманитарные науки, 2016. Т. 18. № 3. С. 9–28.
6. *Юрченко Н. Л.* Некоторые проблемы использования ревизских сказок как источника по исторической демографии // Вспомогательные исторические дисциплины, 1993. Т. XXIV. С. 183–189.
7. *Адамович И. М., Волков О. И.* Структурный подход к связыванию записей в технологии поддержки конкретно-исторических исследований // Системы и средства информатики, 2022. Т. 32. № 1. С. 94–103.
8. *Орлов А. И.* Экспертные оценки. — М., 2002. 31 с.
9. *Коньшева Л. К., Серова Т. А.* Элементы теории нечетких множеств. — Екатеринбург: Рос. гос. проф.-пед. ун-т, 2007. 129 с.
10. *Целых А. А.* Метод оценки разрыва в компетенциях выпускника вуза // Фундаментальные исследования, 2013. № 11-8. С. 1602–1606.
11. *Каид В. А. А.* Методы построения функций принадлежности нечетких множеств // Известия ЮФУ. Технические науки, 2013. № 2(139). С. 144–153.
12. *Гриценко С. А., Храмов В. Ю.* Косвенные методы построения функций принадлежности систем поддержки принятия решений с нечеткой логикой // Вестник ВГУ. Сер.: Системный анализ и информационные технологии, 2017. № 4. С. 114–125.

*Поступила в редакцию 28.02.22*

## APPROACH TO RECORD LINKING IN TECHNOLOGY OF CONCRETE HISTORICAL INVESTIGATION SUPPORT BASED ON FUZZY SETS

*I. M. Adamovich and O. I. Volkov*

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

**Abstract:** The article is devoted to the further development of the distributed technology of concrete historical investigation support based on the principles of crowdsourcing and focused on a wide range of users which are nonprofessional historians and biographers. The development is carried out by modifying the algorithm for automatic record linking when processing nominative documents of family structure based on the Family reconstitution method, successfully used in historical demography, and the structure of family relationship graphs. The modification consists in relying on the theory of fuzzy sets and using the similarity measures to make a decision about records linking in the case when the algorithms of the family reconstitution method do not allow an unambiguous conclusion about the linkage. The article describes the methodology for constructing membership functions and substantiates the use of a similarity measure based on the Lukasiewicz and Zade similarity measures.

**Keywords:** concrete historical investigation; distributed technology; record linking; sources of family structure; fuzzy sets

**DOI:** 10.14357/08696527220213

### References

1. Pomnikova, A. Yu. 2019. Semeynaya istoriya v diskursivnom prostranstve [Family stories in different types of discourse]. *Vestnik of Minin University* 7(1):9. 22 p.
2. Adamovich, I. M., and O. I. Volkov. 2016. Tekhnologiya raspredelennogo avtomatizirovannogo analiza istoricheskikh tekstov [The distributed automated technology of historical texts analysis]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 26(3):148–161.
3. Adamovich, I. M., and O. I. Volkov. 2019. Edinaya tekhnologiya podderzhki konkretno-istoricheskikh issledovaniy [Unified technology of concrete historical research support]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(1):194–205.
4. Antonov, D. N. 2000. Vosstanovlenie istorii semey: metod, istochniki, analiz [Family reconstitution: Method, sources, and analysis]. Moscow: RGGU. PhD Thesis. 290 p.
5. Thorvaldsen, G. 2016. Nominativnye istochniki v kontekste vseмирной istorii perepisey: Rossiya i Zapad [Nominative data and global census history: Russia and the West]. *Izvestiya Ural'skogo federal'nogo universiteta. Ser. 2: Gumanitarnye nauki* [Izvestia. Ural Federal University Journal. Ser. 2. Humanities and Arts] 18(3):9–28.

6. Yurchenko, N. L. 1993. Nekotorye problemy ispol'zovaniya revizskikh skazok kak istochnika po istoricheskoy demografii [Some problems of using census lists as a source for historical demography]. *Vspomogatel'nye istoricheskie distsipliny* [Auxiliary Historical Disciplines] XXIV:183–189.
7. Adamovich, I. M., and O. I. Volkov. 2022. Strukturnyy podkhod k svyazyvaniyu zapisey v tekhnologii podderzhki konkretno-istoricheskikh issledovaniy [Structural approach to record linking in technology of concrete historical investigation support]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 32(1):94–103.
8. Orlov, A. I. 2002. *Ekspertnye otsenki* [Expert assessments]. Moscow. 31 p. Available at: <http://www.aup.ru/books/m154/> (accessed March 10, 2022).
9. Konysheva, L. K., and T. A. Serova. 2007. *Elementy teorii nechetkikh mnozhestv* [Fuzzy-set theory elements]. Ekaterinburg: RGPPU. 129 p.
10. Tselykh, A. A. 2013. Metod otsenki razryva v kompetentsiyakh vypusknika vuza [Method for competency gap analysis of a university graduate]. *Fundamental'nye issledovaniya* [Fundamental Research] 11-8:1602–1606.
11. Qaid, W. A. A. 2013. Metody postroeniya funktsiy prinadlezhnosti nechetkikh mnozhestv [Methods construction membership function of fuzzy sets]. *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU: Engineering Sciences] 2(139):144–153.
12. Gritsenko, S. A., and V. U. Khramov. 2017. Kosvennye metody postroeniya funktsiy prinadlezhnosti sistem podderzhki prinyatiya resheniy s nechetkoy logikoy [Indirect methods of constructing the functions of accessories of decision support systems with fuzzy logic]. *Vestnik VGU. Ser.: Sistemnyy analiz i informatsionnye tekhnologii* [Proceedings of Voronezh State University. Ser.: Systems Analysis and Information Technologies] 4:114–125.

*Received February 28, 2022*

## Contributors

**Adamovich Igor M.** (b. 1934) — Candidate of Science (PhD) in technology, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; [Adam@amsd.com](mailto:Adam@amsd.com)

**Volkov Oleg I.** (b. 1964) — leading programmer, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; [Volkov@amsd.com](mailto:Volkov@amsd.com)

## S-МОДЕЛИРОВАНИЕ И ЦИФРОВЫЕ ТЕХНОЛОГИИ ЭКОНОМИЧЕСКОЙ ДЕЯТЕЛЬНОСТИ

*В. Д. Ильин*

**Аннотация:** Рассмотрено применение теории символьного моделирования произвольных объектов (s-моделирования) для идентификации экономических процессов при решении задач управления ими в цифровой среде. S-модели финансовых, торговых и других процессов экономической деятельности представлены совокупностью самообучающихся цифровых двойников (digital twins). Для построения цифрового двойника используются значения параметров, характеризующих поведение моделируемого объекта. С момента своего создания цифровой двойник накапливает знания о моделируемом объекте, обмениваясь с ним сообщениями и обновляясь. Цифровая технология экономической деятельности определенного вида рассматривается как совокупность методов, средств и правил реализации процессов этой деятельности посредством онлайн-сервисов, работающих в среде цифровых двойников. Приведены примеры обновленных технологий нормализованного экономического механизма (НЭМ).

**Ключевые слова:** символьное моделирование; s-моделирование; цифровая технология; экономический процесс; нормализованный экономический механизм

**DOI:** 10.14357/08696527220214

### 1 Введение

Так сложилось, что о ведении хозяйства в масштабах страны и мира часто рассуждают вовсе не инженеры и ученые, занимающиеся исследованиями природных и изобретенных объектов (из энциклопедии Britannica: «Economics — social science that seeks to analyze and describe the production, distribution, and consumption of wealth (Mark Blaug)»). Это не могло не отразиться на качестве экономических концепций и теорий. Нечеткость определений, псевдоматематизация, оперирование неизмеримыми величинами в «количественных моделях» [1, 2], отсутствие явно указанного целевого назначения и ограничений на применимость модели, неучет реальных информационных технологий — характерные признаки экономических произведений.

О хозяйственной деятельности авторы таких произведений рассуждают, как о стихийных явлениях («кризис разразился» и т. п.) и азартных играх («игроки», «игра на повышение» и пр.).

---

<sup>1</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук, vdilyin@yandex.ru

**Об измеримости и количественных моделях.** Есть вещи, которые научились измерять: ввели единицы измерения, изобрели и построили средства измерения. Умеем измерять длину, массу, температуру и многое другое. Но еще больше вещей, которые измерять пока не умеем.

Исследователи и инженеры на основе количественных моделей формулируют задачи, чтобы, задавая значения переменных, считающихся известными, вычислять значения переменных, считающихся неизвестными. Насколько адекватна количественная модель, выясняют, сопоставляя вычисленные значения с измеренными.

Но изобретателей «количественных теорий денег» [1, 2], видимо, не смущал факт неизмеримости величин, которые включены в их «теории». Толкуя об общем числе сделок, творцы этих «теорий» не различают сделки по видам (подробнее см. [3, разд. 1.4]). Кроме того, в этих «теориях» деньги (средство измерения стоимости товаров) тоже являются товаром, а сделки купли-продажи валюты и различных заменителей денег (так называемых ценных бумаг) неотличимы от сделок купли-продажи реальных товаров.

**О правилах экономического поведения.** Надо отдать должное ротшильдскому клану идеологов финансового сектора экономики. За пару веков глобальная кооперация финансовых «наперсточников» добилась решающего влияния на принятие ключевых решений, связанных с правилами экономического поведения. Представители и идейные последователи этого клана — на ключевых позициях в финансовом секторе, университетах и экспертных группах, готовящих решения для государственных руководителей.

Масса экономических изданий (бумажных и электронных), теле- и радиокomentаторов, говорящих о состоянии рынков и пр., круглосуточно осуществляют финансовое «просвещение» населения.

**О нормализации экономического механизма.** «Чем агрессивнее будут вводиться экономические санкции против России, тем быстрее нормализуется экономический механизм нашей страны. В частности, при отказе России от доллара США (как средства расчетов за поставки нефти) американское экономическое доминирование будет убывать еще интенсивнее. А переход к поливалютному рынку (расчетам в валюте, выбранной участниками торговой сделки) станет решительным шагом к нормализации экономического механизма» [4].

В статье о модели нормализованной экономики (НЭК-модели), опубликованной в 2009 г., содержится предложение о принципиальном изменении подхода к формированию и использованию государственного золотовалютного резерва страны. «Нынешний государственный золотовалютный резерв России уязвим в той мере, в какой неадекватна существующая концепция денег: изменение товарной емкости валют и цены золота — вне области допустимого управления держателя резерва, а средства резерва не работают в экономике России. Целесообразно запретить законом вывод денежных средств государственного резерва и фондов социальной защиты из НЭК-системы той страны, где они были порождены и аккумулированы» [5, с. 134].

Процитированные утверждения автора сохраняют актуальность и в наши дни. Следует заметить, что в последние 3–4 года прирастает поддержка концепции нормализации экономики и среди экономистов-исследователей [6–11].

**Запись формул и выделение фрагментов текста.** Для выделения определений, замечаний и примеров используются средства языка TSM-комплекса (TSM: textual symbolic modeling), разработанного для формализованного описания текстовых s-моделей сообщений<sup>1</sup>.

В статье применены следующие средства выделения фрагментов текста:

□ (фрагмент описания) □ ≈ утверждение (определение, аксиома и др.) (здесь и далее символ ≈ заменяет слово «означает»);

◇ (фрагмент описания) ◇ ≈ замечание;

○ (фрагмент описания) ○ ≈ пример.

Курсивом выделены первые вхождения названий понятий и фрагменты описания, к которым автор хочет привлечь внимание.

**Обсуждаемые результаты.** В статье рассмотрены актуальные примеры применения *теории s-моделирования* при цифровизации экономических процессов.

## 2 S-моделирование процессов нормализованной экономической деятельности

□ *Нормализованный экономический механизм* — рыночный экономический механизм, технологии работы которого реализуют задачи комплексов НЭМ посредством онлайн-сервисов, функционирующих в среде цифровых двойников. Представляет собой (основанную на современных цифровых технологиях) человеко-технологическую систему, связанную отношениями координации и подчинения с государственным механизмом. □

◇ В НЭМ стремление участников экономической деятельности (эд-участников) к получению экстерриториальной выгоды поддерживается в тех случаях, когда сделка (купли-продажи, инвестирования или др.) не противоречит решению задач защиты и развития потенциала страны ◇ [12].

Комплекс технологий нормализованного товарно-денежного обращения включает технологии формирования и реализации договорных отношений в цифровой среде, платежного и товарного кредитования, онлайн-банкинга на основе банков-провайдеров, корпоративных (КЭБов) и персональных электронных банков (ПЭБов).

Ядром комплекса документирования товарно-денежного обращения и имущественных статусов эд-участников является система имущественных статусов.

□ *Система имущественных статусов (ис-система)* — система документального представления денежной и неденежной составляющих имущественных

---

<sup>1</sup>Ильин В. Д. Символьное моделирование // Большая российская энциклопедия — электронная версия. [http://dev.bigenc.ru/technology\\_and\\_technique/text/4010980](http://dev.bigenc.ru/technology_and_technique/text/4010980).

статусов *участников экономической деятельности*. Денежная составляющая выражена значениями сумм нормализованных денег в разделах *уникальных счетов эд-участников (ис-счетов)*. Неденежная — документами, подтверждающими право собственности на недвижимость, транспорт и другое имущество, которое при необходимости может рассматриваться как залоговое. □

□ *Ис-счет* — уникальный банковский счет, размещенный в *персональном электронном банке* физического лица или *корпоративном электронном банке* юридического лица. Владелец ис-счета идентифицируется по ИНН. □

◇ Никакое изменение сумм ис-счета не может быть осуществлено без документируемого зашифрованного подтверждения владельца ис-счета (а в неординарных ситуациях — после документируемого зашифрованного подтверждения определенного законом государственного органа). ◇

В НЭМ это реализуется на программно-аппаратном уровне. Изменение состояния ис-счета может быть реализовано только после получения онлайн-удостоверения допустимости операции над ис-счетом. Удостоверение выдает специальный онлайн-сервис банка-регулятора. Основанием для получения удостоверения служат данные спецификации совершаемой операции (○ платеж при сделке купли-продажи, уплате налога или др. ○). Таким способом исключаются возможности нарушения установленных законом правил товарно-денежного обращения (недопустимые экономические сделки, неуплата установленных законом государственных сборов и др.).

□ *Нормализованные деньги (нд)* — универсальное электронное средство количественного документирования имущественных отношений, удостоверяемых государством. Предназначено для представления стоимости товаров и имущественных статусов эд-участников, для оплаты товаров, инвестирования и накопления богатства. Нд-суммы представлены в ис-счетах записями, отражающими денежно выраженные величины имущественного права и имущественных обязательств. Для представления значений нд-сумм используются вещественные числа с соответствующим знаком (минус — для сумм, которые предстоит вернуть; плюс — для полученных сумм). □

Универсальность нд выражается в применимости во всех внутренних и внешних экономических сделках, разрешенных законом государства, под юрисдикцией которого функционирует экономическая система.

*Банковская система НЭМ (нормализованная банковская система)* включает ПЭБы, КЭБы, банки-провайдеры и банк-регулятор, под управлением которого функционируют все другие банки.

*Банк-регулятор* располагает сетью серверов, размещенных на территории страны, под юрисдикцией которой находится экономическая система.

*Банки-провайдеры* — коммерческие предприятия, учрежденные юридическими лицами (или объединениями юридических лиц, объединениями физических лиц, объединениями юридических и физических лиц), занимающимися производством товаров и/или их реализацией. Банк-провайдер располагает объединенными в сеть серверами, предназначенными для обслуживания запросов

от ПЭБов и КЭБов клиентов и взаимодействия с серверами банка-регулятора.

Персональные и коммерческие электронные банки (как программно-аппаратные средства) — это, как правило, переносные устройства типа планшетов и смартфонов, наделенные надежно защищенными коммуникационными функциями. В ПЭБах и КЭБах в зашифрованной форме хранятся оригиналы ис-счетов и документов по сделкам.

Программно реализуемые записи в файлах ис-счетов имеют право инициировать только владельцы ис-счетов, совершающие сделку. Запросы на обслуживание сделки (удостоверение состояния ис-счета и сохранение обновленной копии ис-счета) получают банки-провайдеры, сервисами которых пользуются совершающие сделку эд-участники [13].

*Технологии электронной долговой торговли и денежной эмиссии* [14]. При электронной долговой торговле все выгоды и риски разделяют продавец и покупатель, а стоимость кредита упакована в стоимость покупки.

Эта технология обеспечивает оформление отсроченной части оплаты товара как долг покупателя продавцу. График выплаты долга фиксируется в договоре. В том же договоре указываются штрафные санкции за нарушение графика выплат долга и ненадлежащее качество проданного товара.

Применительно к долгам, образовавшимся в результате продажи приоритетных товаров, могут действовать правила досрочного возврата долгов продавцам из средств банка-регулятора. Делается это по запросам продавцов в долговой отдел банка-регулятора, направляемых с использованием онлайн-сервиса. При положительном решении на ис-счет продавца переводится сумма долга (полностью или частично), а после этого покупатель выплачивает эту сумму не продавцу, а долговому отделу банка-регулятора. О возможности такой схемы возврата долга покупатель извещается продавцом до реализации сделки.

◇ Досрочным возвратом долгов продавцам приоритетных товаров банк-регулятор реализует текущую государственную программу поддержки производства и продажи таких товаров. ◇

□ *Долговая денежная эмиссия* выполняется банком-регулятором только тогда, когда сумма долгов, возвращенных покупателями долговому отделу, меньше очередной долговой суммы, запрошенной для возврата продавцу. Эмитируемая сумма равна разности запрошенной долговой суммы и суммы на счету долгового отдела. □

◇ Долговая денежная эмиссия служит средством государственного регулирования общей суммы денег в экономической системе. Таким способом осуществляется государственное регулирование товарной емкости денег. ◇

□ *Назначенный платеж* — технологически гарантированная оплата заказанных товаров, правила реализации которой жестко связаны с правилами исполнения заказа. Сумма назначенного платежа переводится на счет получателя сразу после того, как заказчик подтвердил исполнение заказа. В случае неисполнения заказа назначенный платеж отменяется. □

Назначенный платеж не может быть сформирован, если заказ содержит товары, не отнесенные к допустимым. Контроль осуществляется цифровыми двойниками, обеспечивающими функционирование сервисов контроля допустимости заказа. Реализация технологии обеспечивается онлайн-сервисами банков, функционирующими в среде цифровых двойников, которые моделируют состояние платежной системы и процесса исполнения заказа [15].

### 3 Заключение

Представлены основные направления применения методологии s-моделирования для идентификации экономических процессов, реализуемых в современной цифровой среде.

Приведены уточненные определения и краткие характеристики актуальных технологий НЭМ.

### Литература

1. *Friedman M.* The quantity theory of money: A restatement // *Studies in the quantity theory of money* / Ed. M. Friedman. — Chicago, IL, USA: University of Chicago Press, 1956. P. 1–21.
2. *Friedman M., Bordo M.* The optimum quantity of money. — New Brunswick, NJ, USA: Aldine Transaction, 2006. 305 p.
3. *Ильин В. Д.* S-модель нормализованной экономической системы. — М.: ИПИ РАН, 2010. 103 с.
4. *Ильин В. Д.* Следствие экономических санкций // *Cnews клуб*, 26.03.2015. [https://club.cnews.ru/blogs/entry/sledstvie\\_ekonomicheskikh\\_sanktsij](https://club.cnews.ru/blogs/entry/sledstvie_ekonomicheskikh_sanktsij).
5. *Ильин В. Д.* Модель нормализованной экономики (НЭк-модель): основы концепции // *Управление большими системами*, 2009. Вып. 25. С. 116–138.
6. *Бауэр В. П., Еремин В. В., Сильвестров С. И., Смирнов В. В.* Экономическое моделирование процессов цифровой трансформации // *Ж. экономической теории*, 2019. Т. 16. № 3. С. 428–443. doi: 10.31063/2073-6517/2019.16-3.11.
7. *Орлова И. А., Мовчан М. И.* Развитие онлайн-сервисов коммерческих банков // *Инновационные технологии в машиностроении, образовании и экономике*, 2019. Т. 22. № 2. С. 208–211.
8. *Сильвестров С. Н., Бауэр В. П., Еремин В. В., Лапенкова Н. В.* О цифровой трансформации предприятия в контексте системной экономической теории // *Экономическая наука современной России*, 2020. № 2(89). С. 22–45. doi: 10.33293/1609-1442-2020-2(89)-22-45.
9. *Kikavets V. V., Tsaregradskaya Y. K.* The implementation of principles of the budget process in the financing of public procurement in the digital economy // *Artificial intelligence: Anthropogenic nature vs. social origin* / Eds. E. Popkova, B. Sergi. — Advances in intelligent systems and computing ser. — Springer, 2020. Vol. 1100. P. 28–35.
10. *Федосимов Б. А.* Введение налога с оборота и потребления для обеспечения равенства налогообложения различных субъектов финансово-хозяйственной деятельности

- и упрощения проведения налоговых изъятий // Russian Economic Bulletin, 2021. Т. 4. № 6. С. 299–304.
11. Бирюков А. Н. Бюджетирование как инструмент ипотечной кредитной программы банка // Научное обозрение: теория и практика, 2021. Т. 11. № 3(83). С. 794–805.
  12. Ilyin A. V., Ilyin V. D. The normalized economic mechanism in the digital environment // Int. J. Open Information Technologies, 2019. Vol. 7. Iss. 12. P. 77–83.
  13. Ильин В. Д. Нормализация банковских онлайн-сервисов // Системы и средства информатики, 2016. Т. 26. № 2. С. 147–157.
  14. Ilyin A. V., Ilyin V. D. E-trade with direct lending and normalized money // Agris on-line Papers in Economics and Informatics, 2015. Vol. 7. Iss. 4. P. 57–64.
  15. Ильин В. Д. Технология назначенных платежей в среде цифровых двойников // Системы и средства информатики, 2018. Т. 28. № 3. С. 227–235. doi: 10.14357/08696527180318.

Поступила в редакцию 15.03.22

---

---

## S-MODELING AND DIGITAL TECHNOLOGIES OF ECONOMIC ACTIVITY

V. D. Ilyin

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The application of the theory of symbolic modeling of arbitrary objects (s-modeling) to identify economic processes in solving problems of managing them in a digital environment is considered. S-models of financial, trade, and other processes of economic activity are represented by a set of self-learning digital twins. To construct a digital twin, the values of parameters characterizing the behavior of the simulated object are used. Since its creation, digital twin has been accumulating knowledge about the simulated object, exchanging messages with it and updating. Digital technology of economic activity of a certain type is considered as a set of methods, means, and rules for the implementation of the processes of this activity through online services operating in the environment of digital twins. The examples of updated technologies of the normalized economic mechanism are given.

**Keywords:** symbolic modeling; s-modeling; digital technology; economic process; normalized economic mechanism

**DOI:** 10.14357/08696527220214

## References

1. Friedman, M. 1969. The quantity theory of money: A restatement. *Studies in the quantity theory of money*. Ed. M. Friedman. Chicago, IL: University of Chicago Press. 1–21.

2. Friedman, M., and M. Bordo. 2006. *The optimum quantity of money*. New Brunswick, NJ: Aldine Transaction. 305 p.
3. Ilyin, V. D. 2010. *S-model' normalizovannoy ekonomicheskoy sistemy* [S-model of the normalized economic system]. Moscow: Institute of Informatics Problems RAS. 103 p.
4. Ilyin, V. D. 26.03.2015. Sledstvie ekonomicheskikh sanktsiy [The consequence of economic sanctions]. *Cnews club*. Available at: [https://club.cnews.ru/blogs/entry/sledstvie\\_ekonomicheskikh\\_sanktsiy](https://club.cnews.ru/blogs/entry/sledstvie_ekonomicheskikh_sanktsiy) (accessed March 30, 2022)
5. Ilyin, V. D. 2009. Model' normalizovannoy ekonomiki (NEk-model'): osnovy kontseptsii [The model of normalized economics (NEc-model): Basics of framework]. *Upravlenie bol'shimi sistemami* [Large-Scale Systems Control] 25:116–138.
6. Bauer, V. P., V. V. Eremin, S. N. Silvestrov, and V. V. Smirnov. 2019. Ekonomicheskoe modelirovanie protsessov tsifrovoy transformatsii [Economic modeling of digital transformation processes]. *Zh. ekonomicheskoy teorii* [Russ. J. Economic Theory] 16(3):428–443. doi: 10.31063/2073-6517/2019.16-3.11.
7. Orlova, I. A., and M. I. Movchan. 2019. Razvitie onlayn-servisov kommercheskikh bankov [Development of online services of commercial banks]. *Innovatsionnye tekhnologii v mashinostroenii, obrazovanii i ekonomike* [Innovative Technologies in Machine-Building, Education and the Economy] 22(2):208–211.
8. Sil'vestrov, S. N., V. P. Bauer, V. V. Eremin, and N. V. Lapenkova. 2020. O tsifrovoy transformatsii predpriyatiya v kontekste sistemnoy ekonomicheskoy teorii [About digital transformation of the enterprise in the context of systemic economic theory]. *Ekonomicheskaya nauka sovremennoy Rossii* [Economics of Contemporary Russia] 2(89):22–45. doi: 10.33293/1609-1442-2020-2(89)-22-45.
9. Kikavets, V. V., and Y. K. Tsaregradskaya. 2020. The implementation of principles of the budget process in the financing of public procurement in the digital economy. *Artificial intelligence: Anthropogenic nature vs. social origin*. Eds. E. Popkova and B. Sergi. Advances in intelligent systems and computing ser. Springer. 1100:28–35.
10. Fedosimov, B. A. 2021. Vvedenie naloga s oborota i potrebleniya dlya obespecheniya ravenstva nalogooblozheniya razlichnykh sub"ektov finansovo-khozyaystvennoy deyatelnosti i uproshtcheniya provedeniya nalogovykh iz"yatiy [The introduction of a turnover and consumption tax to ensure equal taxation of various subjects of financial and economic activity and simplify the implementation of tax exemptions]. *Russian Economic Bulletin* 4(6):299–304.
11. Biryukov, A. N. 2021. Byudzhethirovanie kak instrument ipotechnoy kreditnoy programmy banka [Budgeting as a tool of the bank's mortgage loan program]. *Nauchnoe obozrenie: teoriya i praktika* [Scientific Review: Theory and Practice] 11(3(83)):794–805.
12. Ilyin, A. V., and V. D. Ilyin. 2019. The normalized economic mechanism in the digital environment. *Int. J. Open Information Technologies* 7(12):77–83.
13. Ilyin, V. D. 2016. Normalizatsiya bankovskikh onlayn-servisov [Normalization of online banking services]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 26(2): 147–157.
14. Ilyin, A. V., and V. D. Ilyin. 2015. E-trade with direct lending and normalized money. *Agris on-line Papers in Economics and Informatics* 7(4):57–64.
15. Ilyin, V. D. 2018. Tekhnologiya naznachennykh platezhey v srede tsifrovyykh dvoynikov [Designated payments technology in digital twins environment]. *Sistemy i Sredstva*

*Informatiki — Systems and Means of Informatics* 28(3):227–235. doi: 10.14357/08696527180318.

*Received March 15, 2022*

## **Contributor**

**Ilyin Vladimir D.** (b. 1937) — Doctor of Science in technology, professor, leading scientist, A. A. Dorodnicyn Computing Center, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 40 Vavilov Str., Moscow 119333, Russian Federation; vdilyin@yandex.ru

## СРАВНИТЕЛЬНЫЙ АНАЛИЗ МНОГОПОТОЧНЫХ ТРАНСПОРТНЫХ ПРОТОКОЛОВ

*Е. П. Степанов<sup>1</sup>, Р. Л. Смелянский<sup>2</sup>*

**Аннотация:** Рассматривается задача обеспечения качества сетевого сервиса при помощи протоколов многопоточной маршрутизации, согласно которым поток данных от приложения разделяется на несколько подпотоков и каждый из них идет по своему маршруту в сети. Выделяют два класса многопоточных протоколов: статический, в котором количество подпотоков не изменяется в течение жизни соединения, и динамический, в котором количество подпотоков зависит от текущего состояния сети и требований приложения. Представлен сравнительный анализ эффективности статического и динамического многопоточных транспортных протоколов. Методика сравнительного анализа инвариантна к топологии сети, нагрузке и настройкам транспортных агентов. Выделены области эффективности динамического и статического многопоточных протоколов. Экспериментальное исследование показало преимущество динамического подхода.

**Ключевые слова:** MPTCP; FDMP; QOS; ISP

**DOI:** 10.14357/08696527220215

### 1 Введение

Требования приложений к скорости передачи данных постоянно растут. Так, согласно прогнозу компании Cisco Systems [1], требования к видеотрафику в ближайшем будущем превысят текущие потребности пользователей в десятки раз. Также следует отметить тенденцию к увеличению количества центров обработки данных (ЦОД). Согласно [2], в России на настоящий момент функционируют 146 ЦОД, а во всем мире их уже более 2500 [3]. Рост числа ЦОД связан с меняющейся организацией приложений: часть вычислений может выполняться на периферии [4], позволяя снижать сетевые задержки на передачу информации. Перечисленные изменения приводят и к увеличению трафика между ЦОД и клиентами. Приведенный тезис подтверждается исследованием TeleGeography [5], которое показало, что в 2017 г. доля трафика между ЦОД на самом популярном маршруте через Атлантический океан достигла 75% от общего трафика, а в 2023 г. она должна превысить 93%. Таким образом, на настоящий момент возрастает актуальность задачи выполнения требований приложения к сети, т. е. повышения качества сетевого сервиса. Далее под качеством сервиса будем понимать скорость транспортного соединения.

---

<sup>1</sup>Московский государственный университет имени М. В. Ломоносова, estepanov@lvk.cs.msu.ru

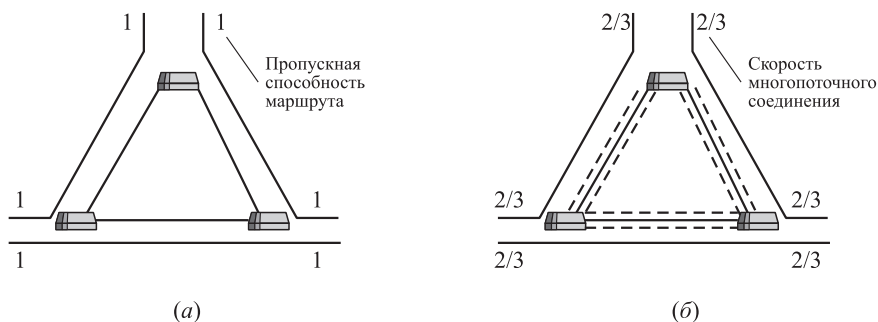
<sup>2</sup>Московский государственный университет имени М. В. Ломоносова, smel@cs.msu.ru

Одним из подходов к увеличению скорости передачи данных может быть разделение потока данных на несколько потоков (называемых далее *подпотоками*), которые передаются через несколько транспортных соединений. Такой подход будем называть *многопоточной маршрутизацией*, а реализующие его протоколы — *многопоточными протоколами*.

Одной из задач многопоточного протокола является выбор количества подпотоков. Скорость транспортного соединения, создаваемого многопоточным протоколом, которое будем называть *многопоточным соединением*, зависит от числа используемых подпотоков. Излишнее количество подпотоков увеличивает накладные расходы на передачу данных, т.е. ресурсы сети используются нерационально.

Можно выделить два класса методов для решения задачи выбора количества подпотоков: статический и динамический. В статических методах количество подпотоков (далее называемых *родственными подпотоками*) одного и того же многопоточного транспортного соединения известно заранее и не изменяется с течением времени. В динамических методах количество родственных подпотоков транспортного соединения зависит от требований приложения к качеству сервиса и от текущей загрузки сети. Когда ресурсов маршрутов текущего множества родственных подпотоков не хватает для обеспечения требуемого качества сервиса, происходит открытие дополнительного подпотока с выделением нового маршрута. Когда качество сервиса транспортного соединения превосходит требования качества приложения, лишние подпотоки закрывают, тем самым освобождая ресурсы сети для обслуживания других потоков.

Возникают вопросы: какой из этих двух методов — статический или динамический — следует использовать, и в каких условиях? Несмотря на простоту статического метода, можно привести пример, когда его использование приводит к неэффективному расходованию ресурсов сети. Рассмотрим топологию, представленную на рис. 1, а. Будем считать, что за каждым маршрутизатором находится пользователь сети и идет передача данных между каждой парой поль-



**Рис. 1** Пример топологии «треугольник», когда используется статический метод многопоточной маршрутизации

зователей. Также будем предполагать, что пропускная способность всех линий связи в представленной топологии равна 1. Тогда суммарная скорость, которую можно получить, используя классические однопоточные протоколы, равна 6 (по 1 в каждую сторону на каждой линии связи). Можно заметить, что между каждой парой пользователей существуют два непересекающихся маршрута, поэтому в статическом методе многопоточной маршрутизации логично использовать два подпотока для каждого многопоточного соединения. Пример организации такой передачи данных изображен на рис. 1, б, где маршрут второго подпотока обозначен штриховой линией.

Чтобы оценить скорость многопоточных соединений, сделаем дополнительное предположение, что пропускная способность каждой линии связи будет делиться поровну между подпотоками, маршруты которых проходят через эту линию связи. Так как пропускная способность каждой линии связи была равна 1, то скорость каждого подпотока будет равна  $1/3$ . Тогда скорость многопоточного соединения будет равна  $2/3$ , а суммарная скорость потоков —  $2/3 \cdot 6 = 4$ . Получаем, что в случае использования статического метода многопоточной маршрутизации скорость будет меньше, чем в случае без использования многопоточных протоколов. Если будем применять динамический метод многопоточной маршрутизации, то подобной ситуации можно избежать, так как второй подпоток не будет создан, его использование не будет увеличивать скорость многопоточного соединения.

Таким образом, задача сравнения эффективности статического и динамического методов многопоточной маршрутизации актуальна. В разд. 2 приведен анализ схожих работ по исследованию эффективности многопоточных протоколов. Раздел 3 посвящен методике проведения сравнительного анализа эффективности статического и динамического методов многопоточной маршрутизации. В разд. 4 приведены результаты экспериментального исследования для сетей интернет-провайдеров.

## 2 Анализ схожих работ

В работах, направленных на анализ эффективности многопоточной маршрутизации, в большинстве случаев в качестве основного критерия исследуют прирост скорости транспортного соединения при статическом многопоточном методе в определенной среде передачи данных или при определенной нагрузке. В работе [6] рассматривается увеличение производительности статического метода при использовании разных алгоритмов планирования сегментов между подпотоками. Авторы статьи [7] исследуют эффективность статического метода для автомобилей, подключенных к сети. Эти автомобили могут координировать свою работу между собой, с городскими службами дорожного движения, а также выполнять вспомогательные вычисления на других сетевых устройствах. Работа [8] посвящена исследованию производительности статического метода в применении к многопоточной маршрутизации в беспроводных децентрализованных самоорганизующихся сетях.

В статье [9] рассматривается влияние разных алгоритмов управления перегрузкой в случае статического многопоточного метода. Авторами этой статьи было проведено исследование массового использования статического метода многопоточной маршрутизации в рамках одной автономной системы. Исследование показало, что применение статического многопоточного транспортного соединения имеет преимущество перед однопоточным протоколом управления передачей (TCP, transmission control protocol) только при низкой загрузке сети, меньшей 37% [10]. Под загрузкой сети понимается среднее значение загрузки по всем линиям связи, где загрузка линии трактуется как отношение суммарной скорости TCP-поток, проходящих через эту линию связи, к ее номинальной пропускной способности.

Из работ по анализу эффективности динамического метода многопоточной маршрутизации можно отметить только работу [11], в которой рассматривается эффективность динамического метода для беспроводных сетей. В работе [12] предлагается динамический метод для автономных систем, однако отсутствует анализ производительности.

Таким образом, в настоящее время отсутствуют работы, посвященные сравнению эффективности статического и динамического методов многопоточной маршрутизации. Анализ схожих работ показал, что наиболее популярным многопоточным протоколом, поддерживающим статический метод, является Multipath TCP (MPTCP) [13]. Несмотря на большое количество различных многопоточных протоколов [14], многие из них являются разными вариантами (улучшениями) MPTCP. Поэтому MPTCP был выбран как основной представитель статического многопоточного протокола для проведения сравнительного анализа эффективности статического и динамического методов многопоточной маршрутизации. Кроме этого, MPTCP имеет доступную открытую реализацию.

В качестве представителя динамического метода в сравнительном анализе использован протокол Flow De-Multiplexing Protocol (FDMP). Его описание дано в [12], где подробно изложено, как устанавливается многопоточное соединение, как выбирается оптимальное число родственных подпотоков, что происходит, если транспортный агент на одном конце соединения — многопоточный, а на другом конце — нет, как идентифицируются родственные подпотоки, относящиеся к одному и тому же транспортному соединению, и т. д. Здесь подробно рассмотрены вопросы маршрутизации родственных подпотоков, методики и инструментарии сравнения и обсуждения результатов экспериментов.

### 3 Методика проведения сравнительного анализа

Для проведения сравнения был использован метод проверки статистических гипотез о попадании случайной величины в заданный интервал. Такое решение было принято, поскольку аналитические методы в данном случае не обеспечат хорошей точности. Дело в том, что нагрузка в сети является случайным процессом, точные математические модели которого пока не известны. Поэтому был

использован метод статистических испытаний, в котором испытаниям подвергалась гипотеза о превосходстве динамического метода над статическим. Для проверки этой гипотезы была сделана оценка числа экспериментов, исходя из желаемого доверительного интервала и уровня значимости.

### 3.1 Входные и выходные данные эксперимента

Сравнительный анализ состоит из проведения серии экспериментов, каждый из которых характеризуется пятеркой  $\langle N, F, T, M, \text{sub}_{\max} \rangle$ :

- $N = \langle G, \rho \rangle$  — транспортная сеть с топологией, представленной ориентированным графом  $G = (V, E)$ , на котором задана функция разметки  $\rho : e \rightarrow (c(e), d(e), \text{loss}(e))$ , где  $e \in E$ ,  $c(e) \in \mathbb{R}^+$ ,  $d(e) \in \mathbb{R}^+$ . Вершины графа — коммутаторы или маршрутизаторы сети; дуги — линии связи. Параметр  $\rho$  характеризует качество сервиса линии связи: пропускную способность  $c(e)$  линии связи, соответствующей дуге  $e$ , задержку передачи  $d(e)$  на этой линии связи и вероятность потери пакета  $\text{loss}(e)$ ;
- $F = \{f_i\}$  — множество потоков данных, передаваемых по сети  $N$ , каждый из которых представлен четверкой  $\langle \text{ip}_i^{\text{src}}, \text{ip}_i^{\text{dst}}, \text{tp}_i^{\text{src}}, \text{tp}_i^{\text{dst}} \rangle$ , где  $\text{ip}_i^{\text{src}}$ ,  $\text{ip}_i^{\text{dst}}$ ,  $\text{tp}_i^{\text{src}}$  и  $\text{tp}_i^{\text{dst}}$  — IP-адреса и транспортные порты отправителя и получателя. Ясно, что такое представление уникально для каждого потока в сети — это фактически адрес на транспортном уровне отправителя и получателя;
- $T = [0, t_{\max}]$  — дискретное время наблюдения за сетью с шагом  $\delta$ ;
- $M = \langle \phi, \psi, \tau \rangle$  — представление свойств потоков. Здесь  $\phi : f_i \rightarrow (v_i^s, v_i^t)$  — распределение полюсов для каждого потока, где под полюсом понимается сток или исток потока (каждый поток обладает своим истоком и стоком, соединенными с сетью в точках  $v_i^s \in V$  и  $v_i^t \in V$  соответственно. Предполагаем, что линии связи между истоком и сетью, а также между сетью и стоком обладают достаточной пропускной способностью для любого потока и нулевой задержкой);  $\psi : f_i \rightarrow (y_i, u_i)$  — требование к качеству потоков — Service Level Agreement (SLA) потока (оно задает объем данных  $y_i \in \mathbb{R}^+$ , который необходимо передать за время  $u_i \in \mathbb{R}^+$ . Эквивалентно можно задать требование к скорости потока  $q_i = y_i/u_i$ );  $\tau : f_i \rightarrow t_i$  — расписание старта потоков, где  $t_i \in T$ ,  $t_i + u_i < t_{\max}$ ;
- $\text{sub}_{\max}$  — максимальное количество родственных подпотоков (в случае МРТСП каждый поток  $f_i$  разбит ровно на  $\text{sub}_{\max}$  родственных подпотоков, а в случае FDMP количество задействованных в передаче данных родственных подпотоков может варьироваться и быть как меньше, так и равно  $\text{sub}_{\max}$ ).

Результатом каждого эксперимента является следующее отображение:

$$\theta : f_i \rightarrow \{L_{ij}, \{r_{ij}\}_t\}, \quad t \in T.$$

Здесь  $L_{ij} = (e_{ij}^1, e_{ij}^2, \dots, e_{ij}^n)$  — маршрут подпотока  $j$  потока  $f_i$ , представленный множеством упорядоченных дуг  $e_{ij}^k \in E$ , где  $k$  — порядковый номер дуги в маршруте;  $\{r_{ij}\}_t$  — временной ряд, элементы  $r_{ij}(t)$  которого представляют среднюю скорость родственного подпотока  $j$  потока  $f_i$  на интервале  $[t - \delta, t]$ ,  $t \in T$ ,  $t \leq t_{\max}$ .

### 3.2 Критерии сравнения

Эффективность статического и динамического методов будем оценивать по следующему критерию: количество потоков, требования SLA которых было удовлетворено.

Введем следующие обозначения. *Средней скоростью потока  $f_i$  на интервале  $[t - \delta, t]$*  назовем величину

$$r_i(t) = \sum_j r_{ij}(t).$$

*Средней скоростью потока  $f_i$*  будем называть величину

$$r_i = \frac{\sum_T r_i(t)}{|\{t_k \in T : t_i \leq t_k \leq t_i + u_i\}|},$$

где  $|\{t_k \in T : t_i \leq t_k \leq t_i + u_i\}|$  — количество дискретных моментов времени, заключенных между временем старта и временем окончания потока.

*Стоимостью подпотока  $j$  потока  $f_i$  на интервале  $[t - \delta, t]$*  назовем величину

$$c_{ij}(t) = r_{ij}(t)|L_{ij}|.$$

Величина  $c_{ij}(t)$  показывает, какую часть пропускной способности линий связи занимает подпоток  $j$  потока  $f_i$  на рассматриваемом интервале. Чем больше длина маршрута, тем больше ресурсов сети затрачивается на передачу этого подпотока. *Стоимостью потока  $f_i$  на интервале  $[t - \delta, t]$*  будет

$$c_i(t) = \sum_j c_{ij}(t).$$

*Средней стоимостью потока  $f_i$*  будем называть величину

$$c_i = \frac{\sum_T c_i(t)}{|\{t_k \in T : c_i(t_k) > 0\}|}.$$

В терминах введенных обозначений количество потоков, для которых требование качества сервиса выполнено, определяет формула

$$sf = |\{f_i : r_i \geq q_i\}|.$$

Тогда стоимость потоков, SLA которых был удовлетворен, можно записать в виде:

$$c = \sum_{i:r_i \geq q_i} c_i.$$

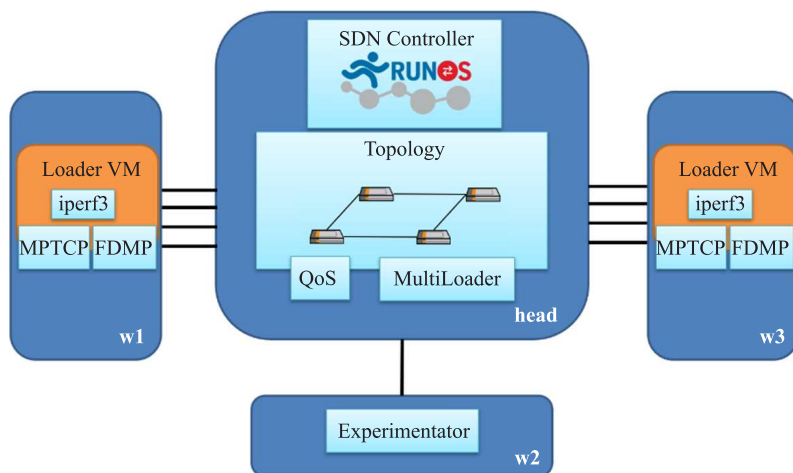
Стоимость потока позволяет различать случаи, когда сравниваемые методы достигают требуемого SLA на одинаковом количестве потоков, но требуют разных затрат ресурсов. Например, рассмотрим топологию сети «треугольник», изображенную на рис. 1. Если в качестве требования к скорости многопоточного соединения выбрать  $2/3$ , то и статический, и динамический метод многопоточной маршрутизации позволят удовлетворить все потоки. Однако в случае статического метода потребуются два родственных подпотока со стоимостью  $1/3 \cdot 1 + 1/3 \cdot 2 = 1$ , а в случае динамического метода — только один подпоток, так как одного маршрута будет достаточно для удовлетворения требуемого качества сервиса, а его стоимость будет  $2/3 \cdot 1 = 2/3$ . Общая стоимость использования статического и динамического методов будет равна 6 и 4 соответственно. В случае статического метода будет задействована вся пропускная способность сети, а динамический метод позволит обслужить дополнительные транспортные потоки.

### 3.3 Описание стенда

Для целей проведения натурных экспериментов был спроектирован и построен стенд, который отвечает следующим требованиям, составленным на основании данного в подразд. 3.1 определения эксперимента.

1. Стенд должен уметь строить имитационную модель сети  $N$  с заданными параметрами качества сервиса линий связи  $\rho$ .
2. Стенд должен уметь генерировать потоки данных  $F$  согласно заданному расписанию  $\tau$  с заданными параметрами SLA  $\psi$ .
3. Стенд должен уметь перенаправлять потоки данных в имитационную модель сети согласно заданному распределению полюсов  $\phi$ .
4. Стенд должен осуществлять отправку потоков данных  $F$  при помощи многопоточных транспортных агентов протоколов MPTCP и FDMP с настраиваемым параметром  $\text{sub}_{\max}$ .
5. Стенд должен уметь строить маршруты  $\{L_{ij}\}$  с наименьшим пересечением для родственных подпотоков.
6. Стенд должен уметь собирать данные о скорости родственных подпотоков  $\{r_{ij}\}_t$  для дискретных моментов времени  $T$ .

Структура экспериментального стенда показана на рис. 2. Стенд состоит из программной и аппаратной частей. Программная часть стенда основана на работе [10], в которой можно найти ее подробное описание.



**Рис. 2** Структура стенда

Подробнее остановимся на требовании 5 о наименьшем пересечении маршрутов родственных подпотоков. Дело в том, что точки пересечения маршрутов — потенциально узкие места, где «сталкиваются» родственные подпотоки. Это приводит к потере эффективности многопоточного соединения. Желательно таких пересечений избегать. Однако возможности управления маршрутизацией родственных подпотоков в традиционных сетях существенно ограничены. В данной статье управление многопоточной маршрутизацией рассматривается применительно к программно-конфигурируемым сетям (ПКС) как наиболее перспективному направлению развития сетей передачи данных, так и в силу вышесказанного. В ПКС контроллер в каждый момент времени знает состояние всей сети и может проложить непересекающиеся маршруты для родственных подпотоков [12].

Управление многопоточной маршрутизацией в ПКС происходит следующим образом. Для каждой пары взаимодействующих вершин контроллер ПКС заранее строит маршруты с минимальным количеством пересечений. Количество маршрутов равно ближайшему наибольшему целому степени демультиплексирования сети, которая равна среднему арифметическому максимального количества непересекающихся маршрутов по всем парам вершин в сети. Контроллер ПКС постоянно мониторит состояние каждого маршрута, где под состоянием маршрута понимается минимальная доступная пропускная способность по всем линиям, составляющим этот маршрут. Это позволяет контроллеру ПКС при появлении многопоточного соединения оценить, хватает ли ресурсов построенных маршрутов для обеспечения требуемого качества сервиса для этого соединения. Если ресурсов хватает, то их закрепляют за этим соединением. В противном случае контроллер ПКС отказывает в многопоточном соединении (посылается пакет с флагом RST).

Для построения маршрутов с наименьшим количеством пересечений был разработан специальный алгоритм MCMF (minimum cost maximum flow), подробно описанный в [10]. Этот алгоритм основан на идее сведения задачи построения маршрутов к задаче поиска максимального потока минимальной стоимости. Алгоритм MCMF, в отличие от более распространенного жадного алгоритма [15], находит оптимальное решение по критерию минимизации штрафа за количество пересечений. Предложенный способ организации многопоточной маршрутизации в ПКС был реализован в стенде для отечественного контроллера ПКС RunOS [16].

Аппаратная часть стенда состоит из 4 серверов: w1, w2, w3 и head. Серверы w1, w2 и w3 имеют процессор Intel Xeon E5-2667 v4 (8 ядер, базовая частота 3,20 ГГц), 32 ГБ оперативной памяти. Сервер head имеет два процессора Intel Xeon E5-2650 v4 (всего 24 ядра, базовая частота 2,20 ГГц), 64 ГБ оперативной памяти. Сервер head соединен с серверами w1 и w4 четырьмя каналами с пропускной способностью 10 Гбит/с, а с сервером w2 — одним каналом с пропускной способностью 10 Гбит/с.

Разработанный стенд отвечает всем шести требованиям, перечисленным в начале подраздела. Использование разных серверов для разных подсистем программной части стенда позволяет исключить гонку за процессорное время для задач генерации потоков данных и имитационного моделирования сети интернет-провайдера.

### 3.4 Организация эксперимента

Перед началом серии экспериментов происходит генерация наборов значений параметров  $\langle N, F, T, M, \text{sub\_max} \rangle$  для каждого эксперимента. Считаем, что допустимые значения параметров имеют равномерное распределение, если не указано иначе. Область допустимых значений для каждого параметра была следующей:

- *топология* ( $G$ ). В качестве топологии используется одна из библиотеки Topology Zoo [17];
- *количество потоков* ( $F$ ). В соответствии с требованиями российского интернет-провайдера Ростелеком количество потоков в сети может достигать 64 000. Поэтому количество потоков в проводимых экспериментах изменялось от 16 000 до 64 000;
- *распределение полюсов* ( $\phi$ ). Количество полюсов, для которых проводилось исследование, изменялось от 10% до 100% от количества вершин в сети с шагом 10. Например, 100% означает, что у каждой вершины в топологии имеется свой полюс. Потоки распределяются равномерно между полюсами. Расположение полюсов в топологии генерируется случайным образом;
- *количество подпотоков* ( $\text{sub\_max}$ ). Основные значения, для которых проводилось исследование, — 2 и 3. Использовались данные значения, так как

- было проведено исследование, которое показало, что степень демультимплексирования для большинства топологий интернет-провайдеров не превосходит 3;
- *расписание старта потоков ( $\tau$ )*. Согласно исследованию [18], время поступления пакетов в автономную систему интернет-провайдера наилучшим образом описывается при помощи распределения Вейбулла. Поэтому в экспериментальном исследовании время старта потока генерируется при помощи распределения Вейбулла от 0 до некоторой константы, которая должна быть меньше времени окончания эксперимента;
  - *классы потоков ( $\psi$ )*. Для всех потоков одной пары виртуальных машин можно задать класс потока. Каждый класс определяет требуемое качество сервиса: объем данных и время, за которое эти данные должны быть переданы. Исходя из параметров класса потока, определяется требование к пропускной способности транспортного соединения. Классы потоков выбираются на основе классов видео (4K, 1440p, 1080p, 720p, 480p, 240p), представленных в [19]. За основу взят видеотрафик, так как, согласно прогнозу Cisco [1], объем видеотрафика значительно возрастет в ближайшем будущем. Требование к времени передачи одного потока данных не превосходит 1 мин;
  - *параметры среды передачи данных ( $\rho$ )*. Все каналы обладают одинаковой пропускной способностью. Значение, для которого проводилось исследование, — 1 Гбит/с. Ограничение в 1 Гбит/с связано с ограничениями стенда, где пропускная способность между серверами не может превышать 40 Гбит/с. Чтобы узкое место образовывалось не на линии связи между серверами, а в виртуальной сети, было выбрано ограничение в 1 Гбит/с;
  - *время окончания эксперимента ( $t$ )* — 3 мин. Так как время окончания потока не превосходит 1 мин, то за 3 мин несколько раз смогут смениться транспортные потоки, проходящие через имитационную модель сети, что позволит лучше исследовать возможности динамического метода многопоточной маршрутизации.

## 4 Результаты экспериментального сравнения

Для объективного выбора оценки числа экспериментов, которые необходимо было провести, чтобы сделать научно обоснованный вывод сравнительного анализа, был применен метод проверки статистических гипотез о числовом значении вероятности события, подробно описанный в [20, 21].

Вначале проверялась гипотеза о преимуществе динамического метода многопоточной маршрутизации над статическим методом. Выдвигалось предположение, что динамический метод не хуже статического с вероятностью не менее 95%. Для проверки этой гипотезы, согласно оценкам из [20, 21], требовалось не менее 422 экспериментов. Всего было проведено 1790 экспериментов, что достаточно для обоснованного принятия решения о выполнимости гипотезы. Из 1790

проведенных экспериментов 1729 (96,59%) экспериментов показывают значение числа удовлетворенных потоков для динамического метода FDMP, больше или равного числу удовлетворенных потоков для статического метода. Для принятия выдвинутой гипотезы с уровнем значимости  $\alpha = 0,01$  необходимо, чтобы значение критерия

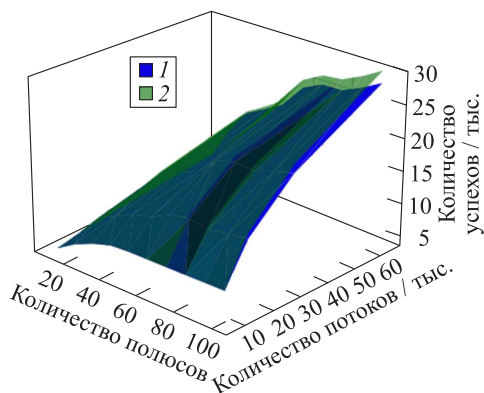
$$\varphi = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}$$

было больше  $x_{\text{лев},\alpha}^{\text{кр}} = -2,34$ . Здесь  $\hat{p} = m/n$ ,  $m$  — количество экспериментов, где динамический метод показал преимущество;  $n$  — общее число экспериментов;  $p_0 = 0,95$ . Для проведенного экспериментального исследования  $\varphi \approx 3,4$ , поэтому гипотеза принимается.

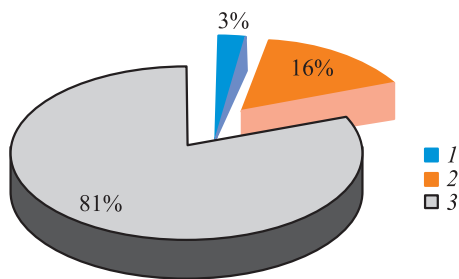
Так как количество экспериментов достаточно велико, то подробные результаты каждого эксперимента о скорости потоков и занимаемых маршрутов заняли бы слишком много места в тексте. Поэтому далее будут приведены только обобщенные результаты по всему экспериментальному исследованию. Однако стенд позволяет оценить эффективность методов многопоточной маршрутизации и для отдельных сетей интернет-провайдеров. Например, для топологии NTT (рис. 3), состоящей из 32 вершин и 63 ребер, из 40 проведенных экспериментов для разного количества потоков и полюсов только 11 экспериментов показали одинаковую эффективность (т. е. равное количество потоков, для которых требования качества сервиса были удовлетворены), а для остальных 29 экспериментов динамический подход оказался лучше. Далее под термином *успех* будем понимать поток, для которого транспортный протокол удовлетворил требования SLA потока. Зависимость количества успехов от параметров эксперимента для статического и динамического методов демультиплексирования в случае топологии NTT показана на рис. 4.



**Рис. 3** Топология NTT ([www.topology-zoo.org](http://www.topology-zoo.org))



**Рис. 4** Зависимость количества успехов от параметров эксперимента для топологии NTT: 1 — статический метод; 2 — динамический метод



**Рис. 5** Сравнение эффективности статического и динамического методов многопоточной маршрутизации: 1 — выигрыш статического метода; 2 — одинаковые результаты; 3 — выигрыш динамического метода

Из 1729 случаев, где динамический метод был не хуже статического, в 289 случаях (16,7%) достигалось равенство по количеству успехов (рис. 5), причем стоимость потоков, определенная в подразд. 3.2, в 102 случаях меньше у динамического метода, в 88 случаях меньше у статического метода, а в остальных случаях одинакова. Стоит отметить, что 80 из 88 случаев равенства количества успехов приходятся на топологии с коэффициентом демультиплексирования 1, т. е. с полным отсутствием альтернативных маршрутов. Большинство остальных случаев равенства приходится на топологии с небольшой степенью демультиплексирования и небольшой долей полосов. Случайное распределение полосов в топологии может привести к ситуации, когда отсутствуют альтернативные маршруты для топологий со степенью демультиплексирования, меньшей 2.

Анализ экспериментов, в которых статический метод оказался лучше, показал, что основная область значений входных параметров, при которых статический метод имеет преимущество, — это топологии с количеством узлов меньше 37 и долей полосов, меньшей или равной 40% от количества узлов в топологии. Для всех остальных случаев динамический

метод будет работать не хуже, а в большинстве случаев — лучше статического метода многопоточной маршрутизации.

Можно сделать предположение, что динамический метод больше обслуживает потоки из класса трафика с небольшим требованием к качеству сервиса, однако анализ результатов показал, что для каждого класса трафика в не менее чем 69% от числа экспериментов наблюдается прирост в числе удовлетворенных потоков при использовании динамического метода.

Возникает вопрос, насколько приведенные результаты будут актуальны в будущем, когда возрастут пропускные способности каналов и требования приложе-

ний. Дополнительное экспериментальное исследование, в котором пропускные способности каналов и требования приложений были увеличены в одинаковое число раз, показало, что эффективность динамического метода многопоточной маршрутизации остается такой же.

## 5 Заключение

В данной статье была предложена методика сравнения эффективности статического и динамического методов многопоточной маршрутизации. Проведенный сравнительный анализ показал преимущество динамического метода с вероятностью не менее 95%, при этом статический метод хорошо работает только на небольших топологиях с количеством узлов менее 36 и небольшой долей полюсов.

Продолжение этой работы может быть направлено на исследование влияния алгоритма управления перегрузкой и алгоритма планирования сегментов в транспортном агенте на эффективность динамического метода многопоточной маршрутизации; исследование динамического многопоточного протокола для подпотоков, работающих по UDP (user datagram protocol), что позволит использовать динамический метод многопоточной маршрутизации в современных протоколах наподобие QUIC (quick UDP Internet connections); исследование динамического метода многопоточной маршрутизации для других видов сетей.

## Литература

1. Cisco Annual Internet Report (2018–2023). <https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.pdf>.
2. Data centers in Russia. <https://cloudscene.com/market/data-centers-in-russia/all>.
3. Locations of Data Centers. <https://www.datacenters.com/locations>.
4. Smeliansky R. Hierarchical edge computing // Conference (International) “Modern Computer Network Technologies.” — IEEE, 2018. 11 p.
5. TeleGeography. Telecommunications market research that’s data-driven. <https://www.telegeography.com/>.
6. Xing Y., Han J., Xue K., Liu J. MPTCP meets Big Data: Customizing transmission strategy for various data flows // IEEE Network, 2020. Vol. 34. No. 4. P. 35–41.
7. Khan I., Chen K. EBA: Efficient bandwidth aggregation for connected vehicles with MPTCP // IEEE Internet Things, 2021. Vol. 9. Iss. 8.
8. Zhang T., Zhao S., Cheng B. Multipath routing and MPTCP-based data delivery over manets // IEEE Access, 2020. Vol. 8. P. 32652–32673.
9. Morawski M., Ignaciuk P. A price to pay for increased throughput in MPTCP transmission of video streams // 24th Conference (International) on System Theory, Control and Computing. — IEEE, 2020. P. 673–678.
10. Stepanov E., Smeliansky R. On analysis of traffic flow demultiplexing effectiveness // Conference (International) “Modern Network Technologies.” — IEEE, 2018. P. 112–122.

11. *Palash M. R., Chen K., Khan I.* Bandwidth-need driven energy efficiency improvement of MPTCP users in wireless networks // IEEE Trans. Green Communications Networking, 2019. Vol. 3. No. 2. P. 343–355.
12. *Chemeritskiy E., Stepanov E., Smeliansky R.* Managing network resources with flow (de) multiplexing protocol // Mathematical Computational Methods Electrical Engineering, 2015. Vol. 53. P. 35–43.
13. *Raiciu C., Paasch C., Barre S.* How hard can it be? Designing and implementing a deployable multipath TCP // USENIX NSDI, 2012. P. 14.
14. *Habib S., Qadir J., Ali A., Habib D., Li M., Sathiaselan A.* The past, present, and future of transport-layer multipath // J. Netw. Comput. Appl., 2016. Vol. 75. P. 236–258.
15. *Kukreja N., Maier G., Alvizu R., Pattavina A.* SDN based automated testbed for evaluating multipath TCP // IEEE Conference (International) on Communications Workshops. — IEEE, 2016. P. 718–723.
16. SDN Controller Runos. <http://github.com/arccn/runos>.
17. *Knight S., Nguyen H. X., Falkner N., Bowden R., Roughan M.* The Internet Topology Zoo // IEEE J. Sel. Area. Comm., 2011. Vol. 29. No. 9. P. 1765–1775. doi: 10.1109/JSAC.2011.111002.
18. *Arfeen M. A., Pawlikowski K., McNickle D., Willig A.* The role of the Weibull distribution in internet traffic modeling // 25th Teletraffic Congress (International Proceedings). — IEEE, 2013. 8 p.
19. Choose live encoder settings, bitrates, and resolutions. <https://support.google.com/youtube/answer/2853702?hl=en>.
20. *Калинина В. Н., Панкин В. Ф.* Математическая статистика. — М.: Высшая школа, 1998. 336 с.
21. *Гмурман В. Е.* Теория вероятностей и математическая статистика. — М.: Высшая школа, 2003. 479 с.

*Поступила в редакцию 25.03.22*

---

---

## COMPARATIVE ANALYSIS OF MULTIPATH TRANSPORT PROTOCOLS

*E. P. Stepanov and R. L. Smeliansky*

M. V. Lomonosov Moscow State University, 1-52 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation

**Abstract:** The paper deals with the problem of ensuring the network service quality by using multipath routing protocols by splitting the application data flow to several subflows, each of which is transmitted through its own route in the network. There are two classes of multipath protocols: the static one with fixed number of subflows during the connection lifetime and the dynamic one which varies the number of subflows depending on the current network state and application requirements. The paper presents the comparative analysis of the effectiveness of static and dynamic multipath transport protocols. The comparison

methodology is invariant to the network topologies, load, and transport agent settings. The areas of static and dynamic multipath protocol efficiency are derived based on study results. The experimental study has shown the advantage of a dynamic approach.

**Keywords:** MPTCP; FDMP; QOS; ISP

**DOI:** 10.14357/08696527220215

## References

1. Cisco Annual Internet Report (2018–2023). Available at: <https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.pdf> (accessed May 20, 2022).
2. Data centers in Russia. <https://cloudscene.com/market/data-centers-in-russia/all> (accessed May 20, 2022).
3. Locations of Data Centers. <https://www.datacenters.com/locations> (accessed May 20, 2022).
4. Smeliansky, R. 2018. Hierarchical edge computing. *Conference (International) "Modern Computer Network Technologies."* IEEE. 11 p.
5. TeleGeography. Telecommunications market research that's data-driven. Available at: <https://www.telegeography.com/> (accessed May 20, 2022).
6. Xing, Y., J. Han, K. Xue, and J. Liu. 2020. MPTCP meets Big Data: Customizing transmission strategy for various data flows. *IEEE Network* 34(4):35–41.
7. Khan, I., and K. Chen. 2021. EBA: Efficient bandwidth aggregation for connected vehicles with MPTCP. *IEEE Internet Things* 9(8).
8. Zhang, T., S. Zhao, and B. Cheng. 2020. Multipath routing and MPTCP-based data delivery over manets. *IEEE Access* 8:32652–32673.
9. Morawski, M., and P. Ignaciuk. 2020. A price to pay for increased throughput in MPTCP transmission of video streams. *24th Conference (International) on System Theory, Control and Computing.* IEEE. 673–678.
10. Stepanov, E., and R. Smeliansky. 2018. On analysis of traffic flow demultiplexing effectiveness. *Conference (International) "Modern Network Technologies."* IEEE. 112–122.
11. Palash, M. R., K. Chen, and I. Khan. 2019. Bandwidth-need driven energy efficiency improvement of MPTCP users in wireless networks. *IEEE Trans. Green Communications Networking* 3(2):343–355.
12. Chemeritskiy, E., E. Stepanov, and R. Smeliansky. 2015. Managing network resources with flow (de)multiplexing protocol. *Mathematical Computational Methods Electrical Engineering* 53:35–43.
13. Raiciu, C., C. Paasch, and S. Barre. 2012. How hard can it be? Designing and implementing a deployable multipath TCP. *USENIX NSDI*. 14.
14. Habib, S., J. Qadir, A. Ali, D. Habib, M. Li, and A. Sathiseelan. 2016. The past, present, and future of transport-layer multipath. *J. Netw. Comput. Appl.* 75:236–258.
15. Kukreja, N., G. Maier, R. Alvizu, and A. Pattavina. 2016. SDN based automated testbed for evaluating multipath TCP. *IEEE Conference (International) on Communications Workshops.* IEEE. 718–723.

16. SDN Controller Runos. Available at: <http://github.com/arccn/runos> (accessed May 19, 2022).
17. Knight, S., H. X. Nguyen, N. Falkner, R. Bowden, and M. Roughan. 2011. The Internet Topology Zoo. *IEEE J. Sel. Area. Comm.* 29(9):1765–1775. doi: 10.1109/JSAC.2011.111002.
18. Arfeen, M. A., K. Pawlikowski, D. McNickle, and A. Willig. 2013. The role of the Weibull distribution in internet traffic modeling. *5th Teletraffic Congress (International) Proceedings*. IEEE. 8 p.
19. Choose live encoder settings, bitrates, and resolutions. Available at: <https://support.google.com/youtube/answer/2853702?hl=en> (accessed May 19, 2022).
20. Kalinina, V. N., and V. F. Pankin. 1998. *Matematicheskaya statistika* [Mathematical statistics]. Moscow: Vysshaya shkola, 1998. 336 p.
21. Gmurman, V. E. 2003. *Teoriya veroyatnostey i matematicheskaya statistika* [Theory of probability and mathematical statistics]. Moscow: Vysshaya shkola. 479 p.

*Received March 25, 2022*

## **Contributors**

**Stepanov Evgeniy P.** (b. 1994) — junior scientist, M. V. Lomonosov Moscow State University, 1-52 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation; [estepanov@lvk.cs.msu.ru](mailto:estepanov@lvk.cs.msu.ru)

**Smeliansky Ruslan L.** (b. 1950) — Doctor of Science in physics and mathematics, professor, Corresponding Member of the Russian Academy of Sciences, head of department, M. V. Lomonosov Moscow State University, 1-52 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation; [smel@cs.msu.ru](mailto:smel@cs.msu.ru)

**Абрамов Алексей Геннадьевич** (р. 1976) — кандидат физико-математических наук, доцент, ведущий научный сотрудник Санкт-Петербургского отделения Межведомственного суперкомпьютерного центра Российской академии наук — филиала Федерального государственного учреждения «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук»

**Адамович Игорь Михайлович** (р. 1934) — кандидат технических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Архипов Павел Олегович** (р. 1979) — кандидат технических наук, директор Орловского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Белозеров Александр Олегович** (р. 2000) — лаборант научно-исследовательской лаборатории моделирования квантовых процессов Тихоокеанского государственного университета

**Волков Олег Игоревич** (р. 1964) — ведущий программист Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Гончар Андрей Андреевич** (р. 1969) — заместитель директора Межведомственного суперкомпьютерного центра Российской академии наук — филиала Федерального государственного учреждения «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук»

**Грушо Александр Александрович** (р. 1946) — доктор физико-математических наук, профессор, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Грушо Николай Александрович** (р. 1982) — кандидат физико-математических наук, старший научный сотрудник Института проблем информатики Феде-

рального исследовательского центра «Информатика и управление» Российской академии наук

**Евсеев Антон Владимирович** (р. 1973) — директор Санкт-Петербургского отделения Межведомственного суперкомпьютерного центра Российской академии наук — филиала Федерального государственного учреждения «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук»

**Егоров Владимир Борисович** (р. 1948) — кандидат технических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Забежайло Михаил Иванович** (р. 1956) — доктор физико-математических наук, доцент, главный научный сотрудник Вычислительного центра им. А. А. Дородницына Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Зацаринный Александр Алексеевич** (р. 1951) — доктор технических наук, профессор, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Зацман Игорь Моисеевич** (р. 1952) — доктор технических наук, заведующий отделом Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Ильин Владимир Дмитриевич** (р. 1937) — доктор технических наук, профессор, ведущий научный сотрудник Вычислительного центра им. А. А. Дородницына Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Ионенков Юрий Сергеевич** (р. 1956) — кандидат технических наук, старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Кириков Игорь Александрович** (р. 1955) — кандидат технических наук, директор Калининградского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Леонтьев Денис Васильевич** (р. 1992) — научный сотрудник Института автоматизации и процессов управления Дальневосточного отделения Российской академии наук

**Листопад Сергей Викторович** (р. 1984) — кандидат технических наук, старший научный сотрудник Калининградского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Мазур Александр Игоревич** (р. 1959) — кандидат физико-математических наук, ведущий научный сотрудник научно-исследовательской лаборатории моделирования квантовых процессов Тихоокеанского государственного университета

**Никишин Дмитрий Александрович** (р. 1976) — кандидат технических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Носова Наталья Юрьевна** (р. 1964) — ведущий инженер Орловского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Синицын Игорь Николаевич** (р. 1940) — доктор технических наук, профессор, заслуженный деятель науки РФ, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук; профессор кафедры Мехатроника и теоретическая механика Московского авиационного института (национального исследовательского университета)

**Смирнов Дмитрий Владимирович** (р. 1984) — бизнес-партнер по ИТ департамента безопасности ПАО «Сбербанк России»

**Смелянский Руслан Леонидович** (р. 1950) — доктор физико-математических наук, профессор, член-корреспондент РАН, заведующий кафедрой автоматизации систем вычислительных комплексов факультета вычислительной математики и кибернетики Московского государственного университета имени М. В. Ломоносова

**Степанов Евгений Павлович** (р. 1994) — младший научный сотрудник Московского государственного университета имени М. В. Ломоносова

**Тимонина Елена Евгеньевна** (р. 1952) — доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Трофименков Андрей Константинович** (р. 1965) — кандидат технических наук, доцент, старший научный сотрудник Орловского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Харитонов Дмитрий Иванович** (р. 1973) — кандидат технических наук, старший научный сотрудник Института автоматики и процессов управления Дальневосточного отделения Российской академии наук

**Цуканов Максим Владимирович** (р. 1995) — младший научный сотрудник Орловского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Чернышов Александр Викторович** (р. 1970) — кандидат технических наук, доцент кафедры Прикладная математика, информатика и вычислительная техника Мытищинского филиала МГТУ им. Н. Э. Баумана

**Шабанов Борис Михайлович** (р. 1954) — доктор технических наук, доцент, директор Межведомственного суперкомпьютерного центра Российской академии наук — филиала Федерального государственного учреждения «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук»

**Шоргин Сергей Яковлевич** (р. 1952) — доктор физико-математических наук, профессор, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

---

## Правила подготовки рукописей статей для публикации в журнале «Системы и средства информатики»

---

Журнал «Системы и средства информатики» публикует теоретические, обзорные и дискуссионные статьи, посвященные научным исследованиям и разработкам в области информационных технологий.

Журнал издается на русском языке. По специальному решению редколлегии отдельные статьи могут печататься на английском языке.

Тематика журнала охватывает следующие направления:

- информационно-телекоммуникационные системы и средства их построения;
- архитектура и программное обеспечение вычислительных машин, комплексов и сетей;
- методы и средства защиты информации.

1. В журнале печатаются статьи, содержащие результаты, ранее не опубликованные и не предназначенные к одновременной публикации в других изданиях.

Публикация предоставленной автором(ами) рукописи не должна нарушать положений глав 69, 70 раздела VII части IV Гражданского кодекса, которые определяют права на результаты интеллектуальной деятельности и средства индивидуализации, в том числе авторские права, в РФ.

Ответственность за нарушение авторских прав, в случае предъявления претензий к редакции журнала, несут авторы статей.

Направляя рукопись в редакцию, авторы сохраняют свои права на данную рукопись и при этом передают учредителям и редколлегии журнала неисключительные права на издание статьи на русском языке (или на языке статьи, если он отличен от русского) и на перевод ее на английский язык, а также на ее распространение в России и за рубежом. Каждый автор должен представить в редакцию подписанный с его стороны «Лицензионный договор о передаче неисключительных прав на использование произведения», текст которого размещен по адресу <http://www.ipiran.ru/publications/licence.doc>. Этот договор может быть представлен в бумажном (в 2-х экз.) или в электронном виде (отсканированная копия заполненного и подписанного документа).

Редколлегия вправе запросить у авторов экспертное заключение о возможности публикации представленной статьи в открытой печати.

2. К статье прилагаются данные автора (авторов) (см. п. 8). При наличии нескольких авторов указывается фамилия автора, ответственного за переписку с редакцией.
3. Редакция журнала осуществляет экспертизу присланных статей в соответствии с принятой в журнале процедурой рецензирования.

Возвращение рукописи на доработку не означает ее принятия к печати.

Доработанный вариант с ответом на замечания рецензента необходимо прислать в редакцию.

4. Решение редколлегии о публикации статьи или ее отклонении сообщается авторам. Редколлегия может также направить авторам текст рецензии на их статью. Дискуссия по поводу отклоненных статей не ведется.

5. Редактура статей высылается авторам для просмотра. Замечания к редакции должны быть присланы авторами в кратчайшие сроки.
6. Рукопись предоставляется в электронном виде в форматах MS WORD (.doc или .docx) или L<sup>A</sup>T<sub>E</sub>X (.tex), дополнительно — в формате .pdf, на дискете, лазерном диске или электронной почтой. Предоставление бумажной рукописи необязательно.
7. При подготовке рукописи в MS Word рекомендуется использовать следующие настройки.

Параметры страницы: формат — A4; ориентация — книжная; поля (см): внутри — 2,5, снаружи — 1,5, сверху и снизу — 2, от края до нижнего колонтитула — 1,3.

Основной текст: стиль — «Обычный», шрифт — Times New Roman, размер — 14 пунктов, абзацный отступ — 0,5 см, 1,5 интервала, выравнивание — по ширине.

Рекомендуемый объем рукописи — не свыше 10 страниц указанного формата. При превышении указанного объема редколлегия вправе потребовать от автора сокращения объема рукописи.

Сокращения слов, помимо стандартных, не допускаются. Допускается минимальное количество аббревиатур.

Все страницы рукописи нумеруются.

Шаблоны примеров оформления представлены в Интернете:

<http://www.ipiran.ru/publications/collected/template.doc>

8. Статья должна содержать следующую информацию на **русском и английском языках**:
  - название статьи;
  - Ф.И.О. авторов, на английском можно только имя и фамилию;
  - место работы, с указанием города и страны и электронного адреса каждого автора;
  - сведения об авторах, в соответствии с форматом, образцы которого представлены на страницах:  
[http://www.ipiran.ru/journal/collected/2019\\_29\\_03\\_rus/authors.asp](http://www.ipiran.ru/journal/collected/2019_29_03_rus/authors.asp) и  
[http://www.ipiran.ru/journal/collected/2019\\_29\\_03\\_eng/authors.asp](http://www.ipiran.ru/journal/collected/2019_29_03_eng/authors.asp);
  - аннотация (не менее 100 слов на каждом из языков). Аннотация — это краткое резюме работы, которое может публиковаться отдельно. Она является основным источником информации в информационных системах и базах данных. Английская аннотация должна быть оригинальной, может не быть дословным переводом русского текста и должна быть написана хорошим английским языком. В аннотации не должно быть ссылок на литературу и, по возможности, формул;
  - ключевые слова — желательно из принятых в мировой научно-технической литературе тематических тезаурусов. Предложения не могут быть ключевыми словами.
  - источники финансирования работы (ссылка на гранты, проекты, поддерживающие организации и т. п.).
9. Требования к спискам литературы.

Ссылки на литературу в тексте статьи нумеруются (в квадратных скобках) и располагаются в каждом из списков литературы в порядке первых упоминаний.

Списки литературы представляются в двух вариантах:

- (1) **Список литературы к русскоязычной части.** Русские и английские работы — на языке и в алфавите оригинала.
- (2) **References.** Русские работы и работы на других языках — в латинской транслитерации с переводом на английский язык; английские работы и работы на других языках — на языке оригинала.

Необходимо для составления списка “References” пользоваться размещенной на сайте <http://www.translit.net/ru/bgn/> бесплатной программой транслитерации русского текста в латиницу.

Список литературы “References” приводится полностью отдельным блоком, повторяя все позиции из списка литературы к русскоязычной части, независимо от того, имеются или нет в нем иностранные источники. Если в списке литературы к русскоязычной части есть ссылки на иностранные публикации, набранные латиницей, они полностью повторяются в списке “References”.

Примеры ссылок на различные виды публикаций в списке “References”:

**Описание статьи из журнала:**

Zhang, Z., and D. Zhu. 2008. Experimental research on the localized electrochemical micromachining. *Russ. J. Electrochem.* 44(8):926–930. doi:10.1134/S1023193508080077.

**Описание статьи из электронного журнала:**

Swaminathan, V., E. Lepkoswka-White, and B. P. Rao. 1999. Browsers or buyers in cyberspace? An investigation of electronic factors influencing electronic exchange. *JCMC* 5(2). Available at: <http://www.ascusc.org/jcmc/vol5/issue2/> (accessed April 28, 2011).

**Описание материалов конференций:**

Usmanov, T. S., A. A. Gusmanov, I. Z. Mullagalin, R. Ju. Muhametshina, A. N. Chervyakova, and A. V. Sveshnikov. 2007. Osobennosti proektirovaniya razrabotki mestorozhdeniy s primeneniem gidrorazryva plasta [Features of the design of field development with the use of hydraulic fracturing]. *Trudy 6-go Mezhdunarodnogo Simpoziuma “Novye resursoberegayushchie tekhnologii nedropol’zovaniya i povyseniya neftegazootdachi”* [6th Symposium (International) “New Energy Saving Subsoil Technologies and the Increasing of the Oil and Gas Impact” Proceedings]. Moscow. 267–272.

**Описание книги (монографии, сборники):**

Lindorf, L. S., and L. G. Mamikonians, eds. 1972. *Ekspluatatsiya turbogeneratorov s neposredstvennym okhlazhdeniem* [Operation of turbine generators with direct cooling]. Moscow: Energy Publ. 352 p.

**Описание переводной книги** (в списке литературы к русскоязычной части необходимо указать: / Пер. с англ. — после названия книги, а в конце ссылки указать оригинал книги в круглых скобках):

1. В русскоязычной части:

Тимошенко С. П., Янг Д. Х., Уивер У. Колебания в инженерном деле / Пер. с англ. — М.: Машиностроение, 1985. 472 с. (*Timoshenko S. P., Young D. H., Weaver W.* Vibration problems in engineering. — 4th ed. — New York, NY, USA: Wiley, 1974. 521 p.)

2. В англоязычной части:

Timoshenko, S. P., D. H. Young, and W. Weaver. 1974. *Vibration problems in engineering*. 4th ed. New York, NY: Wiley. 521 p.

**Описание неопубликованного документа:**

Latypov, A. R., M. M. Khasanov, and V. A. Baikov. 2004. Geology and production (NGT GiD). Certificate on official registration of the computer program No. 2004611198. (In Russian, unpubl.)

**Описание интернет-ресурса:**

Pravila tsitirovaniya istochnikov [Rules for the citing of sources]. Available at: <http://www.scribd.com/doc/1034528/> (accessed February 7, 2011).

**Описание диссертации или автореферата диссертации:**

Semenov, V. I. 2003. Matematicheskoe modelirovanie plazmy v sisteme kompaktnyy tor [Mathematical modeling of the plasma in the compact torus]. Moscow. D.Sc. Diss. 272 p.

Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticheskogo slovary informatsonnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. PhD Thesis. Moscow: IPI RAN. 23 p.

**Описание ГОСТа:**

GOST 8.586.5-2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch'yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. Moscow: Standardinform Publs. 10 p.

**Описание патента:**

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Spособ orientirovaniya po krenu letatel'nogo apparata s opticheskoy golovkoy samonavedeniya [The way to orient on the roll of aircraft with optical homing head]. Patent RF No. 2280590.

10. Присланные в редакцию материалы авторам не возвращаются.
11. При отправке файлов по электронной почте просим придерживаться следующих правил:
  - указывать в поле subject (тема) название журнала и фамилию автора;
  - использовать attach (присоединение);
  - в состав электронной версии статьи должны входить: файл, содержащий текст статьи, и файл(ы), содержащий(е) иллюстрации.
12. Журнал «Системы и средства информатики» является некоммерческим изданием. Плата за публикацию не взимается, гонорар авторам не выплачивается.

**Адрес редакции журнала «Системы и средства информатики»:**

Москва 119333, ул. Вавилова, д. 44, корп. 2, ФИЦ ИУ РАН

Тел.: +7 (499) 135-86-92 Факс: +7 (495) 930-45-05

e-mail: [ssi@frccsc.ru](mailto:ssi@frccsc.ru) (Стригина Светлана Николаевна)

<http://www.ipiran.ru/journal/collected>

---

## Requirements for manuscripts submitted to Journal “Systems and Means of Informatics”

---

Journal “Systems and Means of Informatics” publishes theoretical, review, and discussion articles on the research and development in the field of information technology.

The journal is published in Russian. By a special decision of the editorial board, some articles can be published in English.

Topics covered include the following areas:

- information and communication systems and tools of their design;
- architecture and software of computational complexes and networks; and
- methods and tools of information protection.

1. The Journal publishes original articles which have not been published before and are not intended for simultaneous publication in other editions. An article submitted to the Journal must not violate the Copyright law. Sending the manuscript to the Editorial Board, the authors retain all rights of the owners of the manuscript and transfer the nonexclusive rights to publish the article in Russian (or the language of the article, if not Russian) and its distribution in Russia and abroad to the Founders and the Editorial Board. Authors should submit a letter to the Editorial Board in the following form:

***Agreement on the transfer of rights to publish:***

*“We, the undersigned authors of the manuscript “. . . ,” pass to the Founder and the Editorial Board of the Journal “Systems and Means of Informatics” the nonexclusive right to publish the manuscript of the article in Russian (or in English) in both print and electronic versions of the Journal. We affirm that this publication does not violate the Copyright of other persons or organizations.*

*Author(s) signature(s): (name(s), address(es), date).”*

This agreement should be submitted in paper form or in the form of a scanned copy (signed by the authors).

The Editorial Board has the right to request from the authors an official expert conclusion that the submitted article has no classified data prohibited for publication.

2. A submitted article should be attached with **the data on the author(s)** (see item 8). If there are several authors, the contact person should be indicated who is responsible for correspondence with the Editorial Board and other authors about revisions and final approval of the proofs.
3. The Editorial Board of the Journal examines the article according to the established reviewing procedure. If authors receive their article for correction after reviewing, it does not mean that the article is approved to be published. The corrected article should be sent to the Editorial Board for the subsequent review and approval.
4. The decision on the article publication or its rejection is communicated to the authors. The Editorial Board may also send the reviews on the submitted articles to the authors. Any discussion upon the rejected articles is not possible.
5. The edited articles will be sent to the authors for proofread. The comments of the authors to the edited text of the article should be sent to the Editorial Board as soon as possible.
6. The manuscript of the article should be presented electronically in the MS WORD (.doc or .docx) or L<sup>A</sup>T<sub>E</sub>X (.tex) formats, and additionally in the .pdf format. All documents

may be sent by e-mail or provided on a CD or diskette. A hard copy submission is not necessary.

7. The recommended typesetting instructions for manuscript.

Pages parameters: format A4, portrait orientation, document margins (cm): left — 2.5, right — 1.5, above — 2.0, below — 2.0, footer 1.3.

Text: font — Times New Roman, font size — 14, paragraph indent — 0.5, line spacing — 1.5, justified alignment.

The recommended manuscript size: not more than 10 pages of the specified format. If the specified size exceeded, the editorial board is entitled to require the author to reduce the manuscript.

Use only standard abbreviations. Avoid abbreviations in the title and abstract. The full term for which an abbreviation stands should precede its first use in the text unless it is a standard unit of measurement.

All pages of the manuscript should be numbered.

The templates for the manuscript typesetting are presented on site:

<http://www.ipiran.ru/publication/collected/template.doc>

8. Articles should enclose data both in **Russian and English**:

- title;
- author's name and surname;
- affiliation — organization, its address with ZIP code, city, country, and official e-mail address;
- data on authors according to the format (see site):  
[http://www.ipiran.ru/journal/collected/2019\\_29\\_03\\_rus/authors.asp](http://www.ipiran.ru/journal/collected/2019_29_03_rus/authors.asp) and  
[http://www.ipiran.ru/journal/collected/2019\\_29\\_03\\_eng/authors.asp](http://www.ipiran.ru/journal/collected/2019_29_03_eng/authors.asp);
- abstract (not less than 100 words) both in Russian and in English. Abstract is a short summary of the article that can be published separately. The abstract is the main source of information on the article and it could be included in leading information systems and data bases. The abstract in English has to be an original text and should not be an exact translation of the Russian one. Good English is required. In abstracts, avoid references and formulae.
- Indexing is performed on the basis of keywords. The use of keywords from the internationally accepted thematic Thesauri is recommended.  
Important! Keywords must not be sentences.
- Acknowledgments.

9. References. Russian references have to be presented both in English translation and in Latin transliteration (refer <http://www.translit.net/ru/bgn/>).

Please take into account the following examples of Russian references appearance:

**Article in journal:**

Zhang, Z., and D. Zhu. 2008. Experimental research on the localized electrochemical micromachining. *Russ. J. Electrochem.* 44(8):926–930. doi:10.1134/S1023193508080077.

**Journal article in electronic format:**

Swaminathan, V., E. Lepkoswka-White, and B. P. Rao. 1999. Browsers or buyers in cyberspace? An investigation of electronic factors influencing electronic

exchange. *JCMC* 5(2). Available at: <http://www.ascusc.org/jcmc/vol5/issue2/> (accessed April 28, 2011).

**Conference proceedings:**

Usmanov, T. S., A. A. Gusmanov, I. Z. Mullagalin, R. Ju. Muhametshina, A. N. Chervyakova, and A. V. Sveshnikov. 2007. Osobennosti proektirovaniya razrabotki mestorozhdeniy s primeneniem gidrorazryva plasta [Features of the design of field development with the use of hydraulic fracturing]. *Trudy 6-go Mezhdunarodnogo Simpoziuma "Novye resursoberegayushchie tekhnologii nedropol'zovaniya i povyshe-niya neftegazootdachi"* [6th Symposium (International) "New Energy Saving Subsoil Technologies and the Increasing of the Oil and Gas Impact" Proceedings]. Moscow. 267–272.

**Books and other monographs:**

Lindorf, L. S., and L. G. Mamikonants, eds. 1972. *Ekspluatatsiya turbogenera-torov s neposredstvennym okhlazhdeniem* [Operation of turbine generators with direct cooling]. Moscow: Energy Publs. 352 p.

**Dissertation and Thesis:**

Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticheskogo slovary informat-sionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. Moscow: IPI RAN. PhD Thesis. 23 p.

**State standards and patents:**

GOST 8.586.5-2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichstva zhidkostey i gazov s pomoshch'yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. M.: Standardinform Publs. 10 p.

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Sposob orientirovaniya po krenu letatel'nogo apparata s opticheskoy golovkoy samonavedeniya [The way to orient on the roll of aircraft with optical homing head]. Patent RF No. 2280590.

References in Latin transcription are presented in the original language.

References in the text are numbered according to the order of their first appearance; the number is placed in square brackets. All items from the reference list should be cited.

10. Manuscripts and additional materials are not returned to Authors by the Editorial Board.
11. Submissions of files by e-mail must include:
  - the journal title and author's name in the "Subject" field;
  - an article and additional materials have to be attached using the "attach" function;
  - an electronic version of the article should contain the file with the text and a separate file with figures.
12. "System and Means of Informatics" journal is not a profit publication. There are no charges for the authors as well as there are no royalties.

**Editorial Board address:**

FRC CSC RAS, 44, block 2, Vavilov Str., Moscow 119333, Russia

Ph.: +7 (499) 135 86 92, Fax: +7 (495) 930 45 05

e-mail: [ssi@frccsc.ru](mailto:ssi@frccsc.ru) (to Svetlana Strigina)

[http://www.ipiran.ru/english/journal\\_systems.asp](http://www.ipiran.ru/english/journal_systems.asp)

# SYSTEMS AND MEANS OF INFORMATICS (SISTEMY I SREDSTVA INFORMATIKI)

SCIENTIFIC JOURNAL

**Volume 32 No.2 Year 2022**

**Editor-in-Chief and Chair of Editorial Council**

Academician I. A. Sokolov

## IN THIS ISSUE:

|                                                                                                                                                                                                      |    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| ISSUES OF INCREASING THE NETWORK BANDWIDTH AND TERRITORIAL<br>ACCESSIBILITY OF NATIONAL RESEARCH COMPUTER NETWORK OF RUSSIA<br><i>A. G. Abramov, A. V. Evseev, A. A. Gonchar, and B. M. Shabanov</i> | 4  |
| MACHINE LEARNING FOR EXTRAPOLATION PROBLEMS WITH SMALL DATASET<br><i>A. O. Belozеров and A. I. Mazur</i>                                                                                             | 13 |
| CONSTRUCTING OF THE BRIEF REACHABILITY TREE FOR PROGRAM MODELS<br>IN TERMS OF PETRI NETS<br><i>D. V. Leontyev and D. I. Kharitonov</i>                                                               | 23 |
| ON COMBINING THE SOFTWARE-DEFINED NETWORKING WITH THE NETWORK<br>FUNCTIONS VIRTUALIZATION<br><i>V. B. Egorov</i>                                                                                     | 36 |
| ON THE USE OF EXPERT METHODS IN EVALUATING EFFECTIVENESS<br>AND QUALITY OF INFORMATION SYSTEMS<br><i>A. A. Zatsarinny and Yu. S. Ionenkov</i>                                                        | 47 |
| ANALYTICAL MODELING AND ESTIMATION OF NONSTATIONARY NORMAL<br>PROCESSORS WITH UNSOLVED DERIVATIVES<br><i>I. N. Sinitsyn</i>                                                                          | 58 |
| SOME APPROACHES TO NETWORK DLP ANALYSIS<br><i>A. A. Grusho, N. A. Grusho, M. I. Zabezhalo, D. V. Smirnov,<br/>E. E. Timonina, and S. Ya. Shorgin</i>                                                 | 72 |
| ARCHITECTURES OF INTELLIGENT AGENTS OF COHESIVE HYBRID INTELLIGENT<br>MULTIAGENT SYSTEMS<br><i>S. V. Listopad and I. A. Kirikov</i>                                                                  | 81 |