

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)
профессор В. Хофкирхнер (W. Hofkirchner, Wien, Austria)

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» ВИНТИ
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 № 3 Год 2022

СОДЕРЖАНИЕ

Некоторые методические вопросы оценки уровня технологической готовности проектов информационных систем А. А. Зацаринный, Ю. С. Ионенков	4
Проблемы и угрозы внедрения некоторых новых цифровых технологий В. Е. Гаврилов, А. А. Зацаринный	15
Моделирование выработки полного раннего дифференциального диагноза и прогноза состояния сложных объектов в медицине. Проблемы и перспективы С. Б. Румовская, И. А. Кириков	26
Модели классификации выборки вызванных потенциалов P300 А. М. Самохина, Р. Г. Нейчев, В. В. Гончаренко, Р. К. Григорян, В. В. Стрижов	36
Обзор подходов к пространственно-временному моделированию и выявление основных тенденций развития T-GIS Д. А. Никишин	50
Оптимальные ассоциации устройств в сетях межмашинного взаимодействия стандартов NB-IoT и LTE-M Н. В. Степанов, А. М. Тюрликов, В. О. Бегишев	63
Анализ распределения нагрузок и ресурсов в многопользовательской сети при равных межузловых потоках Ю. Е. Малашенко, И. А. Назарова	71
Варианты самосинхронных регистров сдвига Ю. А. Степченков, Ю. Г. Дьяченко, Ю. В. Рождественский, Н. В. Морозов, Д. Ю. Степченков, Д. Ю. Дьяченко	81

СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ

Том 32 № 3 Год 2022

СОДЕРЖАНИЕ

Методы поиска имплицитных логико-семантических отношений в моноязычных текстах	
<i>А. А. Гончаров</i>	92
Интерфейс пользователя Надкорпусной базы данных иерархий логико-семантических отношений	
<i>А. А. Дурново, О. Ю. Инькова, Н. А. Попкова</i>	103
Стратегия Европейского компьютерного образования: теоретические основания построения	
<i>И. М. Зацман</i>	116
Мультивалютная цифровая платежная система: основы технологического обеспечения	
<i>В. Д. Ильин</i>	128
Подход к поиску аномалий в конкретно-исторических данных	
<i>И. М. Адамович, О. И. Волков</i>	136
О сравнительном генезисе базисных информационных технологий и общественно-экономических формаций Человечества	
<i>С. Н. Гринченко</i>	147
Об авторах	156
Правила подготовки рукописей статей	159
Requirements for manuscripts	163

НЕКОТОРЫЕ МЕТОДИЧЕСКИЕ ВОПРОСЫ ОЦЕНКИ УРОВНЯ ТЕХНОЛОГИЧЕСКОЙ ГОТОВНОСТИ ПРОЕКТОВ ИНФОРМАЦИОННЫХ СИСТЕМ

А. А. Зацаринный¹, Ю. С. Ионенков²

Аннотация: Статья посвящена вопросам оценки уровня технологической готовности (УТГ) проектов информационных систем (ИС). Представлен анализ существующих подходов к оценке уровня готовности проектов за рубежом и в России. Показано, что в настоящее время все большее развитие получает подход, использующий методы оценки технологической/производственной/рыночной готовности (technology/manufacturing/commercialization readiness levels — TRL/MRL/CRL), основанные на 9-уровневой шкале и позволяющие оценить уровень готовности технологии, производства и готовности продукта к выходу на рынок. Рассмотрены преимущества данных методов. Отмечено, что в России они в основном используются в корпоративных проектах и практически не применяются при разработке проектов ИС в интересах государственного управления. Сформулированы предложения по совершенствованию методов оценки проектов ИС, включая методический подход к оценке УТГ проектов.

Ключевые слова: проект; информационная система; уровень готовности проекта; технологическая готовность; производственная готовность

DOI: 10.14357/08696527220301

1 Введение

Современный этап развития цивилизации характеризуется формированием нового технологического уклада — информационного общества, базовыми составляющими которого являются инфокоммуникационные технологии, активное использование информационных ресурсов, изменения в структуре производства и предоставления услуг. Существование информационного общества невозможно без эффективно функционирующих ИС. При этом ИС относятся к классу сложных больших систем с длительным сроком эксплуатации и высокой стоимостью разработки.

Создание ИС, соответствующих современному уровню, требует не только наличия высококвалифицированных специалистов, современной технической и технологической базы, но и повышения роли нормативно-методического обеспечения

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

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

для своевременного принятия обоснованных решений в процессе проектирования и производства систем.

Вопросам совершенствования методического аппарата оценки эффективности и качества ИС посвящен ряд публикаций авторов, в частности [1, 2]. Наряду с этим существенное значение приобретает выработка новых подходов к оценке технологической и производственной готовности инновационных разработок.

В отечественной практике для оценки готовности проектов, как правило, используется подход, предполагающий выполнение и приемку определенных стадий и этапов работы в соответствии с ГОСТ (эскизный проект, технический проект, рабочая документация и т. п.), который основывается на наличии документов, фиксирующих получение определенных результатов работ.

Вместе с тем в настоящее время за рубежом для оценки уровня готовности проектов применяется метод оценки уровней технологической/производственной/рыночной готовности проекта (TRL/MRL/CRL), использующий 9-уровневую шкалу [3–5]. В России также начинает применяться подобный метод, основанный на использовании УТГ [6]. Эти методы представляют собой метрику оценки готовности проектов на разных стадиях и сопоставления различных проектов между собой. Уровни определяются по установленным правилам с учетом концепции проекта, технических и технологических требований.

В данной статье представлены общие подходы к оценке уровня готовности научно-технических проектов ИС в России и за рубежом, а также сформулированы предложения по совершенствованию методов оценки технологической готовности проектов ИС с учетом шкалы TRL.

2 Анализ существующих подходов к оценке уровня готовности проектов информационных систем

В России в настоящее время для оценки уровня готовности проектов преимущественно используется подход, основанный на выполнении определенных стадий и этапов научно-исследовательских и опытно-конструкторских работ (НИОКР) в соответствии с ГОСТ класса 34 (ГОСТ 34.601-90) либо ГОСТ класса 15 (ГОСТ РВ 15.105-2001, ГОСТ РВ 15.203-2001) — эскизный проект, технический проект, рабочая документация и т. п. Этот подход основывается на документальной фиксации выполнения и приемки определенных работ, предусмотренных ГОСТ (техническим заданием (ТЗ), планом выполнения работ и т. п.) для каждой стадии (этапа) НИОКР.

В то же время за рубежом и в ряде организаций в России все большее развитие получают методы оценки готовности проектов — TRL/MRL/CRL [3–5]. Эти методы используют 9-уровневую шкалу, которая позволяет оценить уровень готовности технологии, производства и готовности продукта к выходу на рынок. Оценка выражается в натуральных числах от 1 до 9, при этом 9 соответствует наиболее высокому уровню. Шкала позволяет разработчикам

Таблица 1 Уровни готовности проекта TRL/MRL/CRL

Уровень готовности	Технологическая готовность (TRL)	Производственная готовность (MRL)	Рыночная готовность (CRL)
1	Сформулирована фундаментальная концепция, обоснована полезность технологии	Сделаны выводы по основным требованиям к производству	Проведена оценка востребованности продукта
2	Сформулирована техническая концепция технологии, установлены области применения. Подготовлено ТЗ	Определена концепция производства	Оценены целевые потребительские сегменты, определены ключевые преимущества
3	Разработан макетный образец для демонстрации основных характеристик технологии	Подтверждена концепция производства	Проведен конкурентный анализ
4	Разработан детальный макет, проведены испытания базовых функций	Достигнута возможность изготовления технических средств в лабораторных условиях	Определены поставщики и партнеры, сформирована ценовая политика
5	Работоспособность технологии продемонстрирована на детальном макете	Достигнута возможность изготовления прототипов компонентов систем в реальных условиях	Разработана бизнес-модель
6	Разработан полнофункциональный образец, подтверждены рабочие характеристики в реальных условиях	Достигнута возможность изготовления прототипов систем и подсистем с использованием готовых элементов производства	Получена точная спецификация продукта, уточнена бизнес-модель
7	Прототип технологии продемонстрирован в составе системы в реальных условиях эксплуатации	Достигнута возможность изготовления систем и подсистем в условиях, близких к реальным	Проведен предварительный вывод продукта на рынок
8	Сборка реального устройства и демонстрация в составе системы в реальных условиях	Достигнута готовность к началу мелкосерийного производства	Проанализированы результаты предварительного вывода продукта на рынок, проработаны замечания клиентов
9	Фактическое применение технологии в окончательном виде в реальных условиях	Начато мелкосерийное производство, подготовлена база для полномасштабного производства	Осуществлен вывод продукта на рынок

и заказчикам осуществлять контроль за процессом разработки: проект не перейдет на следующий уровень, пока не будет достигнут предыдущий.

Уровни готовности проекта по основным составляющим: технологическая готовность, производственная готовность и готовность к выходу на рынок — представлены в табл. 1 [7].

Уровень готовности проекта присваивают на основе анализа его качественных и количественных показателей. При этом должно обеспечиваться документальное подтверждение достигнутых показателей. Решение о присвоении уровня принимают научно-исследовательские организации (НИО), занятые в данной области. К оценке также привлекаются независимые эксперты.

Уровень готовности проекта служит полезным инструментом, устанавливающим согласованную терминологию и единообразную методологию оценки уровня зрелости проектов. Данный метод представляет собой систематизированную метрику оценки продукта и позволяет сравнивать уровни готовности различных продуктов. Он обладает гибкостью и возможностью адаптации к конкретному продукту, а также нуждам конкретных отраслей либо организаций.

В России получил распространение метод оценки уровня технологической готовности TRL, который апробирован крупнейшими технологическими лидерами (Росатом, Объединенная авиастроительная корпорация (ОАК), Объединенная двигателестроительная корпорация, фонд «Сколково» и др.). Дирекцией научно-технических программ Минобрнауки России на основе метода TRL разработаны две методологии оценки готовности научно-технических проектов: методология комплексной оценки состояния научно-технических проектов через уровень готовности технологий — TPRL (Technology Project Readiness Level) [8] и модель комплексной оценки технологической готовности инновационных научно-технических проектов [9]. Эти продукты рассчитаны на применение в органах управления сложными комплексными проектами. Они позволяют выявлять динамику развития проектов, оценивать команды исследователей и разработчиков, предсказывать неявные риски комплексных проектов, устранять возможные нарушения по срокам.

Кроме того, Росимуществом разработаны «Методические рекомендации по сопоставлению уровня технологического развития и значений ключевых показателей эффективности акционерных обществ с государственным участием, государственных корпораций, государственных компаний и федеральных государственных унитарных предприятий с уровнем развития и показателями ведущих компаний-аналогов», предусматривающие использование шкал TRL и MRL [10].

Близкий к TRL метод, основанный на использовании 9-уровневой шкалы, предусмотрен в ГОСТ Р 56861-2016 [6], который устанавливает УТГ для проектов (табл. 2).

Таким образом, широко распространенный за рубежом метод оценки готовности проекта TRL/MRL/CRL в некоторой степени получает развитие и в России. При этом в большей степени в корпоративных проектах используются

Таблица 2 Уровни технологической готовности для проектов

Уровень технологической готовности	Примерный состав работ
УТГ 1	Выявлены и зафиксированы фундаментальные принципы технологии
УТГ 2	Концепция или выбор варианта
УТГ 3	Расчетное и (или) экспериментальное обоснование эффективности технологий
УТГ 4	Исследование макетов и (или) компонентов в лабораторных условиях
УТГ 5	Верификация макетов и (или) компонентов при подходящих условиях
УТГ 6	Моделирование систем (подсистем) или испытание трехмерных моделей в подходящих условиях
УТГ 7	Разработка прототипа системы, продемонстрированная на действующем продукте
УТГ 8	Сборка реальной системы и проверка работоспособности в условиях, близких к реальным
УТГ 9	Работа реальной системы в реальных условиях

методы TRL и MRL (Минобрнауки, Росатом, ОАК и др.). Вместе с тем эти методы практически не используются при разработке проектов ИС в интересах государственного управления.

3 Предложения по совершенствованию методов оценки проектов информационных систем

Долговременные проекты разработки инновационных технологий и ИС с их применением требуют существенных ресурсов и связаны со значительными рисками. Начиная с самых ранних стадий разработки таких систем, необходимо применять универсальные показатели, не зависящие от технологий, которые позволяют заказчикам и разработчикам оценивать их состояние и достигнутый прогресс в сравнении с запланированным.

Метод, основанный на применении 9-уровневой шкалы TRL/MRL/CRL, имеет ряд положительных сторон:

- предполагает единый подход к оценке статуса проектов;
- обеспечивает общее понимание статуса проекта;
- позволяет сравнивать уровни готовности различных продуктов;
- позволяет принимать решения о переходе проекта от стадии к стадии;
- используется для принятия решений, касающихся финансирования проектов;

- позволяет выстраивать управление рисками;
- обеспечивает эффективное управление НИОКР (планирование, контроль).

Вместе с тем, как отмечалось выше, этот метод практически не используется при разработке проектов ИС в интересах государственного управления.

Для адаптации данного метода к использованию при разработке проектов ИС в интересах государственного управления может быть предложено несколько групп мероприятий.

Первая группа направлена на создание нормативно-правовой основы применения метода на федеральном уровне на базе постановлений и распоряжений Правительства РФ, соответствующих ГОСТ, а на ведомственном уровне — ведомственных нормативных актов (приказы, распоряжения, методические указания и т. п.).

Вторая группа должна обеспечить увязку с существующими стандартами разработки ИС (ГОСТ классов 15, 19, 34 и др.), что позволит интегрировать подход, основанный на определении уровней готовности проектов, с существующими подходами к разработке систем. Кроме того, это позволит формализовать требования к составу и структуре документов, подтверждающих достижение определенного уровня готовности проекта, и разработать систему подтверждения уровней готовности. При этом следует учитывать, что, например, уровни готовности технологий TRL соответствуют различным стадиям (этапам), определенным в ГОСТ, в частности уровни готовности TRL1–TRL3 соответствуют научно-исследовательским работам, TRL4–TRL6 — стадии опытно-конструкторских работ (ОКР), а TRL7–TRL9 — стадиям ОКР и производства. В соответствии с этим целесообразно планирование этапов НИОКР по результатам с этапами, кратными TRL (разработка технологии, разработка образца, готовность производства и т. п.).

Третья группа охватывает комплекс работ по разработке критериев и показателей оценки уровня готовности проектов. Эти показатели должны быть максимально формализованы, чтобы обеспечить однозначное понимание уровней готовности и учитывать специфику конкретных типов ИС. При этом в число показателей уровня готовности проектов обязательно должны быть включены показатели информационной безопасности. Применение в ИС современных инфокоммуникационных технологий способствует и появлению новых видов опасностей. В этих условиях выполнение проектов в существенной степени зависит от учета возможных уязвимостей и соблюдения всех мер информационной безопасности.

Четвертая группа посвящена формированию системы оценки проектов, включающей как самостоятельную оценку, проводимую разработчиком, так и оценку заказчиком (НИО заказчика) и независимую экспертную оценку третьей стороной. При этом самооценка, проводимая разработчиком, является первичной. Итоговая же оценка определяется рабочей группой, куда входят как представители разработчика, заказчика (НИО заказчика), так и независимые

эксперты. При этом целесообразно определить перечень организаций, из которых выбираются независимые эксперты.

Мероприятия **пятой группы** направлены на совершенствование методического аппарата оценки УТГ проектов. Это совершенствование должно проводиться в направлении разработки методик оценки уровня готовности научно-технических проектов, содержащих перечень показателей уровня их готовности, а также методов их расчета.

Необходимо отметить, что в методиках, изложенных в работах [8, 9], кроме показателей технологической готовности проекта, представленных в табл. 1, предложен ряд дополнительных показателей, учитывающих риски, производственную, инженерную, организационную готовность и т. п. Их разработка вызвана тем, что классическая шкала TRL не охватывает ряд аспектов, которые необходимо учитывать при оценке проекта в целом, например производственную готовность, риски, сроки выполнения проекта и т. п.

4 Методический подход к оценке уровня технологической готовности проектов

С учетом вышеизложенного может быть предложен следующий подход к оценке УТГ проектов. Кроме показателей, характеризующих УТГ проекта (TRL1–TRL9), предлагается ввести ряд дополнительных обобщенных показателей, характеризующих уровень проекта по сравнению с другими проектами, технологические, ресурсные, временные и другие его существующие и потенциальные оценки: потенциальная эффективность результатов; реализуемость; стоимость и сроки выполнения проекта; риски. Каждый из указанных показателей включает в себя ряд частных показателей, которые оцениваются на основе вербально-числовой шкалы (табл. 3).

Оценка УТГ по вербально-числовой шкале проводится экспертами. После получения экспертных оценок проводится усреднение суждений экспертов по формуле

$$a_{ij} = \sqrt[n]{a_{ij}^1 a_{ij}^2 \cdots a_{ij}^n},$$

где a_{ij}^n — оценки, выставленные соответствующими экспертами; n — число экспертов.

Итоговая оценка УТГ проекта по дополнительным показателям может быть получена путем свертки оценок экспертов по группам 1–4 табл. 3 с применением соответствующих весовых коэффициентов. При этом вначале рассчитываются обобщенные показатели по группам 1–3, а затем итоговый показатель по всем четырем группам. Уровень технологической готовности проекта по дополнительным показателям позволяет оценить потенциальную эффективность его результатов, реализуемость и риски, что облегчает принятие решений о финансировании проекта и его дальнейших перспективах.

Таблица 3 Дополнительные показатели для оценки УТГ проекта

Показатели	Вербально-числовая шкала
1. Потенциальная эффективность результатов проекта	
Возможность улучшения характеристик технологии	1 — низкая; 2 — средняя; 3 — высокая
Возможность решения новых задач	1 — низкая; 2 — средняя; 3 — высокая
Уровень технологии по сравнению с мировым	1 — отставание; 2 — соответствие; 3 — опережение
2. Реализуемость технологии	
Новизна технологии	1 — основана на известных методах и моделях; 2 — основана на усовершенствованных методах и моделях; 3 — основана на новых методах и моделях
Обеспечение технологической базы	1 — низкое; 2 — среднее; 3 — высокое
Обеспеченность кадрами	1 — низкая; 2 — средняя; 3 — высокая
3. Стоимость и сроки выполнения проекта	
Оценка стоимости	1 — значительно превышает заявленную; 2 — незначительно превышает заявленную; 3 — соответствует заявленной
Оценка сроков	1 — значительно превышают заявленные; 2 — незначительно превышают заявленные; 3 — соответствуют заявленным
4. Степень риска	
Оценка степени риска	1 — значительный; 2 — допустимый; 3 — минимальный

5 Заключение

Реализация планов развития ИС предполагает определение показателей готовности соответствующих проектов. В связи с этим задача разработки методического аппарата, позволяющего проводить комплексную оценку готовности научно-технических проектов, представляется весьма актуальной.

В настоящее время за рубежом и в некоторой степени в России для оценки уровня готовности проектов получил распространение метод, основанный на применении 9-уровневой шкалы TRL/MRL/CRL и позволяющий оценить уровень готовности технологии, производства и готовности продукта к выходу на рынок и сравнить различные проекты между собой.

В статье рассмотрены основные аспекты данного метода. Отмечено, что в России он в основном используется в корпоративных проектах и практически не применяется при разработке проектов ИС в интересах государственного управления. Сформулированы предложения по совершенствованию методов оценки проектов ИС, включая методический подход к оценке УТГ проектов, позволя-

ющих получить количественные оценки. Эти оценки могут быть использованы для принятия различных управленческих решений, например для разработки плана-графика работ, плана финансирования и т. п.

Литература

1. Зацаринный А. А., Ионенков Ю. С. Оценка эффективности информационно-телекоммуникационных систем. — М.: НИПКЦ Восход-А, 2020. 120 с.
2. Зацаринный А. А., Ионенков Ю. С. Некоторые вопросы оценки качества информационных систем // Системы и средства информатики, 2021. Т. 31. № 4. С. 4–17.
3. Mankins J. C. Technology readiness level. — Advanced Concepts Office of Space Access and Technology NASA, 1995. https://aiaa.kavi.com/apps/group_public/download.php/2212/TRLs_MankinsPaper_1995.pdf.
4. Manufacturing Readiness Level (MRL) Deskbook. Version 2.0. — OSD Manufacturing Technology Program, 2011. https://www.dodmrl.com/MRL_Deskbook_V2.pdf.
5. Метрика «Рыночная готовность CRL». <https://www.rosatom.ru/upload/iblock/fb2/fb23f5ee18bc97e7c0e2f14df9146f6a.pdf>.
6. ГОСТ Р 56861-2016. Система управления жизненным циклом. Разработка концепции изделия и технологий. Общие положения. — М.: Стандартинформ, 2017. 11 с.
7. Уровни готовности продукта/технологии по основным составляющим проекта: технологическая готовность, производственная готовность и готовность к выходу на рынок. http://ctt.volnc.ru/uploads/activity_files/2020/08/14251.pdf.
8. Петров А. Н., Сартори А. В., Филимонов А. В. Комплексная оценка состояния научно-технических проектов через уровень готовности технологий // Экономика науки, 2016. Т. 2. № 4. С. 244–260.
9. Комаров А. В., Петров А. Н., Сартори А. В. Модель комплексной оценки технологической готовности инновационных научно-технических проектов // Экономика науки, 2018. Т. 4. № 1. С. 47–57.
10. Методические рекомендации по сопоставлению уровня технологического развития и значений ключевых показателей эффективности акционерных обществ с государственным участием, государственных корпораций, государственных компаний и федеральных государственных унитарных предприятий с уровнем развития и показателями ведущих компаний-аналогов: Приложение № 2 к протоколу № 2 от 19 сентября 2017 г. заседания Межведомственной рабочей группы по реализации приоритетов инновационного развития президиума Совета при Президенте РФ по модернизации и инновационному развитию России. https://www.economy.gov.ru/material/file/535256e8193016b438e70338a9a0d42d/metodic_ta_mrg.pdf.

Поступила в редакцию 18.02.22

SOME METHODOLOGICAL ISSUES OF ASSESSING THE LEVEL OF TECHNOLOGICAL READINESS OF INFORMATION SYSTEMS PROJECTS

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 issues of assessing the level of technological readiness of information systems (IS) projects. General approaches to assessing the level of readiness of projects abroad and in Russia are presented. It is shown that at present, an approach using the methods of assessing technological/production/market readiness levels (TRL/MRL/CRL) based on a 9-level scale and allowing to assess the level of readiness of technology, production, and product readiness to enter the market is becoming increasingly developed. The advantages of these methods are considered. It is noted that in Russia, they are mainly used in corporate projects and are practically not used in the development of IS projects in the interests of public administration. Proposals are formulated to improve the methods of evaluating IS projects including a methodological approach to assessing the level of technological readiness of projects.

Keywords: project; information system; project readiness level; technological readiness; production readiness

DOI: 10.14357/08696527220301

References

1. Zatsarinny, A. A., and Yu. S. Ionenkov. 2020. *Otsenka effektivnosti informatsionno-telekommunikatsionnykh sistem* [Evaluation of the effectiveness of information and telecommunication systems]. Moscow: NIPKTs Voskhod-A. 120 p.
2. Zatsarinny, A. A., Yu. S. Ionenkov, and A. P. Suchkov. 2021. Nekotorye voprosy otsenki kachestva informatsionnykh sistem [Some issues of information system quality assessment]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 31(4):4–17.
3. Mankins, J. C. 1995. Technology readiness level. Advanced Concepts Office of Space Access and Technology NASA. Available at: https://aiaa.kavi.com/apps/group_public/download.php/2212/TRLs_MankinsPaper_1995.pdf (accessed August 15, 2022).
4. Manufacturing Readiness Level (MRL) Deskbook. Version 2.0. 2011. OSD Manufacturing Technology Program. Available at: https://www.dodmrl.com/MRL-Deskbook_V2.pdf (accessed August 15, 2022).
5. Metrika “Rynochneya gotovnost’ CRL” [The “Market readiness CRL” metric]. Available at: <https://www.rosatom.ru/upload/iblock/fb2/fb23f5ee18bc97e7c0e2f14df9146f6a.pdf> (accessed August 15, 2022).

6. GOST 56861-2016. 2017. Sistema upravleniya zhiznennym tsiklom. Razrabotka kontseptsii izdeliya i tekhnologiy. Obshchie polozheniya [Lifecycle management system. Product concept and technology development. General provisions]. Moscow: Standartinform Publ. 11 p.
7. Urovni gotovnosti produkta/tekhnologii po osnovnym sostavlyayushchim proekta: Tekhnologicheskaya gotovnost', proizvodstvennaya gotovnost' i gotovnost' k vykhodu na rynek [Product/technology readiness levels for the main components of the project: Technology readiness level/manufacturing readiness level/commercialization readiness level]. Available at: http://ctt.volnc.ru/uploads/activity_files/2020/08/14251.pdf (accessed August 15, 2022).
8. Petrov, A. N., A. V. Sartory, and A. V. Filimonov. 2016. Kompleksnaya otsenka sostoyaniya nauchno-tekhnicheskikh proektov cherez uroven' gotovnosti tekhnologiy [Comprehensive assessment of the status of scientific and technical projects using Technology Project Readiness Level]. *Ekonomika nauki* [Economics of Science] 2(4):244–260.
9. Komarov, A. V., A. N. Petrov, and A. V. Sartory. 2018. Model' kompleksnoy otsenki tekhnologicheskoy gotovnosti innovatsionnykh nauchno-tekhnicheskikh proektov [The model of integrated assessment of technological readiness of innovative scientific and technological projects]. *Ekonomika nauki* [Economics of Science] 4(1):47–57.
10. Metodicheskie rekomendatsii po sopostavleniyu urovnya tekhnologicheskogo razvitiya i znacheniy klyuchevykh pokazateley effektivnosti aktsionernykh obshchestv s gosudarstvennym uchastiem, gosudarstvennykh korporatsiy, gosudarstvennykh kompaniy i federal'nykh gosudarstvennykh unitarnykh predpriyatiy s urovnem razvitiya i pokazatelyami vedushchikh kompaniy-analogov: Prilozhenie No. 2 k protokolu No. 2 ot 19 sentyabrya 2017 g. zasedaniya Mezhdedomstvennoy rabochey gruppy po realizatsii prioritetov innovatsionnogo razvitiya prezidiuma Soveta pri Prezidente RF po modernizatsii i innovatsionnomu razvitiyu Rossii [Methodological recommendations for comparing the level of technological development and the values of key performance indicators of state-owned joint-stock companies, state corporations, state-owned companies, and federal state unitary enterprises with the level of development and indicators of leading analog companies: Appendix No. 2 to Protocol No. 2 dated September 19, 2017, of the Meeting of the Interdepartmental Working Group on the Implementation of Innovation Development Priorities of the Presidium of the Council under the President of the Russian Federation for Modernization and Innovative Development of Russia]. Available at: https://www.economy.gov.ru/material/file/535256e8193016b438e70338a9a0d42d/metodic.ta_mrg.pdf (accessed August 15, 2022).

Received February 18, 2022

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

Iononkov 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; uiononkov@ipiran.ru

ПРОБЛЕМЫ И УГРОЗЫ ВНЕДРЕНИЯ НЕКОТОРЫХ НОВЫХ ЦИФРОВЫХ ТЕХНОЛОГИЙ

В. Е. Гаврилов¹, А. А. Зацаринный²

Аннотация: Статья посвящена рассмотрению угроз, связанных с внедрением новых информационных технологий (ИТ) в условиях цифровой трансформации. Рассмотрены некоторые угрозы функционального и системного характера. Показана необходимость и определены направления развития нормативно-технической базы в части обеспечения информационной безопасности.

Ключевые слова: информационная безопасность; угрозы; информационные технологии; функциональная безопасность; системная безопасность; нормативно-методическая база

DOI: 10.14357/08696527220302

1 Введение

Широкое внедрение новых ИТ в самых разных областях экономики, управления, обороны и безопасности порождает новые угрозы информационной безопасности. В условиях взаимозависимости функционирования различных автоматизированных систем, их распределенного характера и общей глобализации цифрового пространства эти угрозы представляют серьезную опасность для функционирования информационных систем и поэтому требуют самого пристального внимания. При этом принципиально изменилась парадигма защиты информации. На первый план, в отличие от конфиденциальности, выходит обеспечение целостности и доступности, корректность реализации алгоритмов функционирования в прикладном программном обеспечении (ПО), а также непротиворечивость и полнота этих алгоритмов.

В статье рассмотрены основные угрозы, которые порождаются применением новых цифровых технологий, определенных программой «Цифровая экономика».

2 Проблемы информационной безопасности в рамках цифровой трансформации

Согласно [1], безопасность ИТ определяется как «состояние информационной технологии, определяющее защищенность информации и ресурсов информацион-

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, vegavrilov@yandex.ru

²Федеральный исследовательский центр «Информатика и управление» Российской академии наук, AZatsarinny@ipiran.ru

ной технологии от действия объективных и субъективных, внешних и внутренних, случайных и преднамеренных угроз, а также способность информационной технологии выполнять предписанные функции без нанесения неприемлемого ущерба субъектам информационных отношений». Однако руководящий документ [2] предусматривает проведение работ только в части защиты от действий нарушителя, направленных на преодоление правил обеспечения конфиденциальности, целостности и доступности информации. Вместе с тем по мере расширения сферы применения ИТ на первый план выходит функциональная безопасность как способность поддерживать корректное функционирование системы управления и управляемого ею оборудования [3].

С учетом актуальности проблемы обеспечения безопасности киберфизических систем был принят закон «О безопасности критической информационной инфраструктуры» [4], но требования по ее защите [5] ограничены только мероприятиями по технической защите информации. При этом предполагается, что функциональная безопасность обеспечивается средствами повышения надежности, противоаварийной защиты и т. п. Однако эти средства строятся на основе предохранительных устройств и блокировок, не защищающих систему от изменения режима работы без перехода оборудования в аварийный режим.

Другой не менее важной представляется проблема обеспечения равнопрочности защиты на всех этапах жизненного цикла системы с учетом того, что общий уровень защиты определяется самым слабым звеном. В последнее время часто пренебрегают проведением работ по обследованию объекта информатизации, предусмотренных [6]. Однако в новых условиях обследование объекта приобретает особую важность, так как именно на этом этапе исследуются алгоритмы автоматизируемых процессов, что позволяет избежать системных ошибок при постановке задачи.

Еще одна проблема состоит в попытках переноса (тиражирования) ИТ, реализованных в защищенных автоматизированных системах для конкретного применения, в другие автоматизированные системы без соответствующей системной проработки с учетом анализа возлагаемых на эту систему функциональных задач.

3 Угрозы и риски применения новых информационных технологий в рамках цифровой трансформации

Программой «Цифровая экономика» [7] предусмотрено развитие и внедрение следующих ИТ: квантовые, распределенного реестра (блокчейн), искусственного интеллекта, биометрические, облачные, большие данные и интернет вещей. Каждая из них порождает те или иные угрозы, которые условно можно разделить на четыре группы:

- (1) угрозы алгоритмического характера — слабые криптографические алгоритмы, уязвимости протоколов и т. д.;

- (2) угрозы технической реализации — ошибки реализации механизмов защиты, уязвимости среды функционирования и др.;
- (3) угрозы функционального характера — ошибки ПО, реализующего алгоритм функционирования системы при решении пользовательских задач;
- (4) угрозы системного характера — ошибки в описании алгоритма функционирования системы и в организации процесса использования ИТ.

Первые две группы характеризуются постоянным противостоянием в формате «щит и меч», но в целом эта ситуация находится под контролем соответствующих регуляторов, принимающих более или менее адекватные меры нормативно-правового и методического характера.

По двум другим группам угроз имеются проблемы, обусловленные во многом отсутствием опыта в обеспечении безопасности ИТ у профильных органов исполнительной власти и ведомственной разобщенности. В связи с этим угрозы при внедрении ИТ рассматриваются ниже в основном с позиций последних двух групп угроз.

3.1 Квантовые технологии

Квантовая криптография. В числе технических проблем квантовой криптографии выделим:

- отсутствие генераторов единичных фотонов (необходимый эффект достигается за счет фильтрации и снижения мощности пучка фотонов);
- ограничения на расстояние передачи данных по квантовому каналу (в настоящее время до 100 км);
- необходимость прямого оптоволокну без разрывов.

С учетом этих ограничений квантовый канал используется только для передачи ключевой информации, а пользовательская информация шифруется традиционной аппаратурой шифрования, и, следовательно, уровень защиты не повышается. Дальнейшая работа с ключами и их хранение в средствах вычислительной техники также не повышают общей надежности такого комбинированного канала шифрованной связи. При этом требуются дополнительное дорогостоящее квантовое оборудование и прямой оптический канал связи только в целях исключения курьерской доставки ключей.

3.2 Технологии искусственного интеллекта

Выделим три вида потенциальных угроз при использовании технологии искусственного интеллекта (ТИИ).

Первый вид угроз обуславливается недостаточно квалифицированной работой по корректной и полной постановке задач при применении ТИИ в различных системах. С учетом этого особую значимость приобретает подготовка соответствующих специалистов — инженеров знаний (онтоинженеров). Однако эта актуальная специальность сегодня отсутствует в официальном реестре.

Второй вид угроз связан с качеством применяемой обучающей выборки (полнота, представительность) и возможностью целенаправленных атак при ее формировании. При обучении «с поощрением» без применения предварительно сформированной обучающей выборки качество результата напрямую зависит от корректности критериев поощрения. Используемый иногда подход, при котором система сама выявляет статистические закономерности и в соответствии с этим модифицирует алгоритм своего функционирования, несет в себе угрозу непредсказуемого развития событий. Получил распространение термин «объяснимый искусственный интеллект». Принятые более чем в 30 странах мира конвенции по использованию искусственного интеллекта также отмечают эту угрозу.

Третий вид угроз заключается в своеобразном гипнозе технологии. Действительно, оператор системы искусственного интеллекта, неоднократно убедившийся в эффективности этой технологии, и в очередной раз при принятии окончательного решения (а в наиболее ответственных системах рекомендован именно такой подход) будет склонен скорее доверять технологии, нежели пытаться объяснить, почему в данном случае этого делать не следует.

3.3 Биометрические технологии

Основная угроза применения биометрических технологий связана с их недостаточной надежностью. Наиболее надежным (помимо ДНК) признается метод идентификации по радужной оболочке глаза — вероятность ложной идентификации на уровне $7 \cdot 10^{-5}$ [8]. Для получения статистически значимой оценки на этом уровне требуется тестовая база объемом 10^9 – 10^{10} образцов, получить которую не представляется возможным. Согласно требованиям [9], вероятность ложной аутентификации не должна превышать 10^{-7} – 10^{-12} в зависимости от класса системы, что делает проблематичной процедуру тестирования биометрических систем.

Стандарты оценки качества биометрических систем [10–12] крайне сложны в применении и не допускают переноса результатов не только на класс однотипных устройств, но и даже на несколько отличающиеся условия использования. Надежность системы в целом определяется наиболее слабым звеном, которым независимо от вида используемых биометрических признаков, скорее всего, станет процедура регистрации пользователя в системе, основанная на решении оператора системы по результатам рассмотрения документов, удостоверяющих личность. При этом вероятность ошибки оператора в теории надежности обычно принимается равной 10^{-3} – 10^{-4} [13]. Таким образом, при использовании технологии биометрической идентификации и аутентификации в высоконадежных системах необходимо применять многофакторные критерии.

3.4 Технологии распределенных реестров

Технология распределенных реестров (блокчейн) в ее первоначальном виде в системах оборота криптовалют характеризуется децентрализацией, аноним-

ностью и невозможностью внесения изменений в уже размещенные блоки данных. Как это часто бывает, ее положительные свойства несут в себе также и угрозы: анонимность обеспечивает удобный канал обезличенного финансирования преступных группировок, неизменность блоков — трудно блокируемый канал распространения вредоносного контента. Предлагаемая иногда технология блокчейна с доверенными центрами, которые вырабатывают очередные блоки, часто не дает ощутимых преимуществ по сравнению с зарекомендовавшей себя технологией электронной подписи (кроме некоторых видов смартконтрактов, позволяющих отслеживать цепочку поставки продукции). Трудности, например, возникают при необходимости внесения изменений в реестр в случае ошибочных записей. Кроме того, доверие к таким системам будет определяться не криптографическими (техническими) средствами, а уровнем доверия к этим центрам.

3.5 Облачные технологии

Разграничение доступа различных групп пользователей в облачных системах достигается путем создания виртуальных подсетей. При этом для каждого из сегментов должна быть создана система технической защиты информации от угроз со стороны внешних и внутренних нарушителей. Однако условия эксплуатации и организационно-технические меры защиты, а значит, и возможности внутреннего нарушителя по проведению атак в различных сегментах могут радикальным образом отличаться, что отражается на построении системы защиты в каждом из сегментов. Таким образом, в случае преодоления средств сегментации общий уровень безопасности в облачной системе будет определяться самым слабым звеном.

Сегментация облачных вычислительных систем достигается путем использования технологии виртуализации, причем даже использование доверенного сертифицированного отечественного гипервизора, обеспечивающего функционирование виртуальных сегментов облачной вычислительной системы, далеко не всегда гарантирует защиту от угроз виртуализации. Если штатный гипервизор, реализованный в недоверенной среде, получит управление раньше, все создаваемые им виртуальные машины попадут под его управление. При этом по внешним признакам не удастся выявить момент передачи ему управления при первоначальной загрузке. Эту угрозу невозможно игнорировать в условиях массового применения зарубежных программно-аппаратных средств.

3.6 Большие данные

Использование аналитической обработки больших объемов информации, получаемой от внешних источников, с точки зрения информационной безопасности создает угрозу деструктивного воздействия на защищаемую систему через входные данные, содержащие исполняемый код, нарушая принцип целостности и замкнутости программной среды. Большую проблему при использовании технологии больших данных представляет выявление фейковой информации.

Проводимые разработчиками различных поисковых систем в этом направлении работы пока не привели к приемлемым результатам [14]. Представляют также опасность попытки целенаправленного воздействия на отбор информации для дальнейшей аналитической обработки. Так, технологии таргетированной рекламы могут использоваться в социальных целях (формирование выборных предпочтений, влияние на результаты соцопросов и т. п.). Особенно актуальна эта атака на входные данные для систем, использующих ТИИ.

3.7 Промышленный интернет, интернет вещей

Появление киберфизических систем, подразумевающих глубокую интеграцию вычислительных ресурсов в физические сущности любого вида, приводит к тому, что проблемы информационной и промышленной, физической безопасности невозможно рассматривать в отрыве друг от друга. При этом происходит резкое расширение сетевой связанности разнородных ресурсов и поверхности атаки. Распространившийся мем «атака на атомную электростанцию через кофеварку» указывает на опасность этого явления. Особое значение приобретает проблема функциональной безопасности систем управления технологическими процессами различного рода при взаимодействии разных сегментов промышленности, транспорта, энергетики. В условиях недостаточного регулирования в области функциональной безопасности (отсутствия методик верификации ПО, проверки корректности описания алгоритмов функционирования технологических процессов, невозможности использования единых компенсирующих мер защиты с учетом радикально отличающихся условий эксплуатации в разных сегментах и др.) проблема безопасности критических объектов в едином информационном пространстве приобретает особую остроту.

4 Предложения по совершенствованию нормативно-правовой базы

Широкое внедрение ИТ в реальный сектор экономики требует принятия адекватных мер по развитию нормативно-правовой базы в области информационной безопасности. С учетом ограничений на применение традиционных методов обеспечения информационной безопасности [15] необходима корректировка нормативно-методической базы с ориентацией на широкое использование методов мониторинга событий в цифровом пространстве с учетом, в том числе, угроз функциональной и системной безопасности.

Представляется целесообразным определить следующие направления:

- обоснование количественных оценок уровней безопасности, основанных на риск-ориентированном подходе [16];
- гармонизация подходов регуляторов в области промышленной и информационной безопасности;
- рассмотрение возможных угроз функционального и системного характера на этапе обследования объекта информатизации;

- переход в классификации уровней защиты информации от привязки к степени секретности (конфиденциальности) к оценке потенциального ущерба [17] как более универсальной характеристике;
- учет в ходе аттестации объектов информатизации факторов их взаимодействия с внешней цифровой средой, в том числе и не имеющей прямых физических подключений;
- обеспечение принципа равнопрочности комплекса мер обеспечения безопасности объекта информатизации, включая системные вопросы организации процесса защиты и компенсирующие организационно-технические меры.

Необходимо активизировать работу по стандартизации в области безопасного в системном плане применения сертифицированных средств и технологий защиты информации (квантовая криптография, использование электронной подписи в электронном документообороте, ведение распределенных реестров на основе использования хеш-функций и др.).

В заключение отметим, что приведенный в статье анализ угроз, связанных с внедрением новых цифровых технологий, требует более детального и внимательного рассмотрения в рамках ведущихся работ, а также при постановке новых работ по созданию автоматизированных систем, прежде всего военного и специального назначения.

Литература

1. ГОСТ Р 50.1.053-2005 Информационные технологии. Основные термины и определения в области технической защиты информации. — М.: Стандартинформ, 2005. 16 с.
2. Автоматизированные системы. Защита от несанкционированного доступа к информации. Классификация автоматизированных систем и требования по защите информации: Руководящий документ, утвержденный решением председателя Государственной технической комиссии при Президенте Российской Федерации от 30 марта 1992 г. — ФСТЭК России, 1992. 29 с. <https://fstec.ru/tekhnicheskaya-zashchita-informatsii/dokumenty/114-spetsialnye-normativnye-dokumenty/384-rukovodyashchij-dokument-reshenie-predsedatelya-gostekkomissii-rossii-ot-30-marta-1992-g>.
3. ГОСТ Р МЭК 62061-2015 Функциональная безопасность систем электрических, электронных, программируемых электронных, связанных с безопасностью. — М.: Стандартинформ, 2016. 72 с.
4. О безопасности критической информационной инфраструктуры Российской Федерации: Федеральный закон от 26.07.2017 № 187-ФЗ. 20 с. <https://fstec.ru/tekhnicheskaya-zashchita-informatsii/obespechenie-bezopasnosti-kii/285-zakony/1610-federalnyj-zakon-ot-26-iyulya-2017-g-n-187-fz>.
5. Об утверждении Требований по обеспечению безопасности значимых объектов критической информационной инфраструктуры Российской Федерации: Приказ ФСТЭК России от 25 декабря 2017 г. № 239. 37 с. <https://fstec.ru/normotvorcheskaya/akty/53-prikazy/1592-prikaz-fstek-rossii-ot-25-dekabrya-2017-g-n-239>.

6. ГОСТ Р 51583-2014 Защита информации. Порядок создания автоматизированных систем в защищенном исполнении. Общие положения. — М.: Стандартинформ, 2018. 15 с.
7. Цифровая экономика Российской Федерации: Программа, утвержденная распоряжением Правительства РФ от 28 июля 2017 г. № 1632-р. <http://d-russia.ru/wpcontent/uploads/2017/07/programma-tsifrov-econ.pdf>.
8. Биометрическая защита, или Берегите ваше тело! // SecurityLab.ru, 12.11.09. <https://www.securitylab.ru/contest/387619.php>.
9. Средства вычислительной техники. Защита от несанкционированного доступа к информации. Показатели защищенности от несанкционированного доступа к информации: Руководящий документ, утвержденный решением председателя Государственной технической комиссии при Президенте Российской Федерации от 30 марта 1992 г. <https://fstec.ru/en/component/attachments/download/297>.
10. ГОСТ Р 58624.1-2019 (ИСО/МЭК 30107-1:2016) Информационные технологии. Биометрия. Обнаружение атаки на биометрическое предъявление. Часть 1. Структура. — М.: Стандартинформ, 2019. 16 с.
11. ГОСТ Р 58624.2-2019 (ИСО/МЭК 30107-2:2017) Информационные технологии. Биометрия. Обнаружение атаки на биометрическое предъявление. Часть 2. Форматы данных. — М.: Стандартинформ, 2019. 16 с.
12. ГОСТ Р 58624.3-2019 (ИСО/МЭК 30107-3:2017) Информационные технологии. Биометрия. Обнаружение атаки на биометрическое предъявление. Часть 3. Испытания и протоколы испытаний. — М.: Стандартинформ, 2019. 45 с.
13. Шубинский И. Б. Функциональная надежность информационных систем. Методы анализа. — М.: Журнал Надежность, 2012. 296 с.
14. Pérez-Rosas V., Kleinberg B., Lefevre A., Mihalcea R. Automatic detection of fake news. Working papers by Cornell University, 2017. arXiv: 1708.07104 [cs.CL]. 10 p.
15. Гаврилов В. Е., Зацаринный А. А. Некоторые системотехнические вопросы использования интеллектуального анализа данных для обеспечения защиты информации в ситуационных центрах // Системы и средства информатики, 2018. Т. 28. № 1. С. 89–98.
16. Гаврилов В. Е., Зацаринный А. А. Исследование проблем нормативно-методического регулирования в области информационной безопасности процессов создания и внедрения информационных технологий, разрабатываемых в рамках программы «Цифровая экономика» // Вестник Воронежского института ФСИИ России, 2020. № 3. С. 30–37.
17. Зацаринный А. А., Гаврилов В. Е. Проблемы нормативно-правового и технического регулирования обеспечения информационной безопасности при создании автоматизированных систем военного назначения // Военная безопасность России: взгляд в будущее: Мат-лы 6-й Междунар. межведомственной науч.-практич. конф. научного отделения № 10 Российской академии ракетных и артиллерийских наук. — М.: МГТУ им. Н. Э. Баумана, 2021. Т. 2. С. 69–75.

Поступила в редакцию 09.07.22

PROBLEMS AND THREATS OF SOME NEW DIGITAL TECHNOLOGIES IMPLEMENTATION

V. E. Gavrilov and A. A. Zatsarinny

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 consideration of threats which arise during the implementation of new digital technologies in frame of digital transformation. Some functional-type and systemic-type threats are presented. The necessity and directions of regulatory and technical base development in the part of information security are shown.

Keywords: information security; threats; information technologies; functional security; systemic security; regulatory and methodical base

DOI: 10.14357/08696527220302

References

1. GOST R 50.1.053-2005. 2005. Informatsionnye tekhnologii. Osnovnye terminy i opredeleniya v oblasti tekhnicheskoy zashchity informatsii [Information technology. General terms and definitions in area of technical protection of information]. Moscow: Standardinform Publ. 16 p.
2. FSTEC Russia. 1992. Avtomatizirovannyye sistemy. Zashchita ot nesanktsionirovannogo dostupa k informatsii. Klassifikatsiya avtomatizirovannykh sistem i trebovaniya po zashchite informatsii [Automated systems. Protection from unauthorized access to information. Classification of automated systems and requirements for data protection]. 29 p. Available at: <https://fstec.ru/tekhnicheskaya-zashchita-informatsii/dokumenty/114-spetsialnye-normativnye-dokumenty/384-rukovodyashchij-dokument-reshenie-predsedatelya-gostekhkommisii-rossii-ot-30-marta-1992-g> (accessed August 17, 2022).
3. GOST R MEK 62061-2015. 2016. Funktsional'naya bezopasnost' sistem elektricheskikh, elektronnykh, programmiruemykh elektronnykh, svyazannykh s bezopasnost'yu [Functional safety of electrical, electronic, and programmable electronic systems related to safety]. Moscow: Standardinform Publ. 72 p.
4. Federal'nyy zakon Rossiyskoy Federatsii ot 26.07.2017 No.187-FZ. O bezopasnosti kriticheskoy informatsionnoy infrastruktury Rossiyskoy Federatsii [State law of the Russian Federation 187-FZ from July 26, 2017. About critical information infrastructure security]. 20 p. Available at: <https://fstec.ru/tekhnicheskaya-zashchita-informatsii/obespechenie-bezopasnosti-kii/285-zakony/1610-federalnyj-zakon-ot-26-iyulya-2017-g-n-187-fz> (accessed August 17, 2022).
5. FSTEC Rossii. December 23, 2017. Ob utverzhdenii Trebovaniy po obespecheniyu bezopasnosti znachimyykh ob"ektov kriticheskoy informatsionnoy infrastruktury

- Rossiyskoy Federatsii: Prikaz No. 239 [On approval of Requirements for ensuring the security of significant objects of critical information infrastructure of the Russian Federation: Order No. 239]. 37 p. Available at: <https://fstec.ru/normotvorcheskaya/akty/53-prikazy/1592-prikaz-fstek-rossii-ot-25-dekabrya-2017-g-n-239> (accessed August 17, 2022).
6. GOST R 51583-2014. 2018. Zashchita informatsii. Poryadok sozdaniya avtomatizirovannykh sistem v zashchishchennom ispolnenii. Obshchie polozheniya [Information protection. The procedure for creating automated systems in a protected version. General provisions]. Moscow: Standardinform Pubs. 15 p.
 7. Tsifrovaya ekonomika Rossiyskoy Federatsii: Programma, utverzhennaya rasporyazheniem Pravitel'stva RF ot 28 iyulya 2017 g. [Digital economy of the Russian Federation: Program approved by the Order No. 1632-r dated July 28, 2017, of the Government of the Russian Federation]. Available at: <http://d-russia.ru/wpcontent/uploads/2017/07/programma-tsfrov-econ.pdf> (accessed August 17, 2022).
 8. SecurityLab.ru. 12.11.09. Biometricheskaya zashchita, ili beregite vashe telo [Biometric protection or take care of your body]. Available at: <https://www.securitylab.ru/contest/387619.php> (accessed August 17, 2022).
 9. FSTEC Russia. 1992. Sredstva vychislitel'noy tekhniki. Zashchita ot nesanktsionirovannogo dostupa k informatsii. Pokazateli zashchishchennosti ot nesanktsionirovannogo dostupa k informatsii [Computer technics means. Protection from unauthorized access to information. Security settings regarding unauthorized access to information]. Available at: <https://fstec.ru/en/component/attachments/download/297> (accessed August 17, 2022).
 10. GOST R 58624.1. 2019. Informatsionnye tekhnologii. Biometriya. Obnaruzhenie ataki na biometricheskoe pred'yavlenie. Chast' 1. Struktura [Information technologies. Biometrics. Detection of an attack on biometric presentation. Part 1. Structure]. Moscow: Standardinform Pubs. 16 p.
 11. GOST R 58624.2. 2019. Informatsionnye tekhnologii. Biometriya. Obnaruzhenie ataki na biometricheskoe pred'yavlenie. Chast' 2. Formaty dannykh [Information technologies. Biometrics. Detection of an attack on biometric presentation. Part 2. Data formats]. Moscow: Standardinform Pubs. 16 p.
 12. GOST R 58624.3. 2019. Informatsionnye tekhnologii. Biometriya. Obnaruzhenie ataki na biometricheskoe pred'yavlenie. Chast' 3. Ispytaniya i protokoly ispytaniy [Information technologies. Biometrics. Detection of an attack on biometric presentation. Part 3. Tests and test reports]. Moscow: Standardinform Pubs. 45 p.
 13. Shubinskiy, I. B. 2012. *Funktsional'naya nadezhnost' informatsionnykh sistem. Metody analiza*. Moscow: Zhurnal Nadezhnost'. 296 p.
 14. Pérez-Rosas, V., B. Kleinberg, A. Lefevre, and R. Mihalcea. 2017. Automatic detection of fake news. *arXiv.org*. 10 p. Available at: <https://arxiv.org/abs/1708.07104> (accessed August 17, 2022).
 15. Gavrilov, V. E., and A. A. Zatsarinny. 2018. Nekotorye sistemotekhnicheskie voprosy ispol'zovaniya intellektual'nogo analiza dannykh dlya obespecheniya zashchity informatsii v situatsionnykh tsentrakh [Some system-technical issues of using data mining to ensure the protection of information in situational centers]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(1):89–98.
 16. Gavrilov, V. E., and A. A. Zatsarinny. 2020. Issledovanie problem normativno-metodicheskogo regulirovaniya v oblasti informatsionnoy bezopasnosti protsessov so-

zdaniya i vnedreniya informatsionnykh tekhnologiy, razrabatyvaemykh v ramkakh programmy “Tsifrovaya ekonomika” [Study of the problems of normative-methodical management in the field of information security processes for the creation and implementation of information technologies developed in the framework of the “Digital economy”]. *Vestnik Voronezhskogo instituta FSIN Rossii* [Vestnik of Voronezh Institute of the Russian Federal Penitentiary Service] 3:30–37.

17. Zatsarinnyy, A. A., and V. E. Gavrilov. 2021. Problemy normativno-pravovogo i tekhnicheskogo regulirovaniya obespecheniya informatsionnoy bezopasnosti pri sozdanii avtomatizirovannykh sistem voennogo naznacheniya [Problems of regulatory and technical regulation of information security in the creation of automated systems for military purposes]. *6th Interdepartmental Scientific and Practical Conference (International) of the Scientific Department No. 10 of the Russian Academy of Rocket and Artillery Sciences “Military Security: A Look into the Future” Proceedings*. Moscow: Bauman University. 2:69–75.

Received July 09, 2022

Contributors

Gavrilov Victor E. (b. 1950) — principal specialist, 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; vegavrilov@yandex.ru

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

МОДЕЛИРОВАНИЕ ВЫРАБОТКИ ПОЛНОГО РАННЕГО ДИФФЕРЕНЦИАЛЬНОГО ДИАГНОЗА И ПРОГНОЗА СОСТОЯНИЯ СЛОЖНЫХ ОБЪЕКТОВ В МЕДИЦИНЕ. ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ

С. Б. Румовская¹, И. А. Кириков²

Аннотация: Рассматривается проблематика персонализированной оценки состояния (ПОС) сложного объекта (СО) как гетерогенной системы и предлагается разработать и исследовать новый подход к повышению эффективности и качества решения проблем медицинской диагностики и прогнозирования в условиях существенного разнообразия, неоднородности и неопределенности информации об СО диагностики, а также гетерогенности процесса персонализированной оценки их состояния. Предлагается создать и исследовать методы и модели нового класса гибридных интеллектуальных систем (ГиИС) для медицинской диагностики и прогнозирования, интегрирующие разнородные модели выработки и интерпретации ПОС СО на разных этапах процесса обследования, а также комбинирующие модели ГиИС двух типов — функциональную и трансформационную.

Ключевые слова: гибридная интеллектуальная система; интеллектуальная система поддержки принятия решений; сложный объект диагностики; медицинская диагностика и прогнозирование

DOI: 10.14357/08696527220303

1 Введение

Медицина и здравоохранение — одна из стратегических и перспективных областей применения технологий искусственного интеллекта для повышения качества профилактических обследований, диагностики, прогнозирования возникновения и развития заболеваний, подбора оптимальных дозировок лекарственных препаратов и др. [1]. Увеличивающиеся объемы и разнообразие симптоматической информации, а также информации об этиологии и патогенезе заболеваний существенно повышают нагрузку на врачей. При этом решения, включающие прогнозирование развития того или иного исхода, включая смерть, зачастую основываются только на личном опыте врача в условиях ограниченности во времени

¹Калининградский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, sophiyabr@gmail.com

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

и существующих в этой области научных данных, что впоследствии приводит к ошибкам в выборе метода терапии, которые могут стать фатальными. Все это приводит к снижению эффективности мер здоровьесбережения и нерациональному распределению человеческих, лекарственных, технических и финансовых ресурсов, а также к выбору не релевантных ситуации терапевтических и диагностических стратегий. Положение можно нивелировать путем повышения уровня поддержки персонализированной медицины информационными технологиями, в частности системами искусственного интеллекта.

Таким образом, актуально исследование проблематики ПОС СО как гетерогенной системы, а также разработки ГиИС персонализированной диагностики и прогнозирования состояния организма человека (СО). Подобные системы — междисциплинарного характера, они комбинируют методы искусственного интеллекта различных классов и отображают функциональную структуру решаемой проблемы, области ее однородных параметров, представленных множеством однородных задач, каждая из которых должна решаться специалистом(ами) соответствующей области знаний. Гибридные интеллектуальные системы персонализированной диагностики и прогнозирования состояния СО будут вырабатывать заключения, релевантные реальной ситуации обследования пациента, а также выведут процессы здоровьесбережения на новый уровень — своевременной ранней диагностики. Такие системы повысят качество диагностических решений и прогноза, а также, как следствие, плана лечения, ошибки в котором могут привести к фатальным результатам.

В данной работе представлены результаты анализа особенностей выработки решений при ПОС СО как гетерогенной системы и описана концепция нового исследования, направленного на моделирование выработки полного раннего дифференциального диагноза и прогноза состояния СО в медицине.

2 Проблема персонализированной оценки состояния сложного объекта как гетерогенной системы

Персонализированная оценка состояния СО в медицине включает две взаимосвязанные проблемы:

- (1) ранней полной дифференциальной диагностики (рис. 1);
- (2) прогнозирования исходов заболевания с обоснованием прогноза, достоверность которого зависит от точности диагноза.

Для первой характерны гетерогенность и полиморбидность (наличие нескольких заболеваний одновременно), что изменяет и осложняет клиническую картину. Проблемы дифференциальной диагностики и прогнозирования состояния здоровья пациента также осложняются:

- (1) ситуациями Р. Ригельмана [2] — наличием атипичных симптомов и симптомов-миражей;

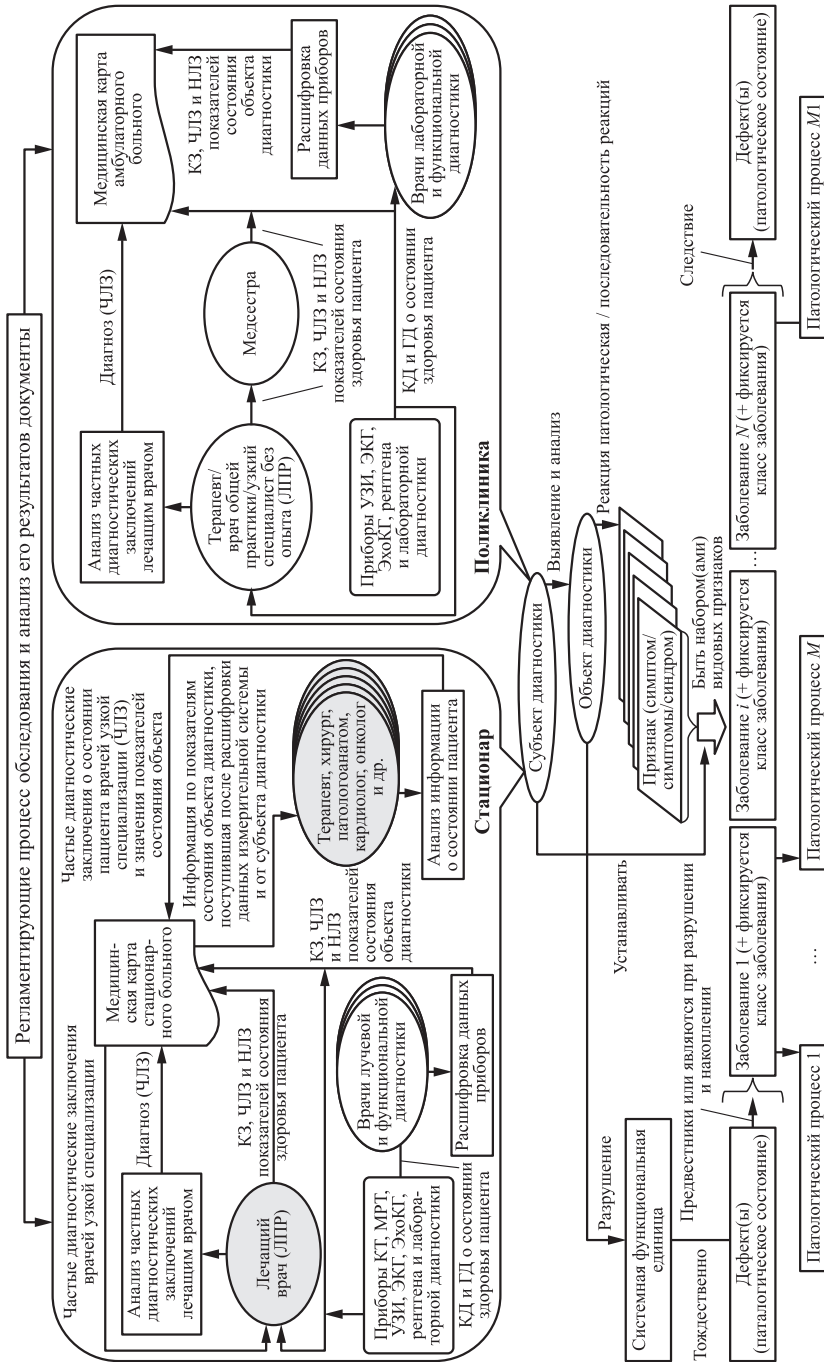


Рис. 1 Схема процесса оценки состояния СО как гетерогенной системы в стационаре и в поликлинике: КЗ, ЧЛЗ и НЛЗ — количественные, четкие и нечеткие лингвистические значения соответственно; КТ и МРТ — компьютерная и магнитно-резонансная томография; УЗИ — ультразвуковое исследование; ЭКГ и ЭхоКГ — электро- и эхокардиография; КД и ГД — количественные и графические данные

- (2) разнообразием обрабатываемой симптоматической информации и, соответственно, направлений медицинских исследований состояния здоровья пациента;
- (3) разнообразием и зачастую непредсказуемостью клинической картины, пересекающейся с рядом других заболеваний.

Данные обстоятельства требуют участия коллектива специалистов разных профилей в принятии диагностического решения, что зачастую доступно только в стационарных больничных учреждениях.

В лечебно-профилактических учреждениях (ЛПУ) амбулаторно-поликлинического типа (см. рис. 1) ввиду ограниченности времени на прием и недостатка врачей с опытом работы более 10 лет, особенно узкой специализации, опираются зачастую на международную классификацию болезней, которая базируется на структурном анализе, поэтому врачи мыслят, преимущественно отталкиваясь от таких понятий, как воспаления, гипо- и гипертрофии, опухоли и т. д. Однако перечисленное — это дефекты (на рис. 1 патологические состояния тождественны накоплению разрушений системных функциональных единиц — клеточных и тканевых структур различного уровня детализации). Дефекты служат предвестниками или следствиями заболеваний (патологических процессов). Заболевание тождественно патологическому процессу, если: последний выступает необходимой стадией заболевания (отложение холестерина в венечных артериях сердца — атеросклероз); указывается локализация (цирроз печени); имеет место зависимость от этиологического фактора (специфическое воспаление, вызываемое микробактериями туберкулеза или лепры, например); ранее идентифицировался как патологический процесс, а теперь — болезнь (ожоговая, постренимационная). Таким образом, в ЛПУ поликлинического типа превалирует анатомо-морфологический анализ. Но еще М. Я. Мудров в XIX в. высказал положение о целостности организма, о единстве психики и соматики, о роли и значении детального расспроса в распознавании болезней [3]. Необходимо уйти от восприятия пациента односторонне, одноаспектно, с позиций одной профессии и/или одной дисциплины и/или одной школы, с позиции патологических состояний, что характерно для гомогенных систем и ведет к решению проблем в медицине несистемно — разнообразные знания не взаимодействуют, не сотрудничают для достижения общей цели, что по определению не создает условий для возникновения синергетического, системного эффекта. Переход к системным и функциональным категориям при рассмотрении объекта диагностики персонализированно как целостной гетерогенной системы позволит выявлять заболевания на ранних стадиях и строить более точные, достоверные и обоснованные прогнозы.

Рассмотрим далее описание концепции создания методов и моделей нового класса ГиИС для медицинской диагностики и прогнозирования — «Виртуальные консилиумы персонализированной диагностики и прогнозирования» (ВКПДП), которые будут вырабатывать заключения, релевантные реальной ситуации обследования пациента.

3 Виртуальные консилиумы персонализированной диагностики и прогнозирования. Подход к разработке

Проблема ПОС СО в медицине относится к области гетерогенных (неоднородных), как уже было отмечено выше. Такие проблемы включают подзадачи, каждая из которых должна решаться специалистом (или специалистами) соответствующей области знаний. С ними лучше всего справляются именно коллективы. Поэтому в рамках нового исследования предлагается разрабатывать интеллектуальные системы поддержки принятия диагностических решений, включающих прогноз развития заболевания и его исходов, такие как ГиИС (рис. 2).

Гибридные интеллектуальные системы позволяют интегрировать разнородные модели знаний о процессе диагностики, прогнозировании исходов и интерпретации прогноза выявленных дисфункций и патологий у СО на всех фазах заболевания и тем самым моделировать взаимодействие рассуждений в реальных

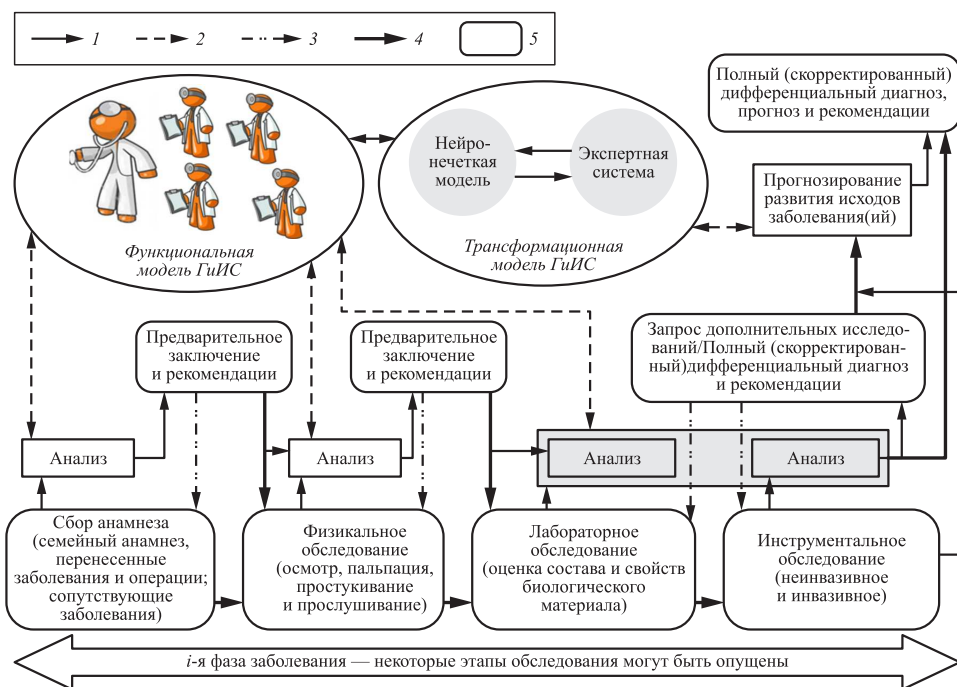


Рис. 2 Схема ВКПДП: 1 — передача входной-выходной информации; 2 — инициализация и запуск функциональной модели ГиИС / запуск трансформационной модели ГиИС; 3 — запрос на уточнение входной информации; 4 — переход к следующему этапу и передача всей полученной на предыдущих этапах входной и выходной информации; 5 — интерфейсы ввода/вывода информации с возможностью корректировки

коллективах врачей, привлекаемых к обследованию пациента, во главе с лицом, принимающим решения (ЛПР). Интегрированная модель будет синтезироваться ситуативно в зависимости от выявленной симптоматики как на разных этапах диагностического процесса (при необходимости), так и в конце обследования — после получения результатов лабораторной и инструментальной диагностики. При этом в рамках исследования предлагается комбинировать модели ГиИС двух типов — функциональную и трансформационную (см. рис. 2).

Первая для повышения качества диагностических решений будет разрабатываться по принципу виртуального консилиума (ВК) [4] в рамках проблемно-структурной методологии разработки ГиИС Калининградского филиала Федерального исследовательского центра «Информатика и управление» РАН [5], сочетающих методы и методики анализа проблемы с динамическим синтезом ее декомпозированного представления и метода ее решения, моделирующих взаимодействие рассуждений в реальных коллективах врачей, привлекаемых к обследованию пациента, во главе с ЛПР.

Проблема прогнозирования решается также ЛПР, поэтому ВК будет расширен трансформационной ГиИС «Нейро-нечеткая модель на базе ANFIS (adaptive neuro-fuzzy inference system, адаптивная сеть на основе системы нечеткого вывода) — экспертная система» (ANFIS–ЭС). ANFIS позволит обеспечить формализацию размытых понятий и связей (нечеткость в рассуждениях врачей), вырабатывая прогноз, который будет интерпретироваться ЭС. В случае отсутствия обоснования полное описание ситуации будет предоставлено ЛПР, который решает: пополнить базу знаний ЭС либо получить прогноз заново на уточненных данных. Кроме того, при использовании нечеткой логики значительно снижается размер базы знаний, а также обучающей и тестирующей выборки [6].

Интегрированные модели выработки ПОС СО и прогноза его возможных состояний будут комбинировать методы из различных классов (аналитические методы, имитационное статистическое моделирование, искусственные нейронные сети, генетические алгоритмы, ЭС, нечеткие системы) в соответствии с их плюсами и минусами и отображать в памяти компьютера функциональную структуру решаемой проблемы, области ее однородных параметров (детерминированных, стохастических, логических, лингвистических и др.), представленных множествами задач из декомпозиции проблемы. В рамках исследования планируется строить прозрачные прогнозы о состоянии СО диагностики с учетом ситуации возможного пополнения знаний об их формировании, а также предполагается системный анализ как СО диагностики, так и проблемы персонализированной оценки его состояния в целом, а также особенностей процессов диагностики и прогнозирования, комбинируя методы постепенной формализации проблемы с системным и объектным подходами к анализу объектов исследования.

Таким образом, исследование будет базироваться на принципах персонализированной медицины [7] — поддержка ранней диагностики заболеваний и предпосылок к нему, а также прогнозирование будущих состояний объекта диагностики (исходов выявленных заболеваний), что впоследствии поможет врачу определить-

ся с правильной дальнейшей терапией. Предлагаемое исследование направлено на развитие междисциплинарных областей знания путем создания и исследования методов и моделей ВКПДП на примере острого панкреатита (ОП) — третьего среди встречающихся острых хирургических заболеваний органов брюшной полости. Острый панкреатит [8–10] составляет в структуре «острого живота» 28%–45%. Летальность при тяжелом ОП (панкреонекрозе) — приблизительно 18%–25%, достигая 40%–70% при инфицированном панкреонекрозе.

4 Заключение

Виртуальные консилиумы персонализированной диагностики и прогнозирования будут интегрировать разнородные модели выработки и интерпретации решения проблем ПОС ситуационно на разных этапах процесса обследования СО, а также комбинировать модели ГиИС двух типов — функциональную и трансформационную. Их разработка позволит повысить релевантность интеллектуальных систем поддержки принятия решений реальным проблемам в СО и гетерогенности процесса обследования, что увеличит оперативность, достоверность и обоснованность рекомендаций системы. Такие системы выведут процессы здоровьесбережения на новый уровень — своевременной ранней диагностики — ввиду системного подхода к анализу и моделированию ПОС СО. Их применение на практике позволит повысить оперативность, достоверность и обоснованность рекомендаций интеллектуальных систем поддержки принятия решений при диагностике и прогнозировании в медицине, в частности в ЛПУ амбулаторно-поликлинического типа (здесь относительно ОП доля диагностических ошибок составляет порядка 30%–60% в зависимости от проявленной симптоматики [11, 12]). Недостаток опытных врачей и большой поток пациентов имеют место и в стационарных многопрофильных больничных учреждениях, что приводит к ошибкам примерно в 10%–26% [11, 12] случаев поступлений с болями в животе, особенно относительно диагностирования тяжелой формы ОП (соответственно, и дальнейших прогнозов). При этом из-за разнообразия и зачастую непредсказуемости клинической картины ОП, пересекающейся с рядом urgentных заболеваний органов брюшной полости, до 17,2% больных ОП ошибочно оперируются из-за подозрения при поступлении в стационар на другие острые хирургические заболевания. Поэтому системы, предлагаемые в рамках исследования, актуальны и для стационара.

Литература

1. О развитии искусственного интеллекта в Российской Федерации: Указ Президента РФ от 10 октября 2019 г. № 490. <http://www.kremlin.ru/acts/bank/44731>.
2. Ригельман Р. Как избежать врачебных ошибок. — М.: Практика, 1994. 208 с.
3. Соловей М. Г. Развитие идей единства и целостности организма в отечественной медицине // Казанский медицинский ж., 1965. Т. 46. № 5. С. 86–89. doi: 10.17816/kazmj63550.

4. Румовская С. Б. Исследование методов поддержки принятия коллективных диагностических решений и разработка инструментальных средств «Виртуальный консилиум» (на примере диагностики артериальной гипертензии): Дис. . . . канд. техн. наук. — М., 2017. 138 с.
5. Колесников А. В., Кириков И. А. Методология и технология решения сложных задач методами функциональных гибридных интеллектуальных систем. — М.: ИПИ РАН, 2007. 387 с.
6. Андриевская Н. В., Резников А. С., Черанев А. А. Особенности применения нейро-нечетких моделей для задач синтеза систем автоматического управления // *Фундаментальные исследования*, 2014. № 11-7. С. 1445–1449.
7. Шляхто Е. В., Конради А. О. Персонализированная медицина. История, современное состояние проблемы и перспективы внедрения // *Российский ж. персонализированной медицины*, 2021. Т. 1. № 1. С. 6–20.
8. Зубарев П. Н., Косачев И. Д., Паскарь Т. В. Причины летальных исходов при остром деструктивном панкреатите // *Вестник Санкт-Петербургского университета. Медицина*, 2009. № 4. С. 161–168.
9. Богданов С. Н., Бабаев С. Ю., Стражнов А. В., Строганов А. Б., Митрофанова Н. Н., Мельников В. Л. Клинико-морфологическая классификация острого панкреатита и его осложнений // *Известия высших учебных заведений. Поволжский регион. Медицинские науки*, 2016. Т. 40. № 4. С. 109–121.
10. Клинические рекомендации — Острый панкреатит — 2020 (20.04.2021). Утверждены Минздравом РФ. 38 с. http://disuria.ru/_ld/9/999_kr20K85mz.pdf.
11. Шевляева М. А. Трудности ранней дифференциальной диагностики острого панкреатита // *Кубанский научный медицинский вестник*, 2013. № 3. С. 141–144.
12. Хаджибаев А. М., Алиджанов Ф. Б., Баймурадов Ш. Э., Акбарова Л. Р. Экспресс-метод определения острого панкреатита и панкреонекроза в условиях экстренной хирургии // *Вестник экстренной медицины*, 2018. Т. 11. № 1. С. 17–20.

Поступила в редакцию 15.08.22

MODELING OF THE DEVELOPMENT OF A COMPLETE EARLY DIFFERENTIAL DIAGNOSIS AND PROGNOSIS OF THE STATE OF COMPLEX OBJECTS IN MEDICINE. PROBLEMS AND PROSPECTS

S. B. Rumovskaya 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 problems of personalized assessment of the state (PSA) of a complex object (CO) as a heterogeneous system are considered. It is proposed to develop and investigate a new approach to improve the efficiency and quality

of solving the problems of medical diagnostics and prediction in conditions of significant diversity, heterogeneity, and uncertainty of information about a complex diagnostic object as well as the heterogeneity of the process of personalized assessment of their state. It is proposed to create and investigate methods and models of a new class of hybrid intelligent systems (HyIS) for medical diagnostics and prediction integrating heterogeneous models for the development and interpretation of PSA of a CO at different stages of the examination process as well as combining two types of HyIS (functional and transformational).

Keywords: hybrid intelligent system; intelligent decision support system; complex object of diagnostics; medical diagnostics and prognosis

DOI: 10.14357/08696527220303

References

1. O razvitii iskusstvennogo intellekta v Rossiyskoy Federatsii: Ukaz Prezidenta ot 10 oktyabrya 2019 g. No. 490 [About strategy of scientific and technological development of the Russian Federation. Presidential Decree No. 490 dated 10.10.2019]. Available at: <http://www.kremlin.ru/acts/bank/44731> (accessed October 11, 2022).
2. Rigelman, R. 1994. *Kak izbezhat' vrachebnykh oshibok* [How to avoid medical errors]. Moscow: Praktika. 208 p.
3. Solovey, M. G. 1965. Razvitie idey edinstva i tselostnosti organizma v otechestvennoy meditsine [Development of ideas of unity and integrity of the body in domestic medicine] *Kazanskiy meditsinskiy zh.* [Kazan Medical J.] 46(5):86–89. doi: 10.17816/kazmj63550.
4. Rumovskaya, S. B. 2017. Issledovanie metodov podderzhki prinyatiya kolektivnykh diagnosticheskikh resheniy i razrabotka instrumental'nykh sredstv "Virtual'nyy konsilium" (na primere diagnostiki arterial'noy gipertenzii) [The research of methods of collective diagnostic decision support and development of instruments of the "Virtual council" (illustrated with the diagnostic of arterial hypertension)]. Moscow. PhD Diss. 138 p.
5. Kolesnikov, A. V., and I. A. Kirikov. 2007. *Metodologiya i tekhnologiya resheniya slozhnykh zadach metodami funktsional'nykh gibridnykh intellektual'nykh sistem* [Methodology and technology for solving of complex problems using the methodology of functional hybrid artificial systems]. Moscow: IPI RAN. 387 p.
6. Andrievskaya, N. V., A. S. Reznikov, and A. A. Cheranov. 2014. Osobennosti primeniya neyro-nechetkikh modeley dlya zadach sinteza sistem avtomaticheskogo upravleniya [Features of the use of neuro-fuzzy models for the problems of synthesis of automatic control systems]. *Fundamental'nye issledovaniya* [Fundamental Research] 11-7:1445–1449.
7. Shlyakhto, E. V., and A. O. Konradi. 2021. Personalizirovannaya meditsina. Istoriya, sovremennoe sostoyanie problemy i perspektivy vnedreniya [Personalized medicine. History, current state and future directions]. *Rossiyskiy zh. personalizirovannoy meditsiny* [Russian J. Personalized Medicine] 1(1):6–20.
8. Zubarev, P. N., I. D. Kosachev, and T. V. Paskar. 2009. Prichiny letal'nykh iskhodov pri ostrom destruktivnom pankreatite [Causes of deaths in acute destructive pan-

- creatitis]. *Vestnik Sankt-Peterburgskogo universiteta. Meditsina* [J. St. Petersburg University. Medicine] 4:161–168.
9. Bogdanov, S.N., S.J. Babaev, A.V. Strazhnov, A.B. Stroganov, N.N. Mitrofanova, and V.L. Melnikov. 2016. Kliniko-morfologicheskaya klassifikatsiya ostrogo pankreatita i ego oslozhneniy [Clinical and morphological classification of acute pancreatitis and its complications]. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Meditsinskie nauki* [University Proceedings. Volga Region. Medical Sciences] 40(4):109–121.
 10. Ministerstvo zdravookhraneniya Rossiyskoy Federatsii. April 20, 2021. Klinicheskie rekomendatsii — Ostryy pankreatit [Clinical guidelines — Acute pancreatitis]. 38 p. Available at: http://disuria.ru/_ld/9/999_kr20K85mz.pdf (accessed August 29, 2022).
 11. Shevlyayeva, M.A. 2013. Trudnosti ranney differentsial'noy diagnostiki ostrogo pankreatita [The difficulties of the early differential diagnosis of acute pancreatitis]. *Kubanskiy nauchnyy meditsinskiy vestnik* [Kuban Scientific Medical Bulletin] 3:141–144.
 12. Khadjibaev, A.M., F.B. Alidjanov, S.E. Baymuradov, and L.R. Akbarova. 2018. Ekspress-metod opredeleniya ostrogo pankreatita i pankreonekroza v usloviyakh ekstremnoy khirurgii [Express method for determining acute pancreatitis and pancreatic necrosis in conditions of emergency surgery]. *Vestnik ekstremnoy meditsiny* [Bulletin Emergency Medicine] 11(1):17–20.

Received August 15, 2022

Contributors

Rumovskaya Sophiya B. (b. 1985) — Candidate of Science (PhD) in technology, 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; sophiyabr@gmail.com

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

МОДЕЛИ КЛАССИФИКАЦИИ ВЫБОРКИ ВЫЗВАННЫХ ПОТЕНЦИАЛОВ P300*

*А. М. Самохина¹, Р. Г. Нейчев², В. В. Гончаренко³, Р. К. Григорян⁴,
В. В. Стрижов⁵*

Аннотация: Решается задача определения объекта внимания пользователя. Исследуется выбор визуального стимула по электроэнцефалограмме (ЭЭГ) с выделенными в ней вызванными потенциалами, связанными с событием, — P300. Измеряются электрические потенциалы мозга во время наблюдения визуальных стимулов. Необходимо выбрать стимул, вызвавший наибольший отклик мозга. Определяется наличие потенциала P300 на отрезке ЭЭГ. Сравниваются модели классификации потенциалов. Предлагается способ обогащения выборки, повышающий качество классификации. Вычислительные эксперименты выполнены на наборе авторских данных потенциалов P300. Набор данных собран на 60 здоровых пользователях, которым предъявлялись визуальные стимулы, и выложен в открытый доступ.

Ключевые слова: классификация; электроэнцефалограмма; потенциал, связанный с событием; выбор модели; интерфейс мозг–компьютер

DOI: 10.14357/08696527220304

1 Введение

Решается задача построения интерфейса мозг–компьютер (ИМК) для определения объекта внимания пользователя по данным ЭЭГ. В [1] ИМК рассматриваются как альтернативный способ коммуникации с пациентами, больными тетраплегией или боковым амиотрофическим склерозом. В [2] адаптировано применение ИМК к здоровым пользователям в рекреационных целях, например в играх.

Для сбора информации о работе мозга при построении ИМК используются как методы, требующие операции для имплантации электродов, такие как электрокортикограмма, так и неинвазивные: функциональная спектроскопия в ближней инфракрасной области, функциональная магнитно-резонансная томография, электроэнцефалография.

*Работа выполнена при частичной поддержке РФФИ (проект 20-37-90050).

¹Московский физико-технический институт, alina.samokhina@phystech.edu

²Московский физико-технический институт, neychev@phystech.edu

³Московский физико-технический институт, vladislav.goncharenko@phystech.edu

⁴Московский государственный университет, grk@eegdude.ru

⁵Федеральный исследовательский центр «Информатика и управление» Российской академии наук, strijov@phystech.edu

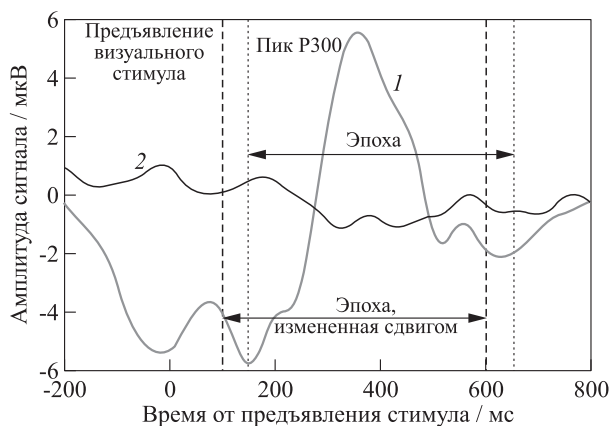


Рис. 1 Вызванный потенциал P300, связанный с событием, (1) и ЭЭГ без P300 (2)

Интерфейсы на основе ЭЭГ имеют низкую стоимость и высокое разрешение метода [3]. Широко используется система размещения электродов 10–10, состоящая из равномерно распределенных по поверхности головы 64 электродов и покрывающая все области мозга. Существуют три типа ИМК: активный, требующий концентрации на специфической для конкретного ИМК задаче; пассивный, не требующий выполнения дополнительных действий; реактивный, считывающий изменения в ЭЭГ пользователя, связанные со специфическими для ИМК изменениями внешней среды, вызывающими реакции мозга, — *стимулами*. Стимулы имеют различную модальность: слуховую, тактильную и зрительную.

Авторы рассматривают реактивный ИМК на основе *вызванных потенциалов, связанных с событием*, — изменения амплитуды электрической активности мозга в ответ на стимул. Рассматриваемые потенциалы — *потенциалы P300*, вызванные визуальными стимулами, — показаны на рис. 1. Пользователю на экране показываются несколько визуальных объектов. Изменение яркости объекта или начало его движения называется *предъявлением стимула*. Пользователь обращает свое внимание на один из объектов при начале его движения, а в ответ на предъявление стимула в ЭЭГ регистрируется увеличение амплитуды сигнала электрической активности мозга. Максимуму амплитуды соответствует пик P300. Для визуально вызванных потенциалов наибольшую амплитуду пик P300 имеет в затылочной доле, где находится зрительная кора головного мозга. Отрезок ЭЭГ после предъявления каждого стимула называется *эпохой*. Потенциал P300 присутствует в эпохах, соответствующих целевому стимулу, — предъявлению стимула на объекте внимания пользователя.

В данной работе предложен эксперимент по записи данных ЭЭГ во время игры «Еноты и демоны». Рассмотрена игра в рамках активного ИМК-P300, где



Рис. 2 Сцена из игры

стимул — начало движения одного из визуальных объектов, представленных на рис. 2. Игра разбита на два этапа: обучение и активное участие. Обучение необходимо для демонстрации пользователю принципов управления ИМК и сбора обучающей выборки модели классификации. В вычислительном эксперименте, приведенном в статье, использована авторская выборка потенциалов $R300$ [4], полученных в этом эксперименте по записи данных во время игры.

По данным эпох ЭЭГ проводится мультиклассовая классификация для определения объекта внимания пользователя среди представленных визуальных стимулов. Цель данной работы — повышение качества классификации.

Решение задачи мультиклассовой классификации основывается на агрегированных результатах бинарной классификации, где бинарная классификация определяет наличие потенциала $R300$ в эпохе. Общее число объектов — пять, но объект внимания пользователя только один. Эпох с потенциалом $R300$ в 4 раза меньше, чем без него. В бинарной задаче присутствует дисбаланс классов.

Для решения проблемы дисбаланса классов предлагается проводить аугментацию данных. Здесь *аугментация данных* — обогащение выборки сдвигом или интерполяцией присутствующих в выборке эпох. После применения аугментации классы для обучения бинарного классификатора сбалансированы. Предполагается, что это повышает качество как бинарной, так и мультиклассовой классификации.

В работах по классификации ЭЭГ применяют различные методы аугментации. В [5] используются пространственные аугментации: варьируется расположение электродов. Сигнал ЭЭГ моделируется с учетом пространственного расположения электродов на голове. Интерполируется сдвиг электродов в пространстве для получения аугментации данных эпох ЭЭГ. Для выборки из 6 чел.

получен прирост мультиклассовой точности около 2,5%. В работе [6] на амплитуды сигнала после преобразования Фурье накладывается гауссовский шум, затем обратное преобразование Фурье возвращает к исходному формату данных. Для задачи представления движения при исследовании выборки из 9 чел. получен прирост точности в 2,3%. В [7] к задаче классификации P300 применяются алгоритмы, основанные на методе синтетической передискретизации меньшинства, SMOTE (synthetic minority oversampling technique) [8]. Для выборки из 44 пользователей при их применении получено 1,3% прироста точности. Вариации SMOTE: SMOTE, Borderline-SMOTE (bSMOTE), Adaptive Synthetic (ADASYN) и Support Vector Machine SMOTE (SVM-SMOTE) [9] — отличаются методами выбора опорного для передискретизации экземпляра эпохи. В [10] рассматриваются аугментации сдвигом времени пика P300 относительно предъявления стимула. Эпоха отсчитывается не от момента предъявления стимула, а с опозданием. Для выборки из 5 пользователей прирост мультиклассовой точности составил 1%. В данной работе авторами рассматриваются два метода аугментаций. Первый метод — SMOTE и его модификации. Второй — сдвиг ЭЭГ относительно стимулов.

2 Процедура построения выборки

Эксперимент проведен на 60 здоровых пользователях, из которых 23 — мужчины без опыта работы с ИМК. Возраст пользователей — от 19 до 45, средний возраст — 28 лет. Все пользователи соответствовали необходимым условиям по состоянию здоровья. Электроэнцефалограммы записывались с помощью энцефалографа NVX-52 с частотой 500 Гц. Для записи использовались 8 губчатых электродов. Стимулы предъявлялись с помощью шлема HTC Vive Pro VR.

Пользователям была предложена игра в виртуальной реальности на управлении с помощью ИМК-P300, разбитая на два этапа: обучение и активное участие. Обучение составляло 5 игровых актов. Акт — последовательность блоков с общим объектом внимания. Число актов в экспериментах — 5 при обучении и 10 на этапе активного участия. Блок — последовательность эпох с общим объектом внимания, где каждый стимул предъявляется один раз, порождая одну эпоху (рис. 3).

Для определения потенциалов P300 в игре пользователю предъявляются визуальные стимулы. В проведенном эксперименте предъявлялись 5 разнесенных в пространстве объектов с разнесенным во времени началом движения — предъявлением стимула. Считается, что объектом внимания может быть только один визуальный объект. Пользователь концентрирует внимание на одном из 5 объектов. Поочередно в случайном порядке с разницей во времени в 450 мс происходит предъявление каждого стимула. Если предъявляется стимул, соответствующий объекту внимания пользователя, в ЭЭГ регистрируется потенциал P300.

Под номерами 1–5 на рис. 2 изображены демоны, служившие объектами во время этапа активного участия. Под номерами 1–3 изображены еноты,

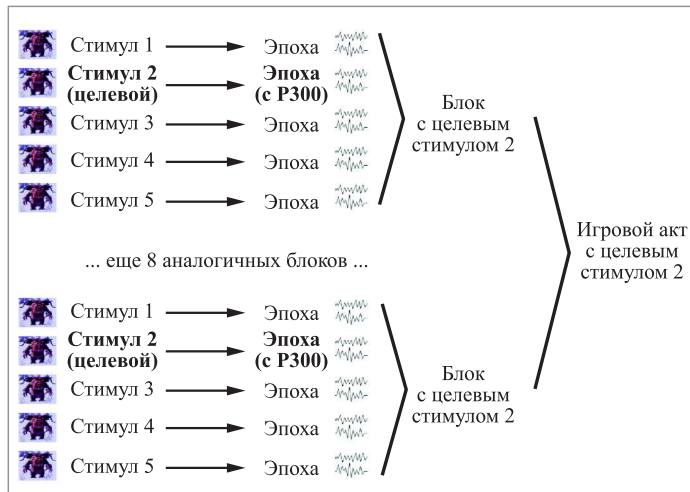


Рис. 3 Иерархическая структура данных

служившие объектами на этапе обучения. На этапе обучения для тренировки модели пользователь смотрел на заранее определенный объект из 1–3 на рис. 2. На этапе активного участия пользователь смотрел на одного из демонов — произвольный объект из 1–5 на рис. 2, а модель определяла, на какой из объектов смотрел пользователь.

3 Постановка задачи классификации

Требуется определить наличие на отрезке ЭЭГ потенциала P300. Задан набор из M измерений ЭЭГ $\mathbf{X} \in \mathbb{R}^{N \times T}$, где N — число электродов; T — число элементов временного ряда, и целевая переменная $Y \in \{0, 1\}$, указывающая на отсутствие или наличие потенциала. Требуется выбрать модель классификации $f(\mathbf{w}, \mathbf{x})$, аппроксимирующую целевую переменную y по описанию \mathbf{x} , полученному по измерениям \mathbf{X} , и оптимизировать ее параметры. Критерием качества модели служит бинарная кросс-энтропия

$$L(\mathbf{w}) = -\frac{1}{M} \sum_{m=1}^M [y_m \log f(\mathbf{w}, \mathbf{x}) + (1 - y_m) \log(1 - f(\mathbf{w}, \mathbf{x}))] - c \|\mathbf{w}\|_2,$$

где c — фиксированный коэффициент регуляризации. Для оптимизации параметров модели классификации решается задача

$$\hat{\mathbf{w}} = \arg \max_{\mathbf{w}} L(\mathbf{w}).$$

Для построения описания \mathbf{x} матрица \mathbf{X} векторизуется одним из двух способов. В первом способе ряды матрицы, временные ряды для каждого канала, записываются друг за другом. Во втором способе используется расширенная матрица ковариации: матрица ковариации потенциалов, связанных с событием, ERPCov [11]. Для одного измерения ЭЭГ \mathbf{X}_m искомая матрица ковариации

$$\mathbf{C}_m = \frac{1}{T-1} \mathbf{P}_m \mathbf{P}_m^T, \quad \mathbf{P}_m = \begin{bmatrix} \mathbf{P}_0 \\ \mathbf{P}_1 \\ \mathbf{X}_m \end{bmatrix}, \quad \mathbf{C}_m \in \mathbb{R}^{3N \times 3N},$$

где \mathbf{P}_0 и \mathbf{P}_1 — средние значения для классов $\{0, 1\}$,

$$\mathbf{P}_c = \frac{\sum_{m=1}^M [y_m = c] \mathbf{X}_m}{\sum_{m=1}^M [y_m = c]}, \quad c \in \{0, 1\}.$$

Пространство матриц ковариации является римановым многообразием [11]. В каждой точке риманова многообразия существует касательная плоскость с определенным на ней скалярным произведением. Общая касательная плоскость для отображения всех матриц ковариации выборки строится в точке среднего геометрического по римановой метрике известных ковариационных матриц

$$\mathbf{C} = \mathbf{C}(\mathbf{C}_1, \dots, \mathbf{C}_M) = \underset{\mathbf{C}}{\operatorname{argmin}} \sum_{m=1}^M \delta_R^2(\mathbf{C}, \mathbf{C}_m),$$

$$\delta_R(\mathbf{C}, \mathbf{C}_m) = \|\log(\mathbf{C}^{-1} \mathbf{C}_m)\|_F = \left[\sum_{i=1}^{3N} \log^2 \lambda_i \right]^{1/2},$$

где λ_i — собственные числа матрицы $\mathbf{C}^{-1} \mathbf{C}_m$.

Согласно [11], для каждой ковариационной матрицы \mathbf{C}_m существует проекция \mathbf{S}_m в касательное пространство

$$\mathbf{S}_m = \mathbf{C}^{1/2} \tilde{\mathbf{S}}_m \mathbf{C}^{1/2}, \quad \tilde{\mathbf{S}}_m = \log \left(\mathbf{C}^{-1/2} \mathbf{C}_m \mathbf{C}^{-1/2} \right),$$

векторное представление которой в пространстве с евклидовой метрикой выражается как записанные последовательно элементы верхней треугольной матрицы от $\tilde{\mathbf{S}}_m$ с коэффициентами 1 для диагональных элементов и коэффициентом $\sqrt{2}$ для недиагональных:

$$\mathbf{x} = \left[\tilde{\mathbf{S}}_{m_{1,1}}, \sqrt{2} \tilde{\mathbf{S}}_{m_{1,2}}, \dots, \sqrt{2} \tilde{\mathbf{S}}_{m_{1,3N}}, \tilde{\mathbf{S}}_{m_{2,2}}, \dots, \tilde{\mathbf{S}}_{m_{3N,3N}} \right], \quad \mathbf{x} \in \mathbb{R}^{3N(3N-1)/2}.$$

После векторизации \mathbf{X}_m одним из предложенных способов, решается задача классификации потенциалов. Внешними критериями качества модели служат точность, полнота, F1-score и AUC. Для игрового процесса требуется по итогу акта игры принять решение о выборе объекта, на который было обращено внимание пользователя. В игровой механике выбирается стимул с наибольшей суммой рассмотренных бинарных вероятностей. Предполагается, что если в течение акта игры пользователь обращал внимание на определенный стимул, то в большинстве эпох у этого стимула будет высокая вероятность наличия потенциала P300. Мультиклассовая точность — доля верно определенных таким образом объектов внимания пользователя. Данный критерий качества также является внешним и используется для определения качества агрегации бинарных предсказаний и финального выбора объекта в игровом процессе.

Модели классификации потенциалов:

1. Линейная модель логистической регрессии (LR, logistic regression) для решения задачи классификации на матрице, векторизованной первым способом:

$$f(\mathbf{w}, \mathbf{x}) = \frac{1}{1 + \exp(-\mathbf{w}^\top \mathbf{x})}. \quad (1)$$

2. Модель опорных векторов (SVM, support vector machine) на матрице, векторизованной первым способом.
3. Линейная модель дискриминантного анализа (LDA, linear discriminant analysis) на матрице, векторизованной первым способом:

$$f(\mathbf{w}, \mathbf{x}) = \mathbf{x}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \frac{1}{2} \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu},$$

$$\boldsymbol{\Sigma} = \sum_{m=1}^M (\mathbf{x}_m - \boldsymbol{\mu})^\top (\mathbf{x}_m - \boldsymbol{\mu}), \quad \boldsymbol{\mu} = \frac{1}{M} \sum_{m=1}^M \mathbf{x}_m.$$

4. Модель логистической регрессии (1) на матрице, векторизованной вторым способом (ERPCov TS LR).
5. Линейная модель (LDA) на матрице, полученной после пространственной фильтрации алгоритмом XDawn [12] и векторизованной первым способом (XDawn LDA).
6. Линейная модель (LR) на матрице, полученной после пространственной фильтрации исходного сигнала алгоритмом XDawn [12] и векторизованной вторым способом (XDawnCov TS LR).

4 Вычислительный эксперимент

Предобработка данных электроэнцефалограмм. Согласно [11], записи ЭЭГ обрабатывались в шесть этапов:

- (1) разделение ЭЭГ на эпохи;
- (2) децимация со сглаживающим фильтром;
- (3) полосовой фильтр Баттерворта;
- (4) передискретизация сигнала с помощью линейной интерполяции;
- (5) клиппинг зашкаливающих значений;
- (6) поканальная нормировка — вычитание среднего и деление на стандартное отклонение всех значений каждого канала.

Авторами были проведены все шаги предобработки данных. Для шагов 1 и 3 были подобраны оптимальные для рассматриваемой выборки гиперпараметры из рассмотренных диапазонов значений согласно табл. 1.

Таблица 1 Значения параметров предобработки

Характеристики	Начало эпохи, с	Конец эпохи, с	Нижняя частота, Гц	Верхняя частота, Гц
Начальное значение	0,1	0,7	0,5	20
Диапазон значений	0,00–0,35	0,45–0,90	0–1	4–22
Оптимальное значение	0,1	0,7	0,5	12,5

Качество моделей. При фиксированных гиперпараметрах предобработки данных для каждой модели было оценено качество классификации. Обучение моделей проводилось на данных обучающего этапа игры. Значения критериев качества моделей приводятся для той части данных, которая была собрана на этапе активной игры.

Аугментация выборки. В работе исследовались два способа аугментации выборки: аугментация выборки интерполяцией с применением метода синтетической передискретизации меньшинства и аугментация выборки сдвигом.

Метод передискретизации меньшинства. В алгоритмах, основанных на SMOTE, новый экземпляр эпохи строится в случайной точке на отрезке между существующими экземплярами. В качестве начала отрезка выбирается один опорный экземпляр из менее представленного класса. Из экземпляров менее представленного класса, соседних опорному, случайным образом выбирается один, который назначается концом отрезка. По результатам вычислительных

Таблица 2 Результаты аугментации выборки методом передискретизации

Модель	Ауг.	Кл.	Мультикл. точность	Точность	Полнота	F1-score	AUC
LR	border	0,4	0,68±0,23	0,36±0,10	0,60±0,11	0,40±0,07	0,62±0,07
LDA	adasyn	0,8	0,71±0,23	0,40±0,12	0,44±0,13	0,40±0,11	0,63±0,07
SVM	adasyn	0,8	0,69±0,24	0,44±0,18	0,41±0,13	0,36±0,11	0,60±0,07
ERPCov TS LR	adasyn	0,8	0,77±0,24	0,46±0,16	0,60±0,14	0,44±0,11	0,65±0,09
Xdawn LDA	adasyn	0,8	0,70±0,23	0,42±0,13	0,47±0,13	0,42±0,11	0,64±0,08
XdawnCov TS LDA	border	0,4	0,71±0,26	0,47±0,17	0,35±0,16	0,38±0,14	0,62±0,08

экспериментов оптимальным числом ближайших экземпляров, считающихся соседними, оказались 5 экземпляров. Параметрами алгоритма служат метод выбора опорного экземпляра и конечное соотношение классов [8].

По результатам вычислительного эксперимента лучшими методами выбора опорного экземпляра оказались граничный (border) и адаптивный (adasyn) варианты SMOTE. Для граничного метода опорный экземпляр выбирается из множества экземпляров менее представленного класса, для которых среди ближайших соседей представлены как экземпляры класса меньшинства, так и экземпляры класса большинства. Для адаптивного метода выбор опорного экземпляра основывается на соотношении экземпляров менее представленного класса и экземпляров более представленного класса среди ближайших соседей. Чем выше отношение числа экземпляров класса большинства к классу меньшинства, тем больше новых экземпляров будут использовать эту точку как опорную. В вычислительном эксперименте варьировалось финальное соотношение классов от исходного 0,25 до 1,0 и число соседей для выбора конца отрезка от 2 до 10.

Для каждой модели выбран вариант метода с лучшим приростом качества. Результаты представлены в табл. 2. Столбец Ауг. указывает на тип аугментации, столбец Кл. — на соотношение мощностей классов. Аугментация выборки на основе SMOTE дает прирост мультиклассовой точности до 3 процентных пунктов. Бинарные метрики в случаях некоторых моделей улучшаются на величину до 15 процентных пунктов.

Метод сдвига. Время от активации стимула до максимума амплитуды потенциала P300 — *латентность*. Оно варьируется как между разными людьми, так и между разными актами игры одного человека [13]. Предлагается обогащать выборку, варьируя время старта и, следовательно, латентность эпох исходной выборки.

Аугментация выборки сдвигом строит несколько новых экземпляров X_m по одной оригинальной эпохе. Границы новых эпох сдвинуты относительно исходных влево или вправо по оси времени (см. рис. 1). Начало исходной эпохи отсчитывается от момента предъявления стимула. Начало измененной эпохи

Таблица 3 Качество моделей на исходной выборке

Модель	Мультикл. точность	Точность	Полнота	F1-score	AUC
LR	0,68±0,24	0,35±0,09	0,44±0,13	0,39±0,10	0,62±0,07
LDA	0,68±0,24	0,40±0,13	0,28±0,13	0,32±0,13	0,59±0,06
SVM	0,67±0,25	0,37±0,18	0,15±0,13	0,20±0,14	0,55±0,05
ERPCov TS LR	0,77±0,25	0,40±0,12	0,47±0,14	0,42±0,12	0,64±0,08
Xdawn LDA	0,69±0,23	0,40±0,13	0,35±0,15	0,37±0,13	0,61±0,07
XdawnCov TS LDA	0,71±0,27	0,46±0,16	0,31±0,16	0,36±0,15	0,61±0,08

Таблица 4 Качество моделей при использовании аугментации сдвигом

Модель	Мультикл. точность	Точность	Полнота	F1-score	AUC
LR	0,76±0,24	0,43±0,10	0,60±0,13	0,50±0,11	0,66±0,09
LDA	0,76±0,23	0,40±0,09	0,67±0,12	0,50±0,09	0,66±0,08
SVM	0,77±0,24	0,41±0,09	0,64±0,13	0,50±0,10	0,66±0,09
ERPCov TS LR	0,78±0,23	0,42±0,10	0,57±0,14	0,48±0,11	0,65±0,09
Xdawn LDA	0,73±0,26	0,40±0,09	0,65±0,12	0,50±0,10	0,66±0,09
XdawnCov TS LDA	0,78±0,25	0,40±0,09	0,64±0,14	0,49±0,10	0,66±0,09

на рис. 1 — за 50 мс до момента предъявления стимула. Параметры данного метода — число новых экземпляров, построенных по одной эпохе, и отступ от момента предъявления стимула.

Первым шагом определяется оптимальная величина сдвига. Подход аналогичен подбору оптимальных параметров предобработки ЭЭГ. Из диапазона от 5 до 40 мс выбирается оптимальный отступ от оригинальной эпохи. По результатам эксперимента оптимальный сдвиг составил 30 мс. Рассматривается оптимальное число аугментаций в границах $[-30, 30]$ мс. По результатам эксперимента получены четыре новые эпохи.

В табл. 3 представлено качество моделей на исходной выборке, а в табл. 4 — после применения оптимальной аугментации. Аугментации стартов по схеме $[-30, -15, 15, 30]$ повышают качество моделей.

5 Заключение

Поставлена задача повышения качества классификации при определении объекта внимания пользователя. Предложенное решение — аугментация выборки сдвигом эпох — значительно повышает качество классификации рассмотренных моделей. Проведено сравнение моделей классификации. По полученным значениям внешних критериев качества сделаны выводы о том, что использование расширенных матриц ковариации потенциалов, связанных с событием,

значительно повышает качество модели. При использовании одновременно аугментации выборки и расширенных матриц ковариации точность мультиклассовой классификации модели логистической регрессии повышается на 10 процентных пунктов. В рамках решения данной задачи был записан набор данных вызванных потенциалов P300 [4].

Литература

1. *Cipresso P., Carelli L., Solca F., Meazzi D., Meriggi P., Poletti B., Lulé D., Albert C. L., Silani V., Riva G.* The use of P300-based BCIs in amyotrophic lateral sclerosis: From augmentative and alternative communication to cognitive assessment // *Brain Behav.*, 2012. Vol. 2. Iss. 4. P. 479–498.
2. *Kaplan A. Y., Shishkin S. L., Ganin I. P., Basyul I. A., Zhigalov A. Y.* Adapting the P300-based brain–computer interface for gaming: A review // *IEEE T. Comp. Intel. AI*, 2013. Vol. 5. Iss. 2. P. 141–149.
3. *Rashid M., Sulaima N., Majeed A. P. P. A., Musa R. M., Nasir A. F. A., Bari B. F., Khatun S.* Current status, challenges, and possible solutions of EEG-based brain–computer interface: A comprehensive review // *Front. Neurorobotics*, 2020. Vol. 14. Art. No. 25. 35 p. doi: 10.3389/fnbot.2020.00025.
4. *Goncharenko V. V.* Racoons vs. demons P300 ERP dataset of Neiry, 2021. <https://gin.g-node.org/v-goncharenko/neiry-demons>.
5. *Krell M. M., Kim S. K.* Rotational data augmentation for electroencephalographic data // 39th Annual Conference (International) of the IEEE Engineering in Medicine and Biology Society Proceedings. — Piscataway, NJ, USA: IEEE, 2017. P. 471–474. doi: 10.1109/EMBC.2017.8036864.
6. *Zhang X., Lei M. Y., Li Y.* An amplitudes-perturbation data augmentation method in convolutional neural networks for EEG decoding // 5th Conference (International) on Information, Cybernetics, and Computational Social Systems Proceedings. — Piscataway, NJ, USA: IEEE, 2018. P. 231–235. doi: 10.1109/ICCSS.2018.8572304.
7. *Lee T., Kim M., Kim S. P.* Data augmentation effects using borderline-SMOTE on classification of a P300-based BCI // 8th Winter Conference (International) on Brain–Computer Interface Proceedings. — Piscataway, NJ, USA: IEEE, 2020. Art. ID: 9061656. 4 p. doi: 10.1109/BCI48061.2020.9061656.
8. *Chawla N. V., Bowyer K. W., Hall L. O., Kegelmeyer W. P.* SMOTE: Synthetic minority over-sampling technique // *J. Artif. Intell. Res.*, 2002. Vol. 16. P. 321–357.
9. *Lemaître G., Nogueira F., Aridas C. K.* Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning // *J. Mach. Learn. Res.*, 2017. Vol. 18. Iss. 17. P. 1–5.
10. *Krell M. M., Seeland A., Kim S.* Data augmentation for brain–computer interfaces: Analysis on event-related potentials data. Working papers by Cornell University, 2018. arXiv:1801.02730 [cs.CV].
11. *Barachant A., Bonnet S., Congedo M., Jutten C.* Multiclass brain–computer interface classification by Riemannian geometry // *IEEE T. Bio-med. Eng.*, 2012. Vol. 59. Iss. 4. P. 920–928.

12. Rivet B., Souloumiac A., Attina V., Gibert G. xDAWN algorithm to enhance evoked potentials: Application to brain-computer interface // IEEE T. Bio-med. Eng., 2009. Vol. 56. Iss. 8. P. 2035–2043.
13. Picton T.W. The P300 wave of the human event-related potential // J. Clin. Neurophysiol., 1992. Vol. 9. Iss. 4. P. 456–479.

Поступила в редакцию 16.01.22

CLASSIFICATION MODELS FOR P300 EVOKED POTENTIALS

*A. M. Samokhina¹, R. G. Neychev¹, V. V. Goncharenko¹, R. K. Grigoryan²,
and V. V. Strijov³*

¹Moscow Institute of Physics and Technology, 9 Institutskiy Per., Dolgoprudny, Moscow Region 141700, Russian Federation

²M. V. Lomonosov Moscow State University, 1 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation

³Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

Abstract: The paper is devoted to the problem of user’s attention detection. It investigates the choice of a visual stimulus by the electroencephalogram (EEG) with the evoked potentials related to the event, P300, highlighted in it. The electrical brain potentials are measured while the user is observing visual stimuli. The goal is to select a stimulus which causes the maximum brain response. A classification model detects if there is a P300 potential in an EEG segment. Various classification models for event-related potentials are compared. The paper proposes a method of data augmentation to improve the quality of classification. Computational experiments use an original real-world dataset of P300 potentials. This dataset was collected on 60 healthy users who are presented with visual stimuli. It is released to the public access.

Keywords: classification; electroencephalogram; event-related potential; model selection; brain–computer interface

DOI: 10.14357/08696527220304

Acknowledgments

This research was partially supported by RFBR (project 20-37-90050).

References

1. Cipresso, P., L. Carelli, F. Solca, D. Meazzi, P. Meriggi, B. Poletti, D. Lule, C. L. Albert, V. Silani, and G. Riva. 2012. The use of P300-based BCIs in amyotroph-

- ic lateral sclerosis: From augmentative and alternative communication to cognitive assessment. *Brain Behav.* 2(4):479–498.
2. Kaplan, A. Y., S. L. Shishkin, I. P. Ganin, I. A. Basyul, and A. Y. Zhigalov. 2013. Adapting the P300-based brain-computer interface for gaming: A review. *IEEE T. Comp. Intel. AI* 5(2):141–149.
 3. Rashid, M., N. Sulaima, A. P. P. A. Majeed, R. M. Musa, A. F. A. Nasir, B. F. Bari, and S. Khatun. 2020. Current status, challenges, and possible solutions of EEG-based brain-computer interface: A comprehensive review. *Front. Neurorobotics* 14:25. 35 p. doi: 10.3389/fnbot.2020.00025.
 4. Goncharenko, V. 2021. Racoons vs. demons P300 ERP dataset of Neiry. Available at: <https://gin.g-node.org/v-goncharenko/neiry-demons> (accessed August 18, 2022).
 5. Krell, M. M., and S. K. Kim. 2017. Rotational data augmentation for electroencephalographic data. *39th Annual Conference (International) of the IEEE Engineering in Medicine and Biology Society Proceedings*. Piscataway, NJ: IEEE. 471–474. doi: 10.1109/EMBC.2017.8036864.
 6. Zhang, X., M. Y. Lei, and Y. Li. 2018. An amplitudes-perturbation data augmentation method in convolutional neural networks for EEG decoding. *5th Conference (International) on Information, Cybernetics, and Computational Social Systems Proceedings*. Piscataway, NJ: IEEE. 231–235. doi: 10.1109/ICCSS.2018.8572304.
 7. Lee, T., M. Kim, and S. P. Kim. 2020. Data augmentation effects using borderline-SMOTE on classification of a P300-based BCI. *8th Winter IEEE Conference (International) on Brain-Computer Interface Proceedings*. Piscataway, NJ: IEEE. 9061656. 4 p. doi: 10.1109/BCI48061.2020.9061656.
 8. Chawla, N. V., K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. 2002. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* 16:321–357.
 9. Lematre, G., F. Nogueira, and C. K. Aridas. 2017. Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning. *J. Mach. Learn. Res.* 18(17):1–5.
 10. Krell, M. M., A. Seeland, and S. Kim. 2018. Data augmentation for brain-computer interfaces: Analysis on event-related potentials data. Working papers by Cornell University. arXiv.org. Available at: <https://arxiv.org/abs/1801.02730> (accessed August 18, 2022).
 11. Barachant, A., S. Bonnet, M. Congedo, and C. Jutten. 2012. Multiclass brain-computer interface classification by Riemannian geometry. *IEEE T. Bio-med. Eng.* 59(4):920–928.
 12. Rivet, B., A. Souloumiac, V. Attina, and G. Gibert. 2009. xDAWN algorithm to enhance evoked potentials: Application to brain-computer interface. *IEEE T. Bio-med. Eng.* 56(8):2035–2043.
 13. Picton, T. W. 1992. The P300 wave of the human event-related potential. *J. Clin. Neurophysiol.* 9(4):456–479.

Received January 16, 2022

Contributors

Samokhina Alina M. (b. 1996) — PhD student, Moscow Institute of Physics and Technology, 9 Institutskiy Per., Dolgoprudny, Moscow Region 141701, Russian Federation; alina.samokhina@phystech.edu

Neychev Radoslav G. (b. 1994) — PhD student, Moscow Institute of Physics and Technology, 9 Institutskiy Per., Dolgoprudny, Moscow Region 141701, Russian Federation; neychev@phystech.edu

Goncharenko Vladislav V. (b. 1994) — PhD student, Moscow Institute of Physics and Technology, 9 Institutskiy Per., Dolgoprudny, Moscow Region 141701, Russian Federation; vladislav.goncharenko@phystech.edu

Grigoryan Rafael K. (b. 1992) — PhD student, Faculty of Biology, M. V. Lomonosov Moscow State University, 1 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation; neuroscientist, Laboratory for Neurophysiology and Neuro-Computer Interfaces, Neiry Ltd., 73-2 Bakuninskaya Str., Moscow 105082, Russian Federation; grk@eegdude.ru

Strijov Vadim V. (b. 1967) — Doctor of Science in physics and mathematics, leading scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; strijov@phystech.edu

ОБЗОР ПОДХОДОВ К ПРОСТРАНСТВЕННО-ВРЕМЕННОМУ МОДЕЛИРОВАНИЮ И ВЫЯВЛЕНИЕ ОСНОВНЫХ ТЕНДЕНЦИЙ РАЗВИТИЯ T-GIS

*Д. А. Никишин*¹

Аннотация: Цель данной работы — рассмотрение характерных особенностей основных подходов и соответствующих им пространственно-временных (темпоральных) моделей геоданных, предложенных в последние десятилетия, а также выявление основных тенденций развития пространственно-временного моделирования для темпоральных геоинформационных систем (T-GIS, temporal geoinformation systems). Отмечено, что говорить об общем подходе к пространственно-временному гео моделированию не приходится ввиду разнообразия требований и особенностей, присущих специфическим задачам конкретных T-GIS. Важными направлениями представляются моделирование событий и процессов на основе графов, агентно-ориентированный подход, использование парадигмы движущихся объектов, а также интеграция этих подходов воедино.

Ключевые слова: моделирование географических явлений; эволюция геообъектов; темпоральная геоинформационная система; пространственно-временная база геоданных; классификация темпоральных геомоделей

DOI: 10.14357/08696527220305

1 Введение

Использование временной (темпоральной) составляющей, позволяющей регистрировать динамику изменения свойств местности в целом или ее отдельных геообъектов и таким образом моделировать протекающие на местности явления и процессы, стало важным и перспективным направлением развития геоинформационных систем (ГИС), конвенционально именуемым T-GIS. Отличительной чертой T-GIS является сохранение в базе геоданных (БГД) истории изменений как семантических, так и метрических свойств и, соответственно, возможность восстанавливать состояние местности на любой необходимый для анализа момент времени.

Цель данной статьи — рассмотрение наиболее характерных черт пространственно-временных моделей геоданных для T-GIS, предложенных в последние десятилетия, и на этой основе проследить основные тенденции развития пространственно-временного моделирования для T-GIS.

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, dmnikishin@mail.ru

В качестве основы для данной работы был использован весьма объемный обзор [1], который классифицирует 186 предложений (подходов) по пространственно-временному моделированию, выделенных при анализе более 1450 статей. Некоторые из подходов являются сугубо теоретическими моделями, другие основаны на конкретных тематических исследованиях, таких как эволюция землепользования, дорожной сети или динамика лесных пожаров. В разд. 2–8 рассмотрены наиболее характерные особенности некоторых из этих предложений. В разд. 9 выделены основные тенденции развития Т-GIS.

2 Ранние модели и предложения Т-GIS

Первые исследования временных (темпоральных) аспектов пространственных систем появились в начале 1980-х гг., а моделирование и интеграция фактора времени в ГИС стали предметом исследований с конца 1980-х гг. (см., например, [2]) и продолжают до настоящего времени.

Исторически для обработки и анализа пространственно-временных данных в ГИС наиболее очевидным решением был подход моментальных снимков (единовременных образов состояния БГД). Здесь не учитывалась природа вызывающих изменения географических процессов, а тем более факторы, обуславливающие эти процессы.

На рис. 1 схематично показан процесс актуализации посредством серии «моментальных снимков» в моменты t_0, \dots, t_5 , стрелками показана преемственность между разновременными моделями отдельных геообъектов. В том числе показаны события, затрагивающие несколько геообъектов: разделение одного объекта S_{m2} на две части S_{m3} и S_{n3} и слияние двух объектов S_{13} и S_{23} в совокупный объект S_{14} . Прочие переходы (вида $S_{ij} \rightarrow S_{i(j+1)}$) подразумевают изменения (как семантики, так и метрики) в контексте одного геообъекта S_i .

Трехдоменная модель (TDM, three domain model) Юаня [3] представляла пространственные, временные и семантические информационные объекты

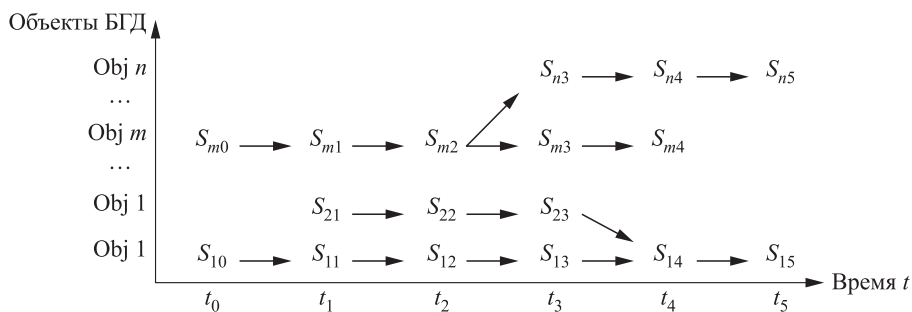


Рис. 1 Схема процесса эволюции ситуации в виде «моментальных снимков» БГД

и ссылки между ними, что позволяло описывать географические явления и процессы, используя только одну запись атрибута для отдельных местоположений с общими свойствами. Это представление, разработанное в контексте исследований лесных пожаров, использовало парадигму «моментального снимка» и зачатки пространственно-временной (ST, spatiotemporal) объектной модели (ST-объекты с отметками времени). Трехдоменная модель предлагалась в качестве подхода для ответа на запросы о моментах и периодах (диапазонах) времени. По сравнению с существовавшими на то время моделями TDM позволяла управлять изменениями одновременно в трех аспектах: семантическом, местоположения (пространственной локализации) и временном.

3 Моделирование на основе событий

Событийно-ориентированный подход «фокусируется на динамическом событии в целом, а не только на времени события» [4]. Другими словами, такой подход помимо отметки момента изменения позволяет различать атрибуты и отношения событий наряду с атрибутами и отношениями объектов. Ключевыми понятиями для моделирования динамических явлений выступают: объект, состояние, процесс и событие [5]. Было предложено большое число пространственно-временных моделей, рассматривающих события и процессы.

Подход к анализу пространственно-временных данных, основанный на времени (time-based analysis of spatiotemporal data), был предложен в [6] как дополнение к подходам на основе объектно-ориентированного анализа и анализа местоположения. Начиная с начального состояния, называемого базовой картой, события записывались в порядке их возникновения, причем каждое событие связывалось со списком изменений, произошедших с момента последнего обновления вектора событий. При обширных изменениях могла быть зарегистрирована новая полная карта области. Этот подход был подтвержден прототипом TEMPEST [7].

Впоследствии была представлена модель ESTDM (Event-oriented Spatio-Temporal Data Model — событийно-ориентированная пространственно-временная модель данных) на основе растра [8]. В ней набор компонентов кодировал местоположения, в которых происходят изменения, и значения атрибутов во время события — таким способом представлялась информация об изменениях в определенных ячейках растра. Модель показала свою эффективность для пространственных и временных запросов. Значимой стала возможность выполнять временные манипуляции с данными (например, изменение временного масштаба) и последовательные сравнения на основе времени. Однако ориентация ESTDM на растр не позволяет сохранять идентичность объекта за пределами исходного местоположения и представлять такие явления, как перемещение, разделение, слияние объекта. Эта модель «сосредоточена на эволюции непрерывных, протяженных, обширных областей, а не на конкретных объектах и их свойствах» [9].

Гибридная пространственно-временная модель и структура данных (Hybrid Spatio-Temporal Data Model and Structure, HST-DMS) [10] стала улучшением,

по сравнению с ESTDM, в отношении эффективности поиска и хранения данных в очень больших базах данных. Улучшения основаны на хранении базовых карт, карт изменений и дополнительных карт во время каждого события; HST-DMS также решала проблему идентификации объекта. В то же время модель требует значительной обработки данных, когда к существующей информации добавляется новый временной шаг. Обе модели (как ESTDM, так и HST-DMS) доказали свою эффективность в сценариях с растровыми данными, однако они ограничены временем действия.

Событийно-ориентированный подход, основанный на расширенном управлении версиями (event-oriented approach based on extended-versioning) моделировал процессы с помощью событий и связанных с ними изменений в виде стандартизованных пространственно-временных операторов [11]. Процессы делились на три категории: основные (появление, исчезновение), трансформация (расширение, сжатие, деформация); движение (перемещение, вращение). Эта модель была «первой успешной попыткой записать индивидуальные характеристики динамических объектов» [7]. «Расширенное управление версиями» заключалось в применении для управления БГД специфических структур данных и процедур обработки.

Подход, основанный на временной логике (temporal logic-based approach) [12], базировался на явном представлении причинно-следственных связей между событиями и объектами. Модель давала инструменты для анализа исторических сценариев и понимания текущего состояния региона в зависимости от причинно-следственных зависимостей в прошлом.

Чисто событийно-ориентированная теория пространства и времени (pure event-oriented theory of space and time) [13] использовала сложные модели геообъектов, обладающих внутренней структурой и поведением, а также взаимосвязями с другими объектами. Модель описывала сложные динамические процессы с помощью парадигм параллелизма и взаимодействия.

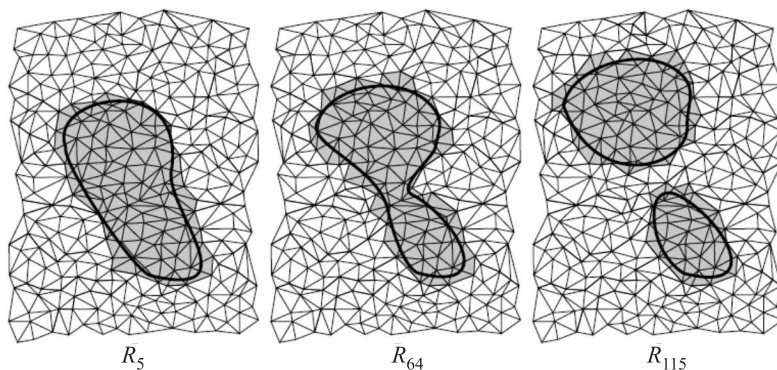


Рис. 2 Иллюстрация эволюции в сети триангуляции из [14]

Подход на основе непрерывных пространственных изменений в ориентированной на события дискретной пространственно-временной структуре (continuous spatial change in an event-oriented discrete spatiotemporal framework) [14] предназначался для обнаружения таких пространственных событий, как слияние, разделение, застой, движение, образование и исчезновение. Изменения географических явлений определялись анализом эволюции сети триангуляции (рис. 2). Предполагалось, что эта модель найдет особое применение в областях, где данные извлекаются из сенсорных сетей, например при мониторинге окружающей среды.

4 Моделирование на основе графов

Простота определения взаимосвязей между географическими объектами и отражения их эволюции в виде графов дает значительную гибкость представления изменений, движений и событий и позволяет комбинировать графы с другими подходами к геоделению.

Модель графа истории (History Graph Model) стала первым подходом к моделированию на основе графов [15]. Она была предназначена для идентификации различных видов временного поведения и визуализации эволюции географической информации. Этот подход отражал концепции как движения, так

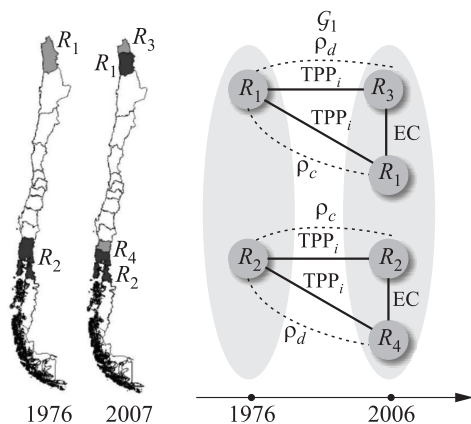


Рис. 3 Иллюстрация графа пространственно-временной эволюции регионов Чили из работы [16]

и изменения и поддерживал большинство пространственно-временных типов запросов. Но при этом он ограниченно описывал взаимодействие между объектами: например, процесс разделения обычно подразумевает не только модификацию исходного, но и создание дополнительного объекта, что в данной модели предусмотрено не было.

Модель на основе пространственно-временного графа для двумерных областей (spatiotemporal graph-based model for two-dimensional regions) [16] позволяла описывать эволюции двумерных областей (рис. 3). Она предполагала ряд операторов для манипулирования графом. Модель была реализована в пространственной базе данных (PostgreSQL + PostGIS) с использованием расширенной реляционной модели через схему, в которой были представлены сущности, отношения между сущностями и отношения филиации. Последние должны были быть четко указаны, поскольку ни одно из таких отношений не могло быть выведено из геометрии, а только из тематической корреляции.

5 Моделирование на основе признаков

Предложение по математическому моделированию на основе признаков (mathematical feature-based modeling) [17] предлагало набор примитивов для представления динамических объектов с учетом топологических и временных связей (последние отражали отношения возникновения явлений, например: «до», «после» и «во время»). Преимуществом предложения заявлена возможность определять взаимосвязи между динамическими явлениями с помощью пространственно-временных функций. Хотя это дает широкий спектр аналитических возможностей и делает модель легко вычислимой, но вместе с тем является недостатком, главным образом потому, что реальные географические явления обычно не описываются правильными геометрическими конструкциями. Но для географических явлений, которые могут быть представлены отдельными географическими точками, данный подход вполне применим.

6 Моделирование на основе процессов

В [18] была представлена модель, основанная на концепции пространственно-временных процессов (Spatio-Temporal Processes, STP) и названная STPs-based approach (подход на основе STPs). В этом подходе события определялись с помощью процессов, которые преобразуют сущности; была предложена типология трех основных классов STP: эволюция одного объекта, функциональные отношения между объектами и эволюция пространственных структур с участием нескольких объектов.

Пространственно-временная модель, основанная на процессах (process-based spatiotemporal model) [19] предназначалась для представления постепенных изменений. Здесь были определены два концепта пространственно-временного объекта: сущность (STEntity) и изменение (STChange). Главное преимущество модели заключалось в возможности идентифицировать изменения на разных уровнях детализации, причем как в пространственном, так и в тематическом аспекте.

Особенность модели геоатома (Geoatom Model) [20] заключалась в рассмотрении как простого объекта (внутренне однородного) — «геоатома», так и объекта с неоднородной внутренней структурой, которая к тому же может развиваться. Также был введен элемент «геодиполь», связанный с двумя местоположениями в пространстве-времени и предназначенный для описания взаимодействия между объектами (как процесса, который изменяет геополя и геообъекты). Но этот подход оказался неприменим к сложным сценариям.

7 Объектно-ориентированное моделирование

Объектно-ориентированный подход получил широкое признание в пространственно-временном моделировании данных, что обусловлено интуитивно близким представлением в виде независимых динамических геообъектов [21], а также гиб-

костью при отслеживании поведения, изменений и эволюции объектов. Объектно-ориентированный подход стал наиболее плодотворным: порядка трети предложенных пространственно-временных моделей объектно-ориентированные.

Применение методологии объектной организации для пространственно-временного моделирования впервые предложено в 1990 г. (см., например, [22]). Такой подход рассматривал географические ST-объекты как независимые единицы. При этом использовались объекты двух типов. ST-симплекс — первичный пространственный объект, представляющий собой точку, отрезок прямой или треугольную область (фрагмент плоскости)¹, связанный с парой моментов времени. Симплекс рассматривался как наименьшая единица, которой присваиваются временные атрибуты. ST-комплекс определялся как набор ST-симплексов и представлял собой географическое явление. Отметки времени отражали эволюции объекта с течением времени.

Интегрированная темпоральная ГИС-система [23] позволяла отслеживать эволюцию географических объектов на основе временных (темпоральных) взаимосвязей между версиями каждого объекта. Кроме того, возможно было определять события и процессы, влияющие на каждый объект, и анализировать их.

8 Моделирование движущихся объектов

Подход к пространственно-временному моделированию и анализу на основе движущихся объектов (Moving Objects) стал весьма плодотворным с учетом экспоненциальной эволюции мобильных и сенсорных технологий. Распространение этих технологий позволило получить представительные наборы пространственно-временных данных, фиксирующих изменения местоположения объектов с течением времени и тем самым описывающих их поведение. Благодаря этому движущиеся объекты стали горячей темой исследований для анализа траектории движущегося объекта или идентификации модели движения [24, 25].

В большинстве предложений рассматриваются объекты в виде движущихся точек. Лишь несколько исследований рассматривали двумерные пространственные объекты (как движущиеся области) или изменения в двумерных пространствах с течением времени (см., например, [19, 26]).

Появился даже специализированный вид базы данных движущихся объектов (см., например, [27]), предоставляющей эффективные методы хранения, индексации и запроса данных о перемещении. Это связано с необходимостью обработки данных о перемещении в реальном времени и разработкой сервисов, основанных на местоположении.

9 Заключение

Рассмотренные модели относятся к следующим основным подходам к пространственно-временному геодекодированию: на основе событий, на основе

¹ Два последних представляют собой элементы полилинии и полиповерхности соответственно.

графов, на основе признаков, на основе процессов, объектно-ориентированный и моделирование движущихся объектов. Это деление достаточно условно, поскольку многие особенности конкретных моделей, отнесенные в обзоре [1] к разным подходам, переключаются между собой.

На основе рассмотренных подходов к пространственно-временному моделированию можно отметить ряд основных тенденций развития T-GIS.

Большинство описанных подходов ориентировано на объектную модель местности, лишь некоторые из них были разработаны для растровых и гибридных моделей, что свидетельствует о доминировании представления динамических географических явлений как объектов. При этом в значительном числе моделей присутствуют элементы топологии (как пространственной, так и временной). Выбор модели обуславливает целевое предназначение T-GIS: растровые модели разрабатывались в контексте экологических приложений, а объектно-ориентированные — в контексте кадастровых и земельных информационных систем.

В исследовательском сообществе сложилось мнение относительно бесперспективности создания общей пространственно-временной модели, применимой к любому сценарию ее использования (см., например, [28]), поскольку разные дисциплины по-разному понимают изменения и, следовательно, предъявляют разные требования к их моделированию.

Вызывает сомнение деление эволюционных преобразований на категории «трансформации» (расширение, сжатие, деформация) и «движения» (перемещение, вращение). Возможно, имеет смысл рассматривать все эти преобразования в контексте единой категории — «эволюции геообъекта».

К характерным приемам в темпоральном моделировании можно отнести:

- атрибутирование не только геообъектов, но и изменений, типа «событие», «процесс»;
- регистрация причинно-следственных связей — использование темпоральной топологии наряду с обычной, пространственной и логической (т. е. семантической) топологией;
- применение пространственно-временных графов (типа сетей Петри);
- использование сложных темпоральных моделей геообъектов, обладающих внутренней структурой и даже предполагающих собственное, индивидуальное поведение («агенты»);
- применение парадигмы движущегося объекта.

Подход к моделированию на основе графов потенциально позволит выйти на новый уровень моделирования событий и процессов, а использование агентов — информационных объектов, обладающих способностью контролировать себя и взаимодействовать в общей среде, — дает разработчикам моделей возможность учитывать новые модели поведения [29]. Интеграция графов и агентов представляется многообещающим направлением исследований: графы будут

сосредоточены на событиях и процессах, а агенты — представлять интеллектуальные объекты и общую среду, в которой взаимодействуют все компоненты.

Движущиеся объекты наряду с агентно-ориентированным подходом также предлагают новые тенденции в моделировании. Рассмотрение сети движущихся объектов, в которой каждый объект представлен агентом, может заложить основы для интеллектуальных ГИС: автономные объекты могут взаимодействовать по общему сценарию, который определяется и ограничен правилами.

Важным моментом для Т-GIS представляется реализация возможности изменения временного масштаба (детализации изменений ситуации в темпоральном аспекте) наряду с использованием традиционной многоуровневой детализации ситуации (в семантическом и пространственном аспектах).

Литература

1. *Siabato W., Claramunt C., Harri S., Manso-Callejo M.A.* A survey of modelling trends in temporal GIS // *ACM Comput. Surv.*, 2018. Vol. 51. Iss. 2. Art. No. 30. 41 p. doi: 10.1145/3141772.
2. *Langran G., Chrisman N.R.* A framework for temporal geographic information // *Cartographica*, 1988. Vol. 25. Iss. 3. P. 1–14.
3. *Yuan M.* Use of a three-domain representation to enhance GIS support for complex spatiotemporal queries // *T. GIS*, 1999. Vol. 3. Iss. 2. P. 137–159.
4. *Yuan M., Hornsby K.S.* Computation and visualization for understanding dynamics in geographic domains: A research agenda. — Boca Raton, FL, USA: CRC Press, 2008. 120 p.
5. *Galton A.P., Worboys M.F.* Processes and events in dynamic geo-networks // *GeoSpatial semantics* / Eds. A. Rodríguez, I. F. Cruz, M. J. Egenhofer, and S. Levashkin. — Berlin: Springer Verlag, 2005. P. 45–59.
6. *Peuquet D.J., Wentz E.A.* An approach for time-based analysis of spatiotemporal data // *6th Symposium (International) on Spatial Data Handling Proceedings* / Eds. T. C. Waugh, R. G. Healey. — London: Taylor & Francis, 1994. P. 489–504.
7. *Pelekis N., Theodoulidis B., Kopanakis I., Theodoridis Y.* Literature review of spatiotemporal database models // *Knowl. Eng. Rev.*, 2004. Vol. 19. Iss. 3. P. 235–274.
8. *Peuquet D.J., Duan N.* An event-based spatiotemporal datamodel (ESTDM) for temporal analysis of geographical data // *Int. J. Geogr. Inf. Syst.*, 1995. Vol. 9. Iss. 1. P. 7–24.
9. *Yuan M.* Representing complex geographic phenomena in GIS // *Cartogr. Geogr. Inf. Sc.*, 2001. Vol. 28. Iss. 2. P. 83–96.
10. *Sengupta R., Yan C.* A hybrid spatio-temporal data model and structure (HST-DMS) for efficient storage and retrieval of land use information // *T. GIS*, 2004. Vol. 8. Iss. 3. P. 351–366.
11. *Claramunt C., Thériault M.* Managing time in GIS: An event-oriented approach // *Recent advances in temporal databases* / Eds. J. Clifford, A. Tuzhilin. — London: Springer Verlag, 1995. P. 23–42.
12. *Allen E., Edwards G., Bédard Y.* Qualitative causal modeling in temporal GIS // *Spatial information theory. A theoretical basis for GIS* / Eds. A. U. Frank, W. Kuhn. — Berlin: Springer Verlag, 1995. P. 397–412.

13. *Worboys M. F.* Event-oriented approaches to geographic phenomena // *Int. J. Geogr. Inf. Sci.*, 2005. Vol. 19. Iss. 1. P. 1–28.
14. *Worboys M. F., Duckham M.* Monitoring qualitative spatiotemporal change for geosensor networks // *Int. J. Geogr. Inf. Sci.*, 2006. Vol. 20. Iss. 10. P. 1087–1108.
15. *Renolen A.* History graphs: Conceptual modeling of spatio-temporal data // *GIS Frontiers in Business and Science Conference Proceedings* / Ed. M. Konecny. — Bern: International Cartographic Association, 1996. Vol. 2. 46 p.
16. *Del Mondo G., Rodríguez M. A., Claramunt C., Bravo L., Thibaud R.* Modeling consistency of spatio-temporal graphs // *Data Knowl. Eng.*, 2013. Vol. 84. P. 59–80.
17. *Ibañez A. M., Hoehne A. V.* Diseño de primitivas geométricas espacio-temporales para describir fenómenos dinámicos // *GeoFocus*, 2010. Vol. 10. Iss. 1. P. 232–251.
18. *Claramunt C., Thériault M., Parent C.* A qualitative representation of evolving spatial entities in two-dimensional topological spaces // *Innovations in GIS 5: 5th National Conference on GIS Research UK Selected Papers* / Ed. S. Carver. — London: Taylor & Francis, 1998. P. 121–131.
19. *Yang Y., Claramunt C.* A process-oriented multi-representation of gradual changes // *J. Geographic Information Decision Analysis*, 2003. Vol. 7. Iss. 1. P. 1–13.
20. *Goodchild M. F., Yuan M., Cova T. J.* Towards a general theory of geographic representation in GIS // *Int. J. Geogr. Inf. Sci.*, 2007. Vol. 21. Iss. 3. P. 239–260.
21. *Bian L.* Object-oriented representation for modelling mobile objects in an aquatic environment // *Int. J. Geogr. Inf. Sci.*, 2000. Vol. 14. Iss. 7. P. 603–623.
22. *Worboys M. F., Hearnshaw H. M., Maguire D. J.* Object-oriented data modelling for spatial databases // *Int. J. Geogr. Inf. Syst.*, 1990. Vol. 4. Iss. 4. P. 369–383.
23. *Khaddaj S., Adamu A., Morad M.* Construction of an integrated object oriented system for temporal GIS // *American J. Applied Sciences*, 2005. Vol. 2. Iss. 12. P. 1584–1594.
24. *Wolfson O., Mena E.* Applications of moving objects databases // *Spatial databases: Technologies, techniques and trends* / Eds. Y. Manolopoulos, A. Papadopoulos, M. G. Vassilakopoulos. — Hershey, PA, USA: IGI Global, 2005. P. 186–203.
25. *Demšar U., Buchin K., Cagnacci F., Safi K., Speckmann B., Van de Weghe N., Weiskopf D., Weibel R.* Analysis and visualisation of movement: An interdisciplinary review // *Movement Ecology*, 2015. Vol. 3. Iss. 1. Art. No. 5. 54 p.
26. *Peña J. L. H., Santos M. Y.* Representing, storing and mining moving objects data // *World Congress on Engineering Proceedings* / Eds. S. I. Ao, L. Gelman, D. W. Hukings, A. Hunter, A. M. Korsunsky. — Kwun Tong: Newswood Ltd., 2011. P. 1823–1828.
27. *Hajari H., Hakimpour F.* A spatial data model for moving object databases // *Int. J. Database Management Systems*, 2014. Vol. 6. Iss. 1. P. 1–20.
28. *Goodchild M. F.* Prospects for a space-time GIS // *Ann. Assoc. Am. Geogr.*, 2013. Vol. 103. Iss. 5. P. 1072–1077.
29. *Yu C., Peuquet D. J.* A geagent-based framework for knowledge-oriented representation: Embracing social rules in GIS // *Int. J. Geogr. Inf. Sci.*, 2009. Vol. 23. Iss. 7. P. 923–960.

Поступила в редакцию 15.08.22

OVERVIEW OF APPROACHES TO SPACE–TIME MODELING AND THE MAIN TRENDS IN THE DEVELOPMENT OF T-GIS

D. A. Nikishin

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

Abstract: The purpose of the work is to consider the characteristic features of a number of basic approaches and their corresponding models of spatiotemporal (temporal) geodata models proposed in recent decades and to identify the main trends in the development of spatiotemporal modeling for temporal geoinformation systems (T-GIS). It is noted that it is not necessary to talk about a general approach to spatiotemporal geomodeling due to the variety of requirements and features inherent in specific tasks of specific T-GIS. Graph-based modeling of events and processes, an agent-oriented approach, the use of the paradigm of moving objects, as well as the integration of these approaches together are important areas.

Keywords: modeling of geographical phenomena; evolution of geo objects; temporal geoinformation system; spatiotemporal geodata database; classification of temporal geomodels

DOI: 10.14357/08696527220305

References

1. Siabato, W., C. Claramunt, S. Ilarri, and M. A. Manso-Callejo. 2018. A survey of modelling trends in temporal GIS. *ACM Comput. Surv.* 51(2):30. 41 p. doi: 10.1145/3141772.
2. Langran, G., and N. R. Chrisman. 1988. A framework for temporal geographic information. *Cartographica* 25(3):1–14.
3. Yuan, M. 1999. Use of a three-domain representation to enhance GIS support for complex spatiotemporal queries. *T. GIS* 3(2):137–159.
4. Yuan, M., and K. S. Hornsby. 2008. *Computation and visualization for understanding dynamics in geographic domains: A research agenda*. Boca Raton, FL: CRC Press. 120 p.
5. Galton, A. P., and M. F. Worboys. 2005. Processes and events in dynamic geonetworks. *GeoSpatial semantics*. Eds. A. Rodríguez, I. F. Cruz, M. J. Egenhofer, and S. Levashkin. Berlin: Springer Verlag. 45–59.
6. Peuquet, D. J., and E. A. Wentz. 1994. An approach for time-based analysis of spatiotemporal data. *6th Symposium (International) on Spatial Data Handling Proceedings*. Eds. T. C. Waugh and R. G. Healey. London: Taylor & Francis. 489–504.
7. Pelekis, N., B. Theodoulidis, I. Kopanakis, and Y. Theodoridis. 2004. Literature review of spatio-temporal database models. *Knowl. Eng. Rev.* 19(3):235–274.

8. Peuquet, D. J., and N. Duan. 1995. An event-based spatiotemporal datamodel (ESTDM) for temporal analysis of geographical data. *Int. J. Geogr. Inf. Syst.* 9(1):7–24.
9. Yuan, M. 2001. Representing complex geographic phenomena in GIS. *Cartogr. Geogr. Inf. Sc.* 28(2):83–96.
10. Sengupta, R., and C. Yan. 2004. A hybrid spatio-temporal data model and structure (HST-DMS) for efficient storage and retrieval of land use information. *T. GIS* 8(3):351–366.
11. Claramunt, C., and M. Thériault. 1995. Managing time in GIS: An event-oriented approach. *Recent advances in temporal databases*. Eds. J. Clifford and A. Tuzhilin. London: Springer Verlag. 23–42.
12. Allen, E., G. Edwards, and Y. Bédard. 1995. Qualitative causal modeling in temporal GIS. *Spatial information theory. A theoretical basis for GIS*. Eds. A. U. Frank and W. Kuhn. Berlin: Springer Verlag. 397–412.
13. Worboys, M. F. 2005. Event-oriented approaches to geographic phenomena. *Int. J. Geogr. Inf. Sci.* 19(1):1–28.
14. Worboys, M. F., and M. Duckham. 2006. Monitoring qualitative spatiotemporal change for geosensor networks. *Int. J. Geogr. Inf. Sci.* 20(10):1087–1108.
15. Renolen, A. 1996. History graphs: Conceptual modeling of spatio-temporal data. *GIS Frontiers in Business and Science Conference Proceedings*. Ed. M. Konecny. Bern: International Cartographic Association. Vol. 2. 46 p.
16. Del Mondo, G., M. A. Rodríguez, C. Claramunt, L. Bravo, and R. Thibaud. 2013. Modeling consistency of spatio-temporal graphs. *Data Knowl. Eng.* 84:59–80.
17. Ibañez, A. M., and A. V. Hoehne. 2010. Diseño de primitivas geométricas espacio-temporales para describir fenómenos dinámicos. *GeoFocus* 10(1):232–251.
18. Claramunt, C., M. Thériault, and C. Parent. 1998. A qualitative representation of evolving spatial entities in two-dimensional topological spaces. *Innovations in GIS 5: 5th National Conference on GIS Research UK Selected Papers*. Ed. S. Carver. London: Taylor & Francis. 121–131.
19. Yang, Y., and C. Claramunt. 2003. A process-oriented multi-representation of gradual changes. *J. Geographic Information Decision Analysis* 7(1):1–13.
20. Goodchild, M. F., M. Yuan, and T. J. Cova. 2007. Towards a general theory of geographic representation in GIS. *Int. J. Geogr. Inf. Sci.* 21(3):239–260.
21. Bian, L. 2000. Object-oriented representation for modelling mobile objects in an aquatic environment. *Int. J. Geogr. Inf. Sci.* 14(7):603–623.
22. Worboys, M. F., H. M. Hearnshaw, and D. J. Maguire. 1990. Object-oriented data modelling for spatial databases. *Int. J. Geogr. Inf. Syst.* 4(4):369–383.
23. Khaddaj, S., A. Adamu, and M. Morad. 2005. Construction of an integrated object oriented system for temporal GIS. *American J. Applied Sciences* 2(12):1584–1594.
24. Wolfson, O., and E. Mena. 2005. Applications of moving objects databases. *Spatial databases: Technologies, techniques and trends*. Eds. Y. Manolopoulos, A. Papadopoulos, and M. G. Vassilakopoulos. Hershey, PA: IGI Global. 186–203.
25. Demšar, U., K. Buchin, F. Cagnacci, K. Safi, B. Speckmann, N. Van de Weghe, D. Weiskopf, and R. Weibel. 2015. Analysis and visualisation of movement: An interdisciplinary review. *Movement Ecology* 3(1):5. 54 p.

26. Peña, J. L. H., and M. Y. Santos. 2011. Representing, storing and mining moving objects data. *World Congress on Engineering Proceedings*. Eds. S. I. Ao, L. Gelman, D. W. Hukings, A. Hunter, and A. M. Korsunsky. Kwun Tong: Newswood Ltd. 1823–1828.
27. Hajari, H., and F. Hakimpour. 2014. A spatial data model for moving object databases. *Int. J. Database Management Systems* 6(1):1–20.
28. Goodchild, M. F. 2013. Prospects for a space–time GIS. *Ann. Assoc. Am. Geogr.* 103(5):1072–1077.
29. Yu, C., and D. J. Peuquet. 2009. A geoagent-based framework for knowledge-oriented representation: Embracing social rules in GIS. *Int. J. Geogr. Inf. Sci.* 23(7):923–960.

Received August 15, 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

ОПТИМАЛЬНЫЕ АССОЦИАЦИИ УСТРОЙСТВ В СЕТЯХ МЕЖМАШИННОГО ВЗАИМОДЕЙСТВИЯ СТАНДАРТОВ NB-IoT И LTE-M*

Н. В. Степанов¹, А. М. Турликов², В. О. Бегисhev³

Аннотация: Новые услуги массового межмашинного взаимодействия (mMTC, massive machine-type communications), такие как мониторинг энергетических сетей, характеризуются пачечной структурой трафика, что оказывает негативное влияние на задержку доставки информации. Кроме того, они также требуют высокой надежности доставки информации и доступности сети, которые могут быть достигнуты за счет использования нескольких сетей радиодоступа (Radio Access Technology, RAT) на одном оконечном устройстве (UE, user equipment). В данной работе построена математическая модель для оценки задержки на радиоинтерфейсе в сетях с низким потреблением энергии (Low Power Wide Area Networks, LPWAN) с несколькими RAT при пачечном поступлении сообщений.

Ключевые слова: 5G; NB-IoT; LTE-M; межмашинное взаимодействие

DOI: 10.14357/08696527220306

1 Введение

Беспроводные сети с низким потреблением энергии (LPWAN), такие как LTE-M (long term evolution for machine) и NB-IoT (narrowband Internet of Things), были разработаны для передачи трафика, генерируемого большим числом оконечных устройств (UE), работающих асинхронно и передающих небольшие объемы данных. Это свойство позволяет мультиплексировать сообщения от множества UE, а в случае NB-IoT даже достигать требований IMT-2020 для массового межмашинного взаимодействия (mMTC) [1, 2].

Услуга мониторинга энергетических сетей представляет собой одну из сложных услуг для сетей LPWAN. В дополнение к повышению требований по задержке доставки информации такая услуга требует постоянной связи с UE, работающими в состоянии поддержки активного соединения (RRC-Connected) для опроса UE через регулярные интервалы времени. Такое поведение приводит к пачечному поступлению сообщений на радиоинтерфейс LPWAN, что увели-

*Работа выполнена при финансовой поддержке Российского научного фонда (проект 22-19-00305 «Пространственно-временные стохастические модели беспроводных сетей с большим числом абонентов»).

¹Санкт-Петербургский государственный университет аэрокосмического приборостроения (ГУАП), stepanov.nikita@guar.ru

²Санкт-Петербургский государственный университет аэрокосмического приборостроения (ГУАП), turlikov@k36.org

³Российский университет дружбы народов (РУДН), begishev-vo@rudn.ru

чивает задержку и может приводить к нарушению требований IMT-2020 для услуг mMTC. Для повышения надежности, а также унификации конструкции UE были предложены UE с несколькими интерфейсами LPWAN. Последнее приводит к проблеме оптимального выбора сети LPWAN таким образом, чтобы задержка доставки информации была минимальной [3, 4].

В данной работе предлагается математическая модель для оценки показателей задержки услуг mMTC, характеризующихся пачечными поступлениями сообщений для сетей LPWAN, LTE-M и NB-IoT. На основе разработанной модели построен алгоритм численной оптимизации для определения количества UE, подключенных в каждой из технологий.

2 Модель системы

Известно, что при пачечном поступлении сообщений процедура случайного доступа (RAP, random access procedure) становится основным фактором, влияющим на качество обслуживания абонентов [5, 6].

Имеется множество абонентов U и две базовые станции (БС): БС₁ (NB-IoT) и БС₂ (LTE-M). Каждый абонент из множества U может иметь доступ к любой из этих БС, но только к одной. Требуется разделить множество абонентов U на два подмножества U_1 и U_2 таким образом, чтобы средняя задержка доступа была минимальной. Для описания предлагаемой системы введем ряд допущений.

Допущение 1. Станции БС₁ и БС₂ обслуживают l_1 и l_2 параллельных каналов случайного множественного доступа соответственно. Время передачи по каналу разделено на окна. Все окна имеют одинаковую длительность, равную $t_{\text{окна}1}$ для БС₁ и $t_{\text{окна}2}$ для БС₂. Абонент может начинать передачу сообщения только в начале очередного окна.

Допущение 2. В каждом окне может произойти одно из трех событий:

- (1) в окне передает один абонент: событие S — успех;
- (2) в окне не передает ни один абонент: событие E — пусто;
- (3) в окне передают два или более абонентов: событие C — конфликт.

Допущение 3. Базовая станция, наблюдая выход канала, к концу окна узнает о событии в окне и передает информацию о событии всем абонентам. Абоненты, у которых есть сообщение, используют эту информацию и в следующем окне принимают решение передавать или нет сообщение в соответствии с вероятностью $P_t = \min(1, l/N_t)$, где N_t — число активных абонентов к началу окна t [7]. На практике это не может быть реализовано, так как число активных абонентов неизвестно. Для любого алгоритма, который может быть использован, средняя задержка будет выше [8, 9].

Допущение 4. Предполагается, что зоны покрытия БС таковы, что потеря данных, вызванная низким уровнем принимаемого сигнала, пренебрежимо мала.

Допущение 5. Для описания процесса возникновения сообщений будем использовать модель с периодическим опросом абонентов. В соответствии с ней новое

Характеристики NB-IoT и LTE-M, используемые в работе

Технология	Размер пакета, байт	Число преамбул	Синхронизация, мс	Время случайного доступа (RAP), мс	Время передачи данных, мс
LTE Cat NB 1	200	$l_1 = 12$	$t_{\text{синх.1}} = 300$	$t_{\text{окна1}} = 40$	$t_{\text{дан.1}} = 350$
LTE Cat M1	200	$l_2 = 36$	$t_{\text{синх.2}} = 300$	$t_{\text{окна2}} = 20$	$t_{\text{дан.2}} = 80$

сообщение для передачи появляется у всех M абонентов одновременно с периодом, равным Δ единиц времени. После появления сообщения абонент пытается осуществить его передачу на БС, с которой он ассоциирован. Если абонент не передаст сообщение до появления нового сообщения, то непереданное сообщение теряется и абонент начинает работать с новым сообщением.

Допущение 6. Для модели с периодическим опросом абонентов предполагается, что интервал между опросами устройств в соте достаточен, чтобы сообщения устройств могли быть переданы и не оставалось ни одного активного устройства в конце интервала.

Таблица иллюстрирует параметры технологий NB-IoT и LTE-M, используемые в работе [6]. Отметим, что на основе данных, представленных в таблице, в дальнейшем осуществляется пересчет временных параметров из окон в единицы времени.

3 Методика расчета

Среднюю задержку для модели с периодическим опросом абонентов можно рассчитать тремя способами с разной степенью точности и сложности вычислений. Первый способ расчета — это точный метод с помощью марковской модели, предложенный в [10]. Второй способ расчета опирается на рекуррентное уравнение постепенного перехода абонентов из активного состояния в состояние без готовых к передаче сообщений [6]. Отметим, что, используя как первый, так и второй метод, нельзя получить явное выражение средней задержки как функции параметров модели. Третий способ, позволяющий получить решение в явном виде, заключается в использовании следующей, более грубой, чем дает рекуррентное уравнение, оценки для числа активных абонентов:

$$\bar{N}_{t+1} = (\bar{N}_t - le^{-1}) I(\bar{N}_t \geq le^{-1}).$$

Для данной оценки будем считать, что все абоненты выходят за время le^{-1} , т. е.

$$\bar{b}(M, \delta) = \begin{cases} \frac{M}{2le^{-1}}, & \text{если } \frac{M}{le^{-1}} \leq \delta; \\ \frac{M}{le^{-1}} - \frac{\delta}{2} & \text{иначе.} \end{cases} \quad (1)$$

Для выбора оптимального распределения числа абонентов M_1 и M_2 для первой и второй технологии соответственно требуется решить следующую оптимизационную задачу:

$$\begin{cases} \max (\tau_1 (M_1 \Delta), \tau_2 (M_2 \Delta)) \rightarrow \min ; \\ M_1 + M_2 = M . \end{cases}$$

Легко показать, что для оптимальных значений выполняется равенство $\tau_1(M_1) = \tau_2(M_2)$. Таким образом, для решения этой оптимизационной задачи требуется проделать следующие действия.

Шаг 1. Численным путем получить зависимость $b_1(M_1, \delta)$.

Шаг 2. Численным путем получить зависимость $b_2(M_2, \delta)$.

Шаг 3. Численным путем решить следующее уравнение:

$$b_1 \left(M_1, \frac{\Delta}{t_{\text{окна } 1}} \right) t_{\text{окна } 1} + t_{\text{дан. } 1} = b_2 \left(M_2, \frac{\Delta}{t_{\text{окна } 2}} \right) t_{\text{окна } 2} + t_{\text{дан. } 2} .$$

Задача решается при следующих ограничениях:

$$\begin{aligned} M_1 + M_2 &= M ; \\ \frac{M_1}{l_1 e^{-1}} &< \frac{\Delta}{t_{\text{окна } 1}} ; & \frac{M_2}{l_2 e^{-1}} &< \frac{\Delta}{t_{\text{окна } 2}} . \end{aligned} \quad (2)$$

Следует отметить, что при ограничениях (2) к концу периода Δ все абоненты успешно передадут свои сообщения.

Для того чтобы упростить решение, можно вместо точных значений $b_1(M_1, \delta)$ и $b_2(M_2, \delta)$ воспользоваться оценками согласно формуле (1) и получить решения в явном виде. Рассмотрим случай $M_1/(l_1 e^{-1}) < \delta$ и $(M - M_1)/(l_2 e^{-1}) < \delta$. Расчет задержки для первой и второй технологии дает

$$\tau_1 = \frac{M_1}{l_1 e^{-1} \cdot 2} t_{\text{окна } 1} + t_{\text{дан. } 1} ; \quad \tau_2 = \frac{M - M_1}{l_2 e^{-1} \cdot 2} t_{\text{окна } 2} + t_{\text{дан. } 2} .$$

Для нахождения оптимального значения M_1 требуется решить уравнение

$$\frac{M_1}{l_1 e^{-1} \cdot 2} t_{\text{окна } 1} + t_{\text{дан. } 1} = \frac{M - M_1}{l_2 e^{-1} \cdot 2} t_{\text{окна } 2} + t_{\text{дан. } 2}$$

при условии выполнении неравенств (2). Получаем

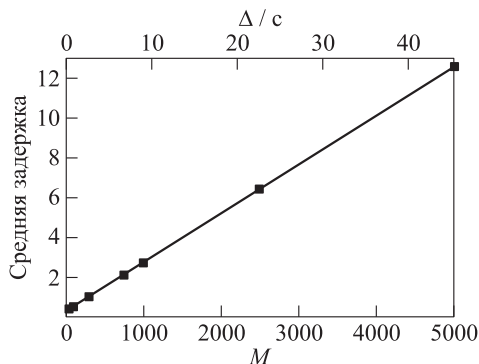
$$M_1 = \frac{M l_1 t_{\text{окна } 2} - 2 l_1 l_2 e^{-1} (t_{\text{дан. } 1} - t_{\text{дан. } 2})}{l_1 t_{\text{окна } 2} + l_2 t_{\text{окна } 1}} .$$

В точке оптимума значения задержки для обеих технологий совпадают и равны

$$\tau_{\text{опт}} = \frac{Ml_1t_{\text{окна}2} - 2l_1l_2e^{-1}(t_{\text{дан.1}} - t_{\text{дан.2}})}{2l_1e^{-1}(l_1t_{\text{окна}2} + l_2t_{\text{окна}1})} t_{\text{окна}1} + t_{\text{дан.1}}.$$

4 Численные результаты

На рисунке показаны значения средней задержки как функции от общего числа устройств в системе M и периода опроса абонентов Δ для оптимального распределения абонентов между технологиями LTE-M и NB-IoT. Отметим, что при рассматриваемой фиксированной нагрузке M/Δ средняя задержка сообщения увеличивается с ростом M и Δ . Такое поведение обусловлено пачечной структурой трафика. Для численных расчетов выбраны 12 преамбул для NB-IoT и 5 преамбул для LTE-M.



Зависимость средней задержки в системе при оптимальном распределении абонентов между технологиями от общего числа абонентов M при фиксированной нагрузке M/Δ

5 Заключение

В работе построена математическая модель для оценки задержки на радиоинтерфейсе в сетях с низким потреблением энергии LPWAN с несколькими RAT при пачечном поступлении сообщений. Также предложен численный алгоритм для оценки оптимальных ассоциаций устройств, минимизирующий задержку передачи информации на радиоинтерфейсе.

Литература

1. Minimum requirements related to technical performance for IMT-2020 radio interface(s) // ITU, 2017. Report M.2410-0. <https://www.itu.int/pub/R-REP-M.2410-2017>.
2. Study on self evaluation towards IMT-2020 submission // 3GPP, 2019. Technical Report 37.910. <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3190>.
3. Petrov V., Samuylov A., Begishev V., Moltchanov D., Andreev S., Samouylov K., Koucheryavy Y. Vehicle-based relay assistance for opportunistic crowdsensing over

- narrowband IoT (NB-IoT) // IEEE Internet Things, 2017. Vol. 5. Iss. 5. P. 3710–3723. doi: 10.1109/JIOT.2017.2670363.
4. *Begishev V., Petrov V., Samuylov A., Moltchanov D., Andreev S., Koucheryavy Y., Samouylov K.* Resource allocation and sharing for heterogeneous data collection over conventional 3GPP LTE and emerging NB-IoT technologies // *Comput. Commun.*, 2018. Vol. 120. P. 93–101. doi: 10.1016/j.comcom.2018.01.009.
 5. *Stepanov N., Moltchanov D., Turlikov A.* Modeling the NB-IoT transmission process with intermittent network availability // *Internet of things, smart spaces, and next generation networks and system. — Lecture notes in computer science ser. — Springer*, 2020. Vol. 7428. P. 241–254. doi: 10.1007/978-3-030-65726-0_22.
 6. *Stusek M., Stepanov N., Moltchanov D., Masek P., Mozny R., Turlikov A., Hosek J.* Optimizing NB-IoT communication patterns for permanently connected mMTC devices // *IEEE WCNC. — Piscataway, NJ, USA: IEEE*, 2022. P. 1413–1418. doi: 10.1109/WCNC51071.2022.9771847.
 7. *Galinina O., Turlikov A., Andreev S., Koucheryavy Y.* Stabilizing multi-channel slotted aloha for machine-type communications // *IEEE Symposium (International) on Information Theory. — Piscataway, NJ, USA: IEEE*, 2013. P. 2119–2123. doi: 10.1109/ISIT.2013.6620600.
 8. *Koseoglu M.* Lower bounds on the LTE-A average random access delay under massive M2M arrivals // *IEEE T. Commun.*, 2016. Vol. 64. Iss. 5. P. 2104–2115. doi: 10.1109/TCOMM.2016.2550526.
 9. *Галинина О. С., Андреев С. Д., Турликов А. М.* Учет специфики доступа большого числа устройств при межмашинном взаимодействии в современных сотовых сетях // *Информационно-управляющие системы*, 2018. № 4(95). С. 105–114. doi: 10.31799/1684-8853-2018-4-105-114.
 10. *Степанов Н. В., Турликов А. М., Бегисhev В. О.* Метод анализа вероятностных характеристик обслуживания пачечного трафика в системах NB-IoT // *Электросвязь*, 2022 (в печати).

Поступила в редакцию 08.07.22

OPTIMAL DEVICE ASSOCIATIONS IN MACHINE-TO-MACHINE NETWORKS OF NB-IoT AND LTE-M STANDARDS

N. V. Stepanov¹, A. M. Turlikov¹, and V. O. Begishev²

¹Saint Petersburg State University of Aerospace Instrumentation, 67 Bolshaya Morskaya Str., Saint Petersburg 190000, Russian Federation

²Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation

Abstract: The new mass machine to machine communication services currently emerging, such as power grid monitoring, are characterized by a fundamentally different nature of the traffic incoming process. By polling the User Equipment

(UE) at regular intervals, these services result in bursting of messages which has a negative impact on the delay in information delivery. In addition, they also require high reliability of information delivery and network availability which can be achieved by using multiple radio access networks (Radio Access Technology, RAT) on one UE. In this paper, the authors develop a mathematical model for estimating radio interface latency in Low Power Wide Area Networks with multiple RATs for bursty messages. The authors also propose a numerical algorithm for estimating optimal UE associations minimizing the information transmission delay on the radio interface.

Keywords: 5G; NB-IoT; LTE-M; machine-to-machine communication

DOI: 10.14357/08696527220306

Acknowledgments

The work was supported by the Russian Science Foundation, project No. 22-19-00305 “Spatial-temporal stochastic models of wireless networks with a large number of users.”

References

1. ITU. 2017. Minimum requirements related to technical performance for IMT-2020 radio interface(s). Report M.2410-0. Available at: <https://www.itu.int/pub/R-REP-M.2410-2017> (accessed August 8, 2022).
2. 3GPP. 2019. Study on self evaluation towards IMT-2020 submission. Technical Report 37.910. Available at: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3190> (accessed August 8, 2022).
3. Petrov, V., A. Samuylov, V. Begishev, D. Moltchanov, S. Andreev, K. Samouylov, and Y. Koucheryavy. 2017. Vehicle-based relay assistance for opportunistic crowd-sensing over narrowband IoT (NB-IoT). *IEEE Internet Things* 5(5):3710–3723. doi: 10.1109/JIOT.2017.2670363.
4. Begishev, V., V. Petrov, A. Samuylov, D. Moltchanov, S. Andreev, Y. Koucheryavy, and K. Samouylov. 2018. Resource allocation and sharing for heterogeneous data collection over conventional 3GPP LTE and emerging NB-IoT technologies. *Comput. Commun.* 120:93–101. doi: 10.1016/j.comcom.2018.01.009.
5. Stepanov, N., D. Moltchanov, and A. Turlikov. 2020. Modeling the NB-IoT transmission process with intermittent network availability. *Internet of things, smart spaces, and next generation networks and system*. Lecture notes in computer science ser. Springer. 7428: 241–254. doi: 10.1007/978-3-030-65726-0_22.
6. Stusek, M., N. Stepanov, D. Moltchanov, P. Masek, R. Mozny, A. Turlikov, and J. Hosek. 2022. Optimizing NB-IoT communication patterns for permanently connected mMTC devices. *IEEE WCNC*. Piscataway, NJ: IEEE. 1413–1418. doi: 10.1109/WCNC51071.2022.9771847.
7. Galinina, O., A. Turlikov, S. Andreev, and Y. Koucheryavy. 2013. Stabilizing multi-channel slotted ALOHA for machine-type communications. *Symposium (International) on Information Theory Proceedings*. Piscataway, NJ: IEEE. 2119–2123. doi: 10.1109/ISIT.2013.6620600.

8. Koseoglu, M. 2016. Lower bounds on the LTE-A average random access delay under massive M2M arrivals. *IEEE T. Commun.* 64(5): 2104-2115. doi: 10.1109/TCOMM.2016.2550526.
9. Galinina, O., S. Andreev, and A. Turlikov. 2018. Uchet spetsifiki dostupa bol'shogo chisla ustroystv pri mezhmashinnom vzaimodeystvii v sovremennykh sotovykh setyakh [On multiple access of a large number of machine-type devices in cellular networks]. *Informatsionno-upravlyayushchiye sistemy* [Information and Control Systems] 4(95):105–114. doi: 10.31799/1684-8853-2018-4-105-114.
10. Stepanov, N., A. Tyurlikov, and V. Begishev. 2022 (in press). Metod analiza veroyatnostnykh kharakteristik obsluzhivaniya pachechnogo trafika v sistemakh NB-IoT [A method for analyzing the probabilistic characteristics of burst traffic service in NB-IoT systems]. *Electrosvyaz'* [Electrosvyaz Magazine].

Received July 8, 2022

Contributors

Stepanov Nikita V. (b. 1996) — assistant professor, Department of Infocommunication Systems, Saint Petersburg State University of Aerospace Instrumentation, 67 Bolshaya Morskaya Str., Saint Petersburg 190000, Russian Federation; stepanov.nikita@guap.ru

Turlikov Andrey M. (b. 1957) — Doctor of Science in Technology, professor, head of department, Saint Petersburg State University of Aerospace Instrumentation, 67 Bolshaya Morskaya Str., Saint Petersburg 190000, Russian Federation; turlikov@guap.ru

Begishev Vyacheslav O. (b. 1988) — Candidate of Science (PhD) in physics and mathematics, associate professor, Department of Applied Probability and Informatics, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; begishev-vo@rudn.ru

АНАЛИЗ РАСПРЕДЕЛЕНИЯ НАГРУЗОК И РЕСУРСОВ В МНОГОПОЛЬЗОВАТЕЛЬСКОЙ СЕТИ ПРИ РАВНЫХ МЕЖУЗЛОВЫХ ПОТОКАХ

Ю. Е. Малащенко¹, И. А. Назарова²

Аннотация: Анализируется метод распределения ресурсов для формирования сети связи при равных межузловых потоках. В рамках формальной математической модели пропускные способности ребер рассматриваются как компоненты вектора ресурсов различных типов, которые требуются для передачи потоков разных видов. Граф сети и сумма пропускных способностей ребер считаются заданными. Предлагается алгоритмическая процедура выбора пропускных способностей, при которых все межузловые потоки равны и достигается предельная загрузка всех ребер сети. При поиске соответствующих распределений используются значения максимальных однопродуктовых потоков, которые определяются для каждой пары узлов-корреспондентов. Вычислительные эксперименты проводились для сетей с различными структурными особенностями.

Ключевые слова: многопродуктовая потоковая модель; метод распределения ресурсов сети; уравнильное распределение межузловых потоков; предельная загрузка сети

DOI: 10.14357/08696527220307

1 Введение

Данная работа продолжает исследования методов анализа функциональных возможностей телекоммуникационных систем [1, 2]. В рамках многопродуктовой сетевой модели рассматривается вычислительная процедура оценки распределения нагрузок и ресурсов при равных совместно допустимых межузловых потоках. Под ресурсом, выделяемым некоторой паре узлов-корреспондентов, понимается суммарная величина пропускных способностей, необходимых для обеспечения связи. Сумма дуговых потоков разных видов трактуется как нагрузка на некотором ребре, возникающая при одновременном соединении корреспондентов. Суммарный ресурс сети считается заданным, и осуществляется поиск распределения, при котором межузловые потоки равны друг другу. В ходе эксперимента для каждой пары узлов вычисляется максимальный однопродуктовый поток. Маршруты передачи для каждой пары корреспондентов прокладываются по ребрам

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, malash09@ccas.ru

²Федеральный исследовательский центр «Информатика и управление» Российской академии наук, igina-nazar@yandex.ru

соответствующих минимальных разрезов и подсчитываются нагрузки при равных значениях потоков. На основе найденных значений нагрузок определяются требуемые пропускные способности. Анализ результатов экспериментов позволяет оценить распределение ресурсов при предельной загрузке сети и равных межузловых потоках в многопользовательских сетях [1, 2].

В настоящее время для создания, развития и эксплуатации телекоммуникационных систем используются потоковые модели и разрабатываются специальные методы решения [3, 4]. Математические модели передачи многопродуктового потока применяются для поиска недискриминирующих правил распределения ресурсов в системах связи [5]. На основе формализма многокритериальной оптимизации создаются методы анализа функциональных характеристик с учетом вектора требований всех равноправных и невзаимозаменяемых пользователей [6]. В рамках методологии исследования операций и теории игр рассматриваются *справедливые правила* распределения потоков и ресурсов: решается задача на максимум и/или находятся гарантированные оценки [7–10]. В русле указанных работ лежит алгоритмическая схема получения уравнивательных распределений межузловых потоков, приведенная в разд. 3. В разд. 4 обсуждаются результаты экспериментов и сравниваются достижимые межузловые потоки и предельные распределения пропускных способностей в сетях с различной структурой. Предложенный метод имеет полиномиальную оценку требуемого числа операций.

2 Математическая модель

Для описания многопользовательской сетевой системы связи воспользуемся следующей математической записью модели передачи многопродуктового потока. Сеть G задается множествами $\langle V, R, U, P \rangle$: узлов (вершин) сети $V = \{v_1, v_2, \dots, v_n, \dots, v_N\}$; неориентированных ребер $R = \{r_1, r_2, \dots, r_k, \dots, r_E\}$; ориентированных дуг $U = \{u_1, u_2, \dots, u_k, \dots, u_{2E}\}$; пар узлов-корреспондентов $P = \{p_1, p_2, \dots, p_M\}$. Предполагается, что в сети отсутствуют петли и двоянные ребра. В многопользовательской сети G рассматривается $M = N(N - 1)$ независимых, невзаимозаменяемых и равноправных межузловых потоков различных видов.

Ребро $r_k \in R$ соединяет *смежные* вершины v_{n_k} и v_{j_k} . Каждому ребру r_k ставятся в соответствие две ориентированные дуги u_k и u_{k+E} из множества U . Дуги $\{u_k, u_{k+E}\}$ определяют прямое и обратное направления передачи потока по ребру r_k между концевыми вершинами v_{n_k} и v_{j_k} .

Каждой паре узлов-корреспондентов p_m из множества P ставятся в соответствие: вершина-источник с номером s_m , из s_m входной поток m -го вида поступает в сеть; вершина-приемник с номером t_m , из t_m поток m -го вида покидает сеть. Обозначим через z_m величину *межузлового* потока m -го вида, поступающего в сеть через узел с номером s_m и покидающего сеть из узла с номером t_m ; x_{mk} и $x_{m(k+E)}$ — поток m -го вида, который передается по дугам u_k и u_{k+E} согласно направлению передачи, $x_{mk} \geq 0$, $x_{m(k+E)} \geq 0$, $m = \overline{1, M}$, $k = \overline{1, E}$; $S(v_n)$ —

множество номеров исходящих дуг, по ним поток покидает узел v_n ; $T(v_n)$ — множество номеров входящих дуг, по ним поток поступает в узел v_n .

Состав множеств $S(v_n)$ и $T(v_n)$ однозначно формируется в ходе выполнения следующей процедуры. Пусть некоторое ребро $r_k \in R$ соединяет вершины с номерами n и j , такими что $n < j$. Тогда ориентированная дуга $u_k = (v_n, v_j)$, направленная из вершины v_n в v_j , считается *исходящей* из вершины v_n и ее номер k заносится в множество $S(v_n)$, а дуга u_{k+E} , направленная из v_j в v_n , — *входящей* для v_n и ее номер $k + E$ помещается в список $T(v_n)$. Дуга u_k является *входящей* для v_j , и ее номер k попадает в $T(v_j)$, а дуга u_{k+E} — *исходящей*, и номер $k + E$ вносится в список исходящих дуг $S(v_j)$.

Во всех узлах сети $v_n \in V$, $n = \overline{1, N}$, для каждого вида потока должны выполняться условия сохранения потоков:

$$\sum_{i \in S(v_n)} x_{mi} - \sum_{i \in T(v_n)} x_{mi} = \begin{cases} z_m, & \text{если } v_n = v_{s_m}; \\ -z_m, & \text{если } v_n = v_{t_m}; \\ 0 & \text{в остальных случаях,} \end{cases} \\ n = \overline{1, N}, \quad m = \overline{1, M}, \quad x_{mi} \geq 0, \quad z_m \geq 0. \quad (1)$$

Величина z_m равна входному межузловому потоку m -го вида, проходящему от источника к приемнику пары p_m при распределении потоков x_{mi} по дугам сети.

Каждому ребру $r_k \in R$ приписывается неотрицательное число d_k , определяющее суммарный предельно допустимый поток, который можно передать по ребру r_k в обоих направлениях. В исходной сети компоненты вектора пропускных способностей $\mathbf{d} = (d_1, d_2, \dots, d_k, \dots, d_E)$ — наперед заданные положительные числа $d_k > 0$. Обозначим через D суммарную пропускную способность сети G :

$$D = \sum_{k=1}^E d_k.$$

Вектор \mathbf{d} задает следующие ограничения на сумму потоков всех видов, передаваемых по ребру r_k одновременно:

$$\sum_{m=1}^M (x_{mk} + x_{m(k+E)}) \leq d_k, \quad x_{mk} \geq 0, \quad x_{m(k+E)} \geq 0, \quad k = \overline{1, E}. \quad (2)$$

В рамках данной модели пропускная способность ребер сети измеряется в условных единицах потока и трактуется как *ресурсное* ограничение. Сумма дуговых потоков (2) рассматривается как *требование* на предоставление ресурса k -го типа, а $y(k)$ — суммарная *нагрузка* на ребро r_k :

$$y(k) = \sum_{m=1}^M (x_{mk} + x_{m(k+E)}), \quad k = \overline{1, E}. \quad (3)$$

Сумма реберных дуговых потоков m -го вида

$$y_m = \sum_{i=1}^{2E} x_{mi}, \quad m = \overline{1, M},$$

характеризует *нагрузку* на сеть при передаче межузлового потока величины z_m из узла-источника s_m в узел-приемник t_m . Величина y_m показывает, какая суммарная пропускная способность сети потребуется для передачи межузлового потока z_m . В рамках модели отношение реберных и межузловых потоков

$$w_m = \frac{y_m}{z_m}, \quad m = \overline{1, M},$$

можно трактовать как удельные *затраты* ресурсов сети при передаче единичного потока m -го вида между узлами s_m и t_m при соответствующих дуговых потоках x_{mi} .

Ограничения (1) и (2) задают множество допустимых значений компонент вектора межузловых потоков $\mathbf{z} = (z_1, z_2, \dots, z_m, \dots, z_M)$:

$$\mathcal{Z}(\mathbf{d}) = \{\mathbf{z} \geq 0 \mid \exists \mathbf{x} \geq 0 : (\mathbf{z}, \mathbf{x}) \text{ удовлетворяют (1), (2)}\}.$$

Допустимые распределения реберных потоков принадлежат подмножеству

$$Y(\mathbf{d}) = \{\mathbf{y} \geq 0 \mid (\mathbf{z}, \mathbf{x}, \mathbf{y}) \text{ (1)–(3)}\}.$$

В множестве пар P выделяется подмножество $P(\mathbf{M}_+)$ смежных пар узлов-корреспондентов, расположенных в концевых вершинах ребра r_k , $k = \overline{1, E}$. Список номеров \mathbf{M}_+ формируется по следующему правилу. Пусть ребро r_k соединяет вершины с номерами n и j , такими что $n < j$. Для пары узлов-корреспондентов с номером k , расположенных в узлах $\{v_{n_k}, v_{j_k}\}$, узел v_{n_k} считается источником s_k , а узел v_{j_k} — приемником t_k потока k -вида (поток z_k передается в сети G из узла с номером n в узел с номером j для пары p_k). Для пары с номером $k + E$ узел v_{j_k} считается источником s_{k+E} , а узел v_{n_k} — приемником t_{k+E} . Таким образом формируется $\mathbf{M}_+ = \{1, 2, \dots, E, E + 1, \dots, 2E\}$ — список номеров смежных пар, $|\mathbf{M}_+| = 2E$ — общее число пар. Обозначим через $P(\mathbf{M}_-)$ подмножество несмежных пар, для которых кратчайший маршрут соединения содержит более одного ребра. Тогда

$$P = P(\mathbf{M}_+) \cup P(\mathbf{M}_-), \quad P(\mathbf{M}_+) \cap P(\mathbf{M}_-) = \emptyset.$$

3 Метод оценки нагрузки при равных межузловых потоках и предельной загрузке ребер

В рамках модели проводились вычислительные эксперименты оценки требуемых пропускных способностей при предельной загрузке всех ребер сети и равных

межузловых потоках. При подготовке данных формировался вектор исходных пропускных способностей $\mathbf{d}(0)$ для некоторой заданной сети $G(0)$, в которой

$$D(0) = \sum_{k=1}^E d_k(0).$$

В сети $G(0)$ при заданных $d_k(0)$ последовательно решалась задача поиска максимального независимого однопродуктового потока [11] для каждой пары узлов $p_m \in P$, $m = \overline{1, M}$.

Задача 1. Для некоторой пары узлов p_a найти

$$z_a^0(0) = \max \{z_a | (\mathbf{z}, x) \in \mathcal{Z}(\mathbf{d}(0))\}$$

при дополнительных условиях

$$z_m = 0, \quad m \neq a, \quad m = \overline{1, M}.$$

При последовательном решении задачи 1 для каждой пары $p_a \in P$ вычисляются максимальный междуузловой поток $z_a^0(0)$ и соответствующие дуговые потоки $(x_{ak}^0(0), x_{a(k+E)}^0(0))$, $k = \overline{1, E}$. Для $z_a^0(0) \neq 0$ определяется нормирующий коэффициент $\xi_a^0(0) = 1/z_a^0(0)$ и дуговые потоки $x_{ak}^0 = \xi_a^0(0)x_{ak}^0(0)$, $k = \overline{1, E}$. При передаче дуговых потоков x_{ak}^0 соответствующий междуузловой поток z_a^0 равен единице.

Задача 1 решается последовательно для всех $p_m \in P$, и для найденных дуговых потоков $x_{mk}^0(0)$ определяется нагрузка на ребра сети при передаче единичных потоков:

$$\alpha(0) \left[\sum_{m=1}^M \xi_m^0(0) \left(x_{mk}^0(0) + x_{m(k+E)}^0(0) \right) \right] = y^0(k). \quad (4)$$

Величина нагрузки численно равна требуемой суммарной пропускной способности, которая считается заданной:

$$\sum_{k=1}^E y^0(k) = D(0). \quad (5)$$

Из соотношений (4) и (5) находится значение $\alpha(0)$:

$$\alpha(0) \sum_{k=1}^E \left[\sum_{m=1}^M \xi_m^0(0) \left(x_{mk}^0(0) + x_{m(k+E)}^0(0) \right) \right] = D(0)$$

и вычисляются значения $y^0(k)$, $k = \overline{1, E}$.

На основе найденных значений формируется сеть $G(1)$, в которой

$$d_k^*(1) := y^0(k), \quad k = \overline{1, E}.$$

В сети $G(1)$ для всех пар узлов $p_a \in P$, $a = \overline{1, M}$, определяется максимальный однопродуктовый поток.

Задача 2. Найти

$$z_a^0(1) = \max \{z_a | (\mathbf{z}, x) \in \mathcal{Z}(\mathbf{d}^*(1))\}$$

при дополнительных условиях

$$z_m = 0, \quad m \neq a, \quad m = \overline{1, M}.$$

На основе последовательности решений задачи 2 вычисляются коэффициенты нормировки

$$\xi_m^0(1) = \frac{1}{z_m^0(1)} \text{ для всех } z_m^0(1) > 0, \quad m = \overline{1, M},$$

и формируется система уравнений для поиска распределения нагрузок

$$\alpha^*(1) \sum_{m=1}^M \xi_m^0(1) \left(x_{mk}^0(1) + x_{m(k+E)}^0(1) \right) = y^1(k), \quad k = \overline{1, E}, \quad \sum_{k=1}^E y^1(k) = D(0).$$

Вычисляем $\alpha^*(1)$ и $y^1(k)$, $k = \overline{1, E}$, и формируем сеть $G(2)$, в которой пропускные способности $d_k^*(2)$ полагаем равными $y^1(k)$, т. е.

$$d_k^*(2) := y^1(k), \quad k = \overline{1, E}.$$

Для полученных решений находим $z_m^*(1) = \alpha^*(1)$ для всех пар $p_m \in P$. Таким образом, все межузловые потоки равны $\alpha^*(1)$, а суммарный поток

$$\sum_{m=1}^M z_m^*(1) = M\alpha^*(1), \quad \|\mathbf{z}^*(1)\| = \alpha^*(1)\sqrt{M}.$$

Все ребра сети $G(2)$ загружены полностью, выделенный ресурс совпадает с нагрузкой на ребро при равных межузловых потоках. Норма вектора ресурсов в «новой» сети $G(2)$:

$$\|\mathbf{y}^1\| = \|\mathbf{d}^*(2)\| = \left[\sum_{k=1}^E d_k^*(2)^2 \right]^{1/2}.$$

4 Вычислительный эксперимент

Вычислительный эксперимент проводился на моделях сетевых систем, представленных на рис. 1. В каждой сети 69 узлов. Пропускные способности ребер — значения $d_k(0)$ — выбирались случайным образом из отрезка $[900, 999]$ и совпадали для ребер, общих для обеих сетей. В кольцевой сети пропускная способность всех добавленных ребер внутреннего кольца составила 900. В ходе вычислительного эксперимента проводилась нормировка и суммарная пропускная способность в обеих сетях была одинакова:

$$\sum_{k=1}^E d_k(0) = D(0) = 68\,256.$$

Результаты вычислительных экспериментов представлены на рис. 2 и в таблице. Толщина ребер на рисунках пропорциональна результирующей нагрузке. В состав внутреннего кольца на рис. 1, б и 2, б вошли четыре узла, которые в базовой сети были висячими, и соответствующие ребра образовали «мостики»

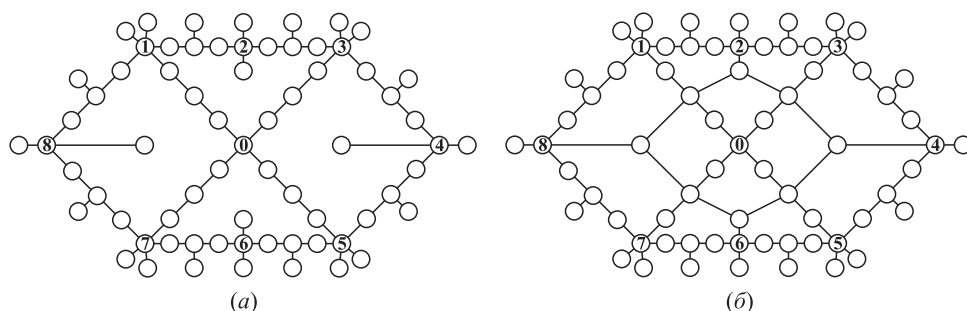


Рис. 1 Базовая (а) и кольцевая (б) сети

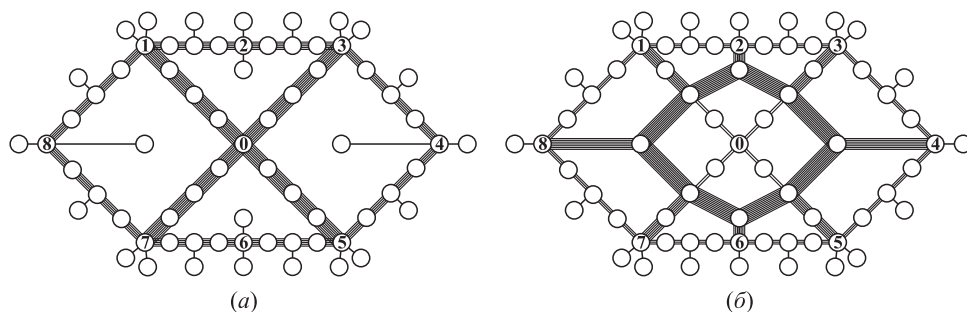


Рис. 2 Результирующая нагрузка на ребра базовой (а) и кольцевой (б) сетей

Межузловые потоки

Сеть	$z^*(1)$	$\ z^*(1)\ $	$Mz^*(1)$	$\ y^*(1)\ $
Базовая	1,54	105	7 220	9 270
Кольцевая	2,26	155	10 600	9 430

для транзитных потоков. Нагрузка на ребра-мостики составляет 1630 единиц, наибольшие значения достигаются на ребрах внутреннего кольца — 2230 единиц, а на внешнем не превышает 830. Напротив, в базовой сети основная нагрузка — 1600 единиц — приходится на радиальные ребра, исходящие из центра, а на ребрах внешнего кольца не превышает 1270 единиц (см. рис. 2, а). В обеих сетях минимальная нагрузка достигается на висячих ребрах: 210 единиц для базовой и 310 единиц для кольцевой. Указанные значения отвечают собственным информационным потокам, которые передаются из каждого узла без учета транзитных. Нагрузки на ребрах внутреннего кольца на порядок превышают собственные потоки инцидентных узлов.

Результаты, представленные в таблице, свидетельствуют, что как суммарный, так и каждый отдельно взятый межузловой поток на 50% больше в кольцевой сети при равной суммарной пропускной способности обеих. Дело в том, что при поиске и передаче максимального потока используются все пути, проходящие через минимальные сечения. В кольцевой сети маршруты соединения короче (не длиннее) для многих пар узлов, что уменьшает транзитную нагрузку и удельные затраты ресурсов. В базовой сети 46 ребер пригодны для транзитной передачи, а в кольцевой — 58, что только на 25% больше, однако при той же пропускной способности это позволяет увеличивать потоки на 50%.

5 Заключение

Предложенный метод позволяет оценивать допустимые распределения межузловых потоков при изменении структуры сети. Построение нескольких путей соединения дает возможность анализировать нагрузку на сеть с учетом обходов для транзитных потоков. Полученные агрегированные показатели для потоков, проходящих через узлы сети, могут быть использованы при построении маршрутно-адресных таблиц и при подготовке графиков-обходов для работы в аварийных режимах.

Литература

1. Малащенко Ю. Е., Назарова И. А. Неоднородность распределения потоков при предельной загрузке многопользовательской сети // Известия РАН. Теория и системы управления, 2022. № 3. С. 81–96.
2. Малащенко Ю. Е., Назарова И. А. Оценка предельных распределений пропускной способности в многопользовательской сети при передаче межузловых потоков по

- кратчайшим маршрутам // Известия РАН. Теория и системы управления, 2022. № 5. С. 90–101.
3. *Ogryczak W., Luss H., Pioro M., Nace D., Tomaszewski A.* Fair optimization and networks: A survey // *J. Appl. Math.*, 2014. Vol. 3. P. 1–25.
 4. *Salimifard K., Bigharaz S.* The multicommodity network flow problem: State of the art classification, applications, and solution methods // *Oper. Res. — Ger.*, 2020. Vol. 22. Iss. 2. P. 1–47.
 5. *Luss H.* Equitable resource allocation: Models, algorithms, and applications. — Hoboken, NJ, USA: John Wiley & Sons, 2012. 420 p.
 6. *Balakrishnan A., Li G., Mirchandani P.* Optimal network design with end-to-end service requirements // *Oper. Res.*, 2017. Vol. 65. Iss. 3. P. 729–750.
 7. *Hahne E. L.* Round-robin scheduling for max-min fairness in data networks // *IEEE J. Sel. Area. Comm.*, 1991. Vol. 9. Iss. 7. P. 1024–1039.
 8. *Radunovic B., Le Boudec J.-Y.* A unified framework for max-min and min-max fairness with applications // *IEEE ACM T. Network.*, 2007. Vol. 15. Iss. 5. P. 1073–1083.
 9. *Nace D., Doan L. N., Klopfenstein O., Bashllari A.* Max-min fairness in multicommodity flows // *Comput. Oper. Res.*, 2008. Vol. 35. Iss. 2. P. 557–573.
 10. *Ros-Giralt J., Tsai W. K.* A lexicographic optimization framework to the flow control problem // *IEEE T. Inform. Theory*, 2010. Vol. 56. Iss. 6. P. 2875–2886.
 11. *Йенсен П., Барнес Д.* Потокоевое программирование / Пер. с англ. — М.: Радио и связь, 1984. 392 с. (*Jensen P. A., Barnes J. W.* Network flow programming. — New York, NY, USA: Wiley, 1980. 408 p.)

Поступила в редакцию 15.08.22

ANALYSIS OF THE DISTRIBUTION OF LOADS AND RESOURCES IN A MULTIUSER NETWORK WITH EQUAL-SHARE INTERNODAL FLOWS

Yu. E. Malashenko and I. A. Nazarova

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

Abstract: An iterative method of resource distribution for the formation of a communication network with equal internodal flows is considered. Within the framework of a formal mathematical model, the throughput capacities of edges are considered as components of a vector of resources of various types that are required for the transmission of different types of flows. The network graph and the sum of the edge capacity are considered to be given. An algorithmic procedure for selecting throughput capacities is proposed in which all internodal flows are equal and the maximum load of all network edges is achieved. When searching for the corresponding distributions, the values of the maximum single-product flows are used which are determined for each pair of nodes. The computational experiments were carried out for the networks with various structural features.

Keywords: multicommodity flow model; network resource distribution method; equalizing distribution of interstitial flows; network peak load

DOI: 10.14357/08696527220307

References

1. Malashenko, Yu. E., and I. A. Nazarova. 2022. Heterogeneous flow distribution at the peak load in the multiuser network. *J. Comput. Sys. Sc. Int.* 61(3):372–387.
2. Malashenko, Yu. E., and I. A. Nazarova. 2022. Estimate of resource distribution with the shortest paths in the multiuser network. *J. Comput. Sys. Sc. Int.* 61(4):599–610.
3. Ogryczak, W., H. Luss, and M. Pioro. 2014. Fair optimization and networks: A survey. *J. Appl. Math.* 3:1–25.
4. Salimifard, K., and S. Bigharaz. 2020. The multicommodity network flow problem: State of the art classification, applications, and solution methods. *Oper. Res. — Ger.* 18(3):1–47.
5. Luss, H. 2012. *Equitable resource allocation: Models, algorithms, and applications*. Hoboken, NJ: John Wiley & Sons. 420 p.
6. Balakrishnan, A., G. Li, and P. Mirchandani. 2017. Optimal network design with end-to-end service requirements. *Oper. Res.* 65(3):729–750.
7. Hahne E. L. 1991. Round-robin scheduling for max-min fairness in data networks. *IEEE J. Sel. Area. Comm.* 9(7):1024–1039.
8. Radunovic, B., and J.-Y. Le Boudec. 2007. A unified framework for max-min and min-max fairness with applications. *IEEE ACM T. Network.* 15(5):1073–1083.
9. Nace, D., L. Doan, and O. Klopfenstein. 2008. Max-min fairness in multicommodity flows. *Comput. Oper. Res.* 35(2):557–573.
10. Ros-Giralt, J., and W. K. Tsai. 2010. A lexicographic optimization framework to the flow control problem. *IEEE T. Inform. Theory* 56(6):2875–2886.
11. Jensen, P. A., and J. W. Barnes. 1980. *Network flow programming*. New York, NY: Wiley. 408 p.

Received August 15, 2022

Contributors

Malashenko Yuri E. (b. 1946) — Doctor of Science in physics and mathematics, principal scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; malash09@ccas.ru

Nazarova Irina A. (b. 1966) — Candidate of Science (PhD) in physics and mathematics, scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; irina-nazar@yandex.ru

ВАРИАНТЫ САМОСИНХРОННЫХ РЕГИСТРОВ СДВИГА*

*Ю. А. Степченков¹, Ю. Г. Дьяченко², Ю. В. Рождественский³,
Н. В. Морозов⁴, Д. Ю. Степченков⁵, Д. Ю. Дьяченко⁶*

Аннотация: Затрагиваются проблемы проектирования и использования самосинхронных (СС) регистров сдвига (РС). Самосинхронные схемы имеют свою специфику: двухфазную дисциплину работы, избыточное кодирование информации и др. За счет этого они обладают рядом преимуществ в сравнении с синхронными аналогами: независимость поведения от задержек элементов, обнаружение лобых константных неисправностей и др. Статья рассматривает варианты реализации СС-регистра с различными опциями, включая установку спейсера и предустановку заданного значения в каждом разряде РС. Предлагаемые варианты обладают разной функциональностью, сложностью и быстродействием. Регистры сдвига на основе RS-триггеров обладают минимальными аппаратными затратами. Регистры сдвига на основе гистерезисных триггеров (Г-триггеров) характеризуются лучшим быстродействием. Статья анализирует их характеристики и обосновывает рекомендации по их использованию в качестве последовательно-параллельного или параллельно-последовательного порта, FIFO (First Input, First Output).

Ключевые слова: самосинхронная схема; гистерезисный триггер; RS-триггер; регистр сдвига; FIFO; последовательно-параллельный порт; аппаратные затраты; быстродействие

DOI: 10.14357/08696527220308

1 Введение

Подавляющее большинство цифровых устройств в современном мире реализовано в базисе синхронной схемотехники. Самосинхронная схемотехника [1, 2] стала перспективной альтернативой синхронным схемам в областях применения,

* Исследование выполнено при поддержке Российского научного фонда (проект 22-19-00237).

¹ Федеральное исследовательское учреждение «Информатика и управление» Российской академии наук, YStepchenkov@ipiran.ru

² Федеральное исследовательское учреждение «Информатика и управление» Российской академии наук, diauga@mail.ru

³ Федеральное исследовательское учреждение «Информатика и управление» Российской академии наук, YRogdest@ipiran.ru

⁴ Федеральное исследовательское учреждение «Информатика и управление» Российской академии наук, NMorozov@ipiran.ru

⁵ Федеральное исследовательское учреждение «Информатика и управление» Российской академии наук, Stepchenkov@mail.ru

⁶ Федеральное исследовательское учреждение «Информатика и управление» Российской академии наук, diaden87@gmail.com

характеризующихся нестабильным и низким напряжением питания, изменчивыми и неблагоприятными условиями эксплуатации цифрового оборудования. Самосинхронные схемы обладают рядом преимуществ в сравнении с синхронными аналогами [3]: отсутствием «гонок» между сигналами, устойчивым функционированием при любых условиях эксплуатации, независимостью поведения схемы от реальных задержек ее элементов, отсутствием «дерева» глобальной синхронизации, лучшей устойчивостью к логическим сбоям [4] и др.

В то же время СС-схемы имеют и ряд недостатков [3], к числу которых относятся избыточные аппаратные затраты, в ряде случаев ухудшение быстродействия из-за двухфазной дисциплины работы и наличия индикаторной подсхемы, обеспечивающей контроль окончания переключения СС-схемы в каждую фазу работы. Однако в критических областях применения недостатки СС-схем компенсируются их преимуществами.

Некомпетентный логический синтез СС-схемы способен привести к сверхизбыточным аппаратным затратам и дополнительному снижению быстродействия. Цифровые устройства с регулярной структурой целесообразно вставлять в проектируемую СС-схему в виде готовых сложных функциональных блоков, специфицированных под конкретное применение [5, 6]. Данный подход обеспечивает реализацию СС-схемы произвольного типа с минимальными аппаратными затратами и/или максимальным быстродействием в пределах эквифазной зоны [7] в произвольном логическом базисе.

В связи с этим становится актуальной задача разработки типовых схемотехнических решений СС-регистров сдвига. В данной статье обосновывается базовый набор СС-регистров сдвига, предлагаются варианты их схемотехнической реализации и дается сравнительный анализ их характеристик.

2 Типовые самосинхронные регистры сдвига

Регистры сдвига используются для реализации последовательно-параллельных портов, интерфейсов последовательной передачи данных (FIFO) последовательных устройств многоразрядного сдвига. В синхронной схемотехнике они работают по принципу последовательной передачи бита информации по цепочке двухтактных D-триггеров. В любой момент времени каждый D-триггер содержит бит информации.

В СС-схемотехнике типовой РС может быть организован аналогично синхронному аналогу — на двухтактном D-триггере (D0RE24, рис. 1, а) или RS-триггере (S0RRE2, рис. 1, б). Большая сложность триггера D0RE24 (54 комплементарных металл–диэлектрик–полупроводник (КМДП) транзисторов против 48 у S0RRE2) обеспечивает возможность сокращения информационных сигналов в схеме за счет использования унарного информационного входа вместо парафазного или бифазного [2].

Реализации трехразрядного РС на D- и RS-триггере представлены на рис. 2. В сравнении с синхронными аналогами они имеют дополнительный индикатор-

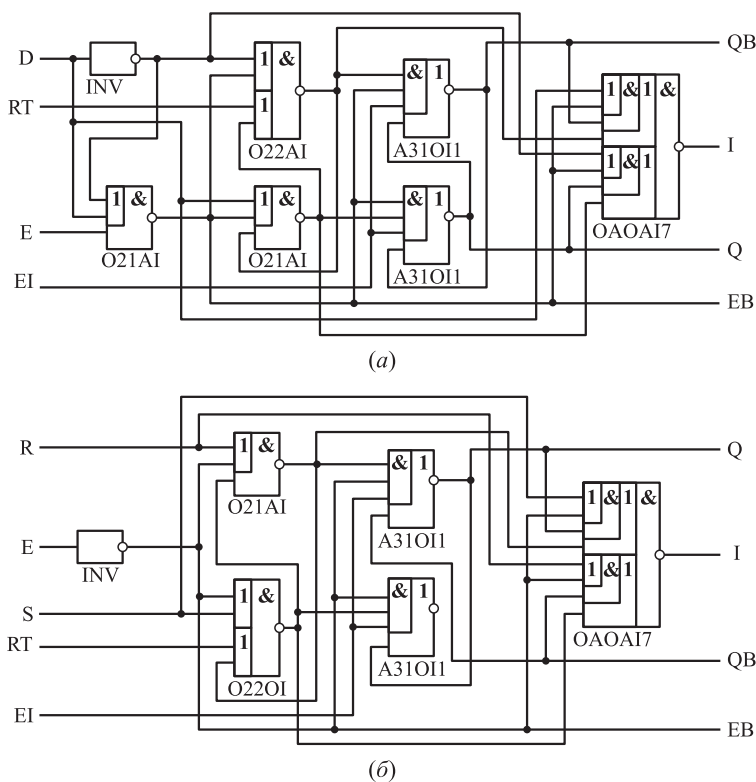


Рис. 1 Функциональные схемы разряда СС-регистра сдвига на D-триггере (а) и на RS-триггере (б)

ный элемент — трехвходовый Г-триггер ГІЗ [8]. При построении регистров с числом разрядов больше трех индикация реализуется пирамидальной схемой из Г-триггеров.

Возможным решением является реализация РС на Г-триггерах. Прототипом такого РС служит СС-конвейер, в котором хранение информации оказывается полуплотным. Схема разряда РС на Г-триггерах приведена на рис. 3. Ее сложность равна 32 транзисторам. Вход RT устанавливает оба Г-триггера в нулевой спейсер при $RT = 0$. На рис. 4 показана схема трехразрядного РС на Г-триггерах.

Алгоритм работы РС на Г-триггерах заключается в следующем:

- вначале все разряды РС устанавливаются в нулевой спейсер ($RT = 0$);
- на парафазный вход (X, XB) подается рабочее состояние, оно записывается по цепочке во все разряды, заполняя все элементы регистра D1–D6;

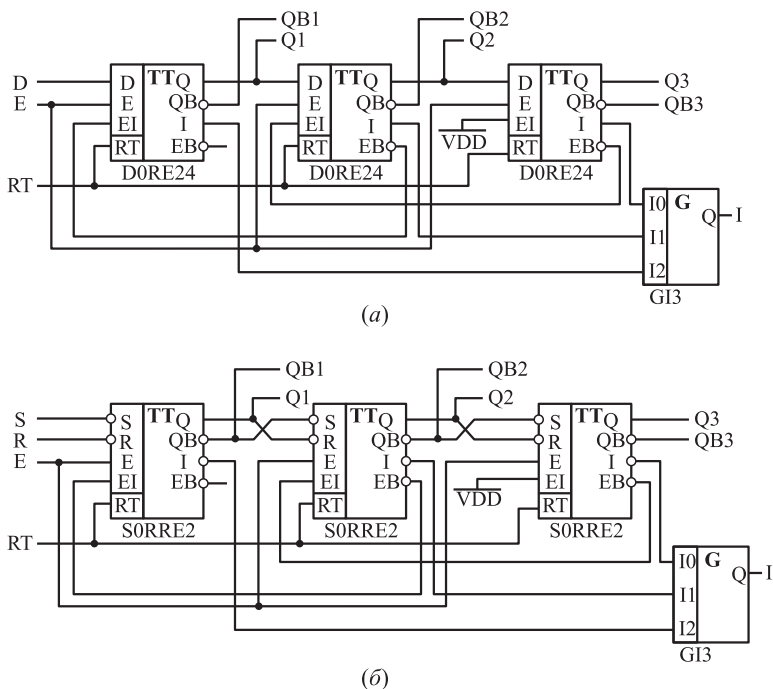


Рис. 2 Схема трехразрядного СС-регистра сдвига на D-триггере (а) и на RS-триггере (б)

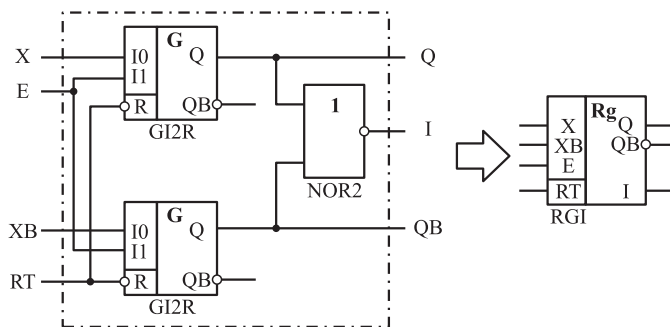


Рис. 3 Схема разряда СС-регистра сдвига на Г-триггерах

- на вход (X, XB) поступает нулевой спейсер, который распространяется по элементам D1–D5, но элемент D6 сохраняет свое рабочее состояние;
- новое рабочее состояние, поступающее на вход (X, XB), запишется в первые четыре элемента регистра (D1–D4);

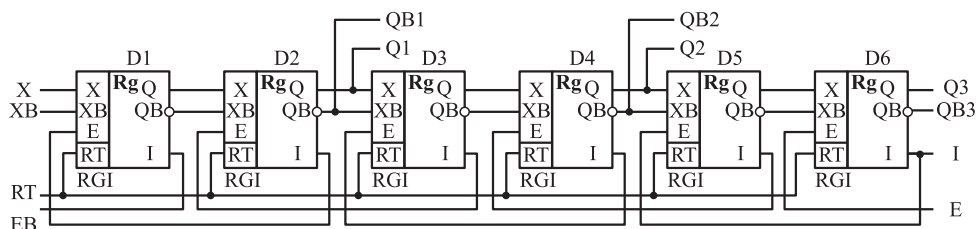


Рис. 4 Схема трехразрядного СС-регистра сдвига на Г-триггерах

- дальнейшее чередование спейсера и рабочего состояния на входе (X, XB) приведет к заполнению регистра тремя рабочими состояниями в элементах D2, D4 и D6 и спейсером в элементах D1, D3 и D5;
- переключение входа управления E из «1» в «0» разрешит элементу D6 перейти в спейсер, что, в свою очередь, разрешит рабочему состоянию из D4 записаться в D5, затем спейсеру из D3 записаться в D4 и т. д. В результате биты информации, хранимые регистром, сдвинутся на один элемент регистра и будет разрешена запись нового бита информации в элемент D1.

Регистр на рис. 4 функционирует как «полуплотное FIFO» [9]. Для реализации n -разрядного РС требуется $2n$ разрядов, и общие аппаратные затраты на рис. 4 оказываются в 1,2–1,3 раза больше, чем в регистрах на рис. 2. Его преимущества в сравнении с ними: парафазное кодирование информационных входов и выходов и лучшее в 2,2 раза быстродействие.

Одно из практических применений РС — реализация последовательного выходного интерфейса. Оно требует параллельной записи передаваемой информации в регистр перед началом передачи.

3 Самосинхронный регистр сдвига с параллельной записью

Устройства последовательного выходного интерфейса основаны на параллельной записи в РС информации и «выталкивании» ее по последовательному каналу. На рис. 5 показана реализация разряда РС с параллельной записью. Он отличается от разряда РС на рис. 1, б наличием входов разрешения параллельной записи WE (WE = 1 — разрешение записи; WE = 0 — отсутствие записи, режим сдвига), СС-установки P (P = 1 — установка выхода разряда РС в состояние Q = 1, QB = 0; P = 0 — отсутствие установки), СС-сброса C (C = 1 — сброс выхода разряда регистра в состояние Q = 0, QB = 1; C = 0 — отсутствие сброса) и дополнительного выхода IW, индицирующего успешное завершение СС-записи (IW = 0 — запись завершена; IW = 1 — запись отсутствует). Аппаратные затраты разряда регистра на рис. 5 составляют 64 КМДП-транзистора.

На рис. 6 показана реализация трехразрядного РС с параллельной СС-записью, разряды которого реализованы схемой на рис. 5. Элемент D7 формирует общий индикаторный выход II регистра с учетом режима его работы.

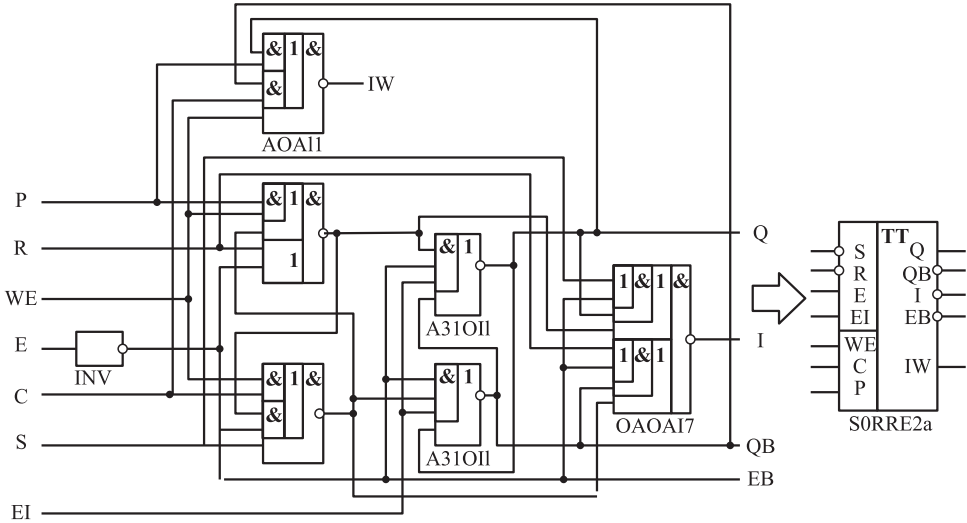


Рис. 5 Схема одного разряда РС с СС параллельной записью

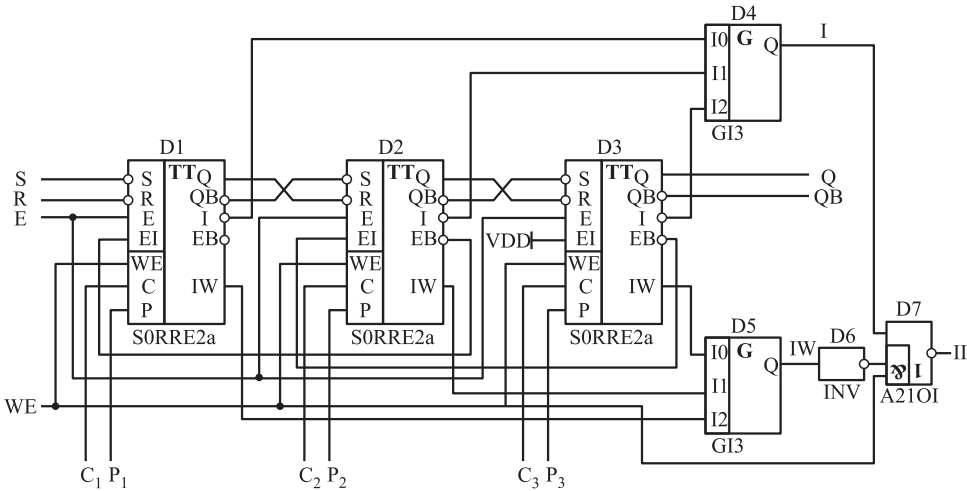


Рис. 6 Схема трехразрядного РС с СС параллельной записью

В режиме сдвига вход разрешения записи $WE = 0$ и цепи параллельной записи не влияют на индикаторный выход II . Параллельная запись в регистр выполняется при $E = 0$ путем подачи на входы установки C_i и P_i требуемых значений ($C_i = 1$ и $P_i = 0$ для установки в i -м разряде регистра «0»; $C_i = 0$ и $P_i = 1$ для установки в i -м разряде регистра «1») и задании $WE = 1$. Сигнал WE должен индексировать входы установки регистра. Аппаратные затраты СС-регистра

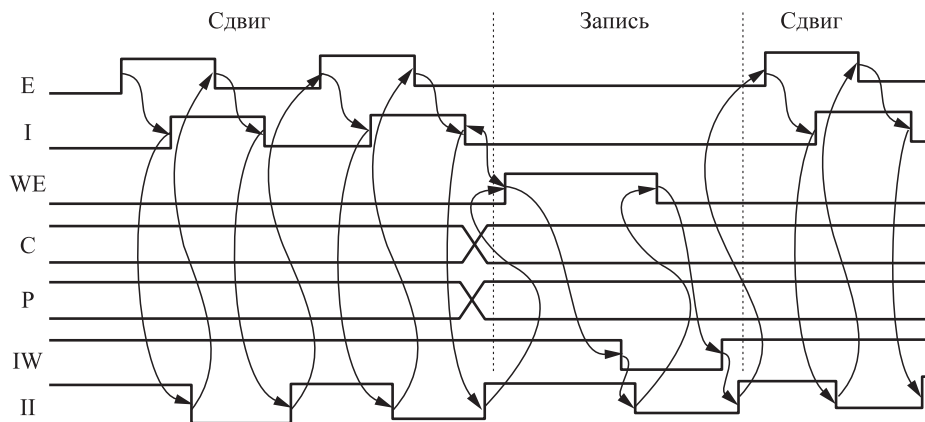


Рис. 7 Диаграмма работы СС-регистра сдвига с параллельной СС-записью

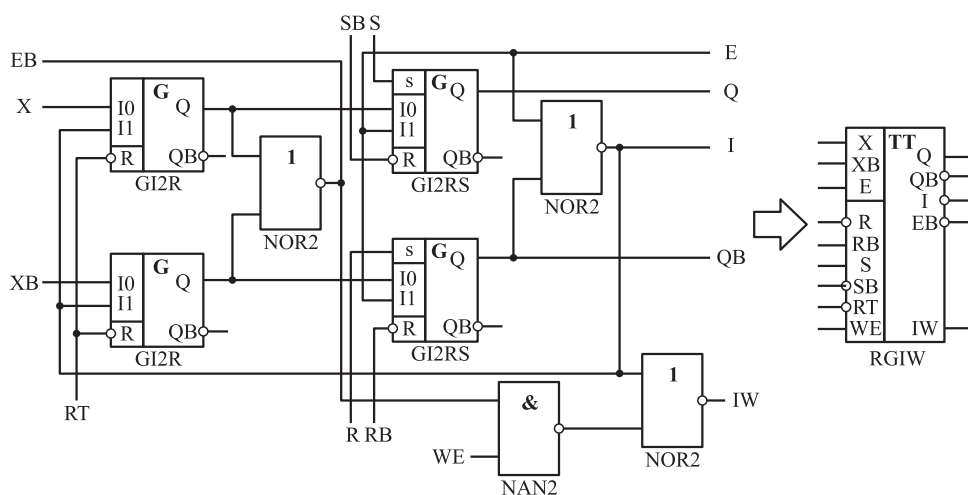


Рис. 8 Схема одного разряда РС на Г-триггерах с параллельной записью

на рис. 6 составляют 232 КМДП-транзистора. Рисунок 7 показывает временную диаграмму работы СС-регистра сдвига в терминах входов управления и записи и индикаторных выходов. Тонкие линии со стрелками показывают причинно-следственные зависимости между сигналами схемы регистра.

Разряд РС на Г-триггерах с параллельной записью показан на рис. 8. Он устанавливается в нулевой спейсер при $R = RB = RT = S = SB = 0$, записывает бит «0» при $R = RB = 1$ и $RT = S = SB = 0$ или бит «1» при $R = RB = RT = 0$ и $S = SB = 1$.

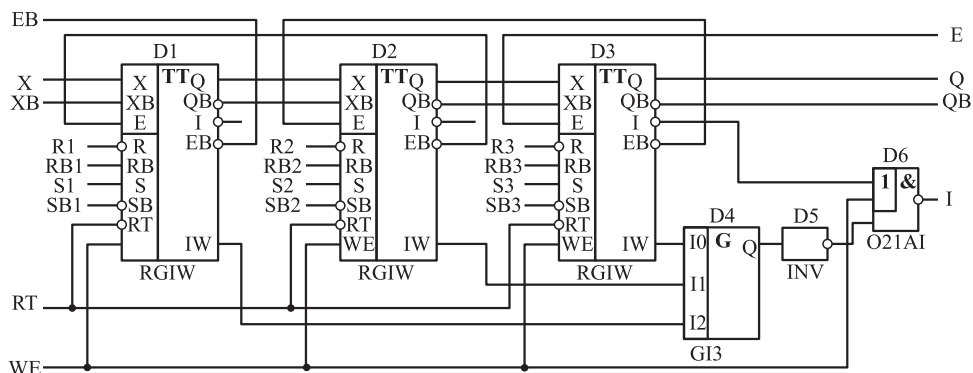


Рис. 9 Регистр сдвига на Г-триггерах с параллельной записью

Параллельная запись выполняется после переключения разряда РС в спейсер ($Q = QB = 0$). После инициации параллельной записи $WE = 1$ и подачи сигналов $RT = 0$ и требуемой комбинации входов R, RB, S и SB переключение выхода IW в значение $IW = 1$ подтвердит успешное окончание параллельной записи в разряд РС. Установка $WE = 0$ и затем $RT = 1$ и $R = S = 0, RB = SB = 1$ возвращает разряд РС в спейсер. Сложность реализации схемы составляет 76 КМДП-транзисторов. На рис. 9 показан 3-разрядный СС-регистр сдвига с параллельной записью на Г-триггерах.

Индикаторная подсхема, включающая элементы D4–D6, подтверждает успешное окончание параллельной записи в регистр. В результате общая сложность реализации СС-регистра сдвига на рис. 9 составляет 252 КМДП-транзистора, что на 9% выше сложности регистра на рис. 6.

С учетом того, что для параллельной записи в регистр на рис. 9 требуется сформировать вдвое больше сигналов предустановки R^* и S^* , суммарные затраты варианта регистра на Г-триггерах окажутся еще хуже в сравнении с затратами варианта на RS-триггере (см. рис. 6).

Однако моделирование вариантов СС-регистра сдвига, изображенных на рис. 6 и 9 с учетом их окружения, обеспечивающего работу РС в режимах параллельной СС-записи и СС-сдвига, показывает, что суммарная длительность рабочей и спейсерной фаз в РС на рис. 6 в 1,5 раза больше, чем в регистре на рис. 9.

Таким образом, РС с параллельной записью (параллельно-последовательный порт) на RS-триггере (см. рис. 6) обеспечивает минимальную сложность, а регистр на Г-триггерах (см. рис. 9) гарантирует лучшее быстродействие.

4 Заключение

Рассмотренные варианты реализации разряда СС-регистра сдвига и самого СС-регистра сдвига не исчерпывают всего многообразия сдвиговых СС-регист-

ров. Ввод в описанные схемы входов асинхронной или СС-предустановки позволяет получить широкую номенклатуру практически значимых реализаций СС-регистра сдвига.

Самосинхронный РС типа FIFO или последовательно-параллельный порт целесообразно реализовывать на Г-триггерах с предустановкой спейсера, если требуется минимальная латентность и парафазное кодирование параллельного выхода регистра. В противном случае лучшим решением будет использование двухтактного RS-триггера.

Самосинхронный РС с параллельной записью целесообразно строить на RS-триггерах с СС-установкой и сбросом, если целевым критерием служит минимальная сложность схемы. Аналогичный регистр на Г-триггерах оказывается в 1,09 раза сложнее, однако последний обладает в 1,5 раза лучшим быстродействием.

Литература

1. Muller D., Bartky W. A theory of asynchronous circuits // Annals Computation Laboratory Harvard University, 1959. Vol. 29. P. 204–243.
2. Varshavsky V., Kishinevsky M., Marakhovsky V., et al. Self-timed control of concurrent processes. — Kluwer Academic Publs., 1990. 245 p.
3. Плеханов Л. П., Степченков Ю. А. Экспериментальная проверка некоторых свойств строго самосинхронных электронных схем // Системы и средства информатики, 2006. Т. 16. № 2. С. 476–485.
4. Соколов И. А., Степченков Ю. А., Рождественский Ю. В., Дьяченко Ю. Г. Приближенная оценка эффективности синхронной и самосинхронной методологий в задачах проектирования сбоеустойчивых вычислительно-управляющих систем // Автоматика и телемеханика, 2022. № 2. С. 122–132.
5. Плеханов Л. П., Денисов А. Н., Дьяченко Ю. Г., Степченков Ю. А., Мамонов Д. И., Степченков Д. Ю. Синтез самосинхронных схем в базе БМК // Наноиндустрия, 2020. № S96-2. С. 460–470. doi: 10.22184/1993-8578.2020.13.3s.460.470.
6. Плеханов Л. П., Денисов А. Н., Дьяченко Ю. Г., Мамонов Д. И., Морозов Н. В., Степченков Д. Ю. Реализация синтеза самосинхронных схем в базе БМК // Наноиндустрия, 2021. Т. 14. № S7. С. 395–397.
7. Zakharov V., Stepchenkov Y., Diachenko Y., Rogdestvenski Y. Self-timed circuitry retrospective // Conference (International) Engineering Technologies and Computer Science Proceedings. — Piscataway, NJ, USA: IEEE, 2020. P. 58–64. doi: 10.1109/EnT48576.2020.00018.
8. Степченков Ю. А., Денисов А. Н., Дьяченко Ю. Г. и др. Библиотека функциональных ячеек для проектирования самосинхронных полузаказных БМК микросхем серий 5503/5507. — Сер. «Полузаказные БИС на БМК серий 5503 и 5507». — М.: Техносфера, 2017. Т. 4. 376 с. <http://www.technosphera.ru/lib/book/497>.
9. Автоматное управление асинхронными процессами в ЭВМ и дискретных системах / Под ред. В. И. Варшавского. — М.: Наука, 1986. 400 с.

Поступила в редакцию 13.06.22

SELF-TIMED SHIFT REGISTER CASES

*Yu. A. Stepchenkov, Yu. G. Diachenko, Yu. V. Rogdestvenski, N. V. Morozov,
D. Yu. Stepchenkov, and D. Yu. Diachenko*

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

Abstract: The paper discusses the problems of designing and using self-timed (ST) shift registers. Self-timed circuits have their specifics: two-phase work discipline, redundant information coding, etc. Due to this, they have some advantages compared with synchronous counterparts: independence of behavior from cell delays, detection of any stuck faults, etc. The article considers implementation options for the ST shift register with various options, including setting to a spacer and presetting a fixed value in each bit of the shift register. The proposed options have different functionality, complexity, and performance. Shift registers based on RS-flip-flops have minimal hardware costs, while shift registers based on hysteretic triggers have better performance. The article analyzes shift register’s characteristics and substantiates recommendations for their use as a serial-to-parallel port, parallel-to-serial port, or FIFO (First Input, First Output).

Keywords: self-timed circuit; hysteretic trigger; RS-flip-flop; shift register; FIFO; serial-to-parallel port; hardware costs; performance

DOI: 10.14357/08696527220308

Acknowledgments

The research was supported by the Russian Science Foundation (project No. 22-19-00237).

References

1. Muller, D., and W. Bartky. 1959. A theory of asynchronous circuits. *Annals Computation Laboratory Harvard University* 29:204–243.
2. Varshavsky, V., M. Kishinevsky, V. Marakhovsky, et al. 1990. *Self-timed control of concurrent processes*. Kluwer Academic Publs. 245 p.
3. Plekhanov, L. P., and Yu. A. Stepchenkov. 2006. Eksperimental’naya proverka nekotorykh svoystv strogo samosinkhronnykh skhem [Experimental verification of some properties of strictly self-timed circuits]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 16:476–485.
4. Sokolov, I. A., Yu. A. Stepchenkov, Yu. V. Rozhdestvenski, and Yu. G. Diachenko. 2022. Approximate evaluation of the efficiency of synchronous and self-timed methodologies in problems of designing failure-tolerant computing and control systems. *Automat. Rem. Contr.* 83(2):264–272.
5. Plekhanov, L. P., A. N. Denisov, Yu. G. Diachenko, Yu. A. Stepchenkov, D. I. Mamonov, and D. Yu. Stepchenkov. 2020. Sintez samosinkhronnykh skhem v bazise BMK

- [Self-timed circuit synthesis in gate array basis]. *Nanoindustriya* [Nanoindustry] S96-2:460–470. doi: 10.22184/1993-8578.2020.13.3s.460.470.
6. Plekhanov, L. P., A. N. Denisov, Yu. G. Diachenko, D. I. Mamonov, N. V. Morozov, and D. Yu. Stepchenkov. 2021. Realizatsiya sinteza samosinkhronnykh skhem v bazise BMK [Implementing self-timed circuits synthesis in gate array basis]. *Nanoindustriya* [Nanoindustry] 14(S7):395–397.
 7. Zakharov, V., Y. Stepchenkov, Y. Diachenko, and Y. Rogdestvenski. 2020. Self-timed circuitry retrospective. *Conference (International) Engineering Technologies and Computer Science Proceedings*. Piscataway, NJ: IEEE. 58–64. doi: 10.1109/EnT48576.2020.00018.
 8. Stepchenkov, Yu. A., A. N. Denisov, Yu. G. Diachenko, et al. 2017. *Biblioteka funktsional'nykh yacheek dlya proektirovaniya samosinkhronnykh poluzakaznykh BMK mikroskhem seriy 5503/5507* [Library of functional cells for designing self-timed semicustom chips of the 5503 and 5597 series]. Ser. “Poluzakaznye BIS na BMK seriy 5503 i 5504” [Ser. Semicustom chips of the 5503 and 5597 series]. Moscow: Tekhnosfera. Vol. 4. 376 p.
 9. Varshavsky, V. I., ed. 1986. *Avtomatnoe upravlenie asinkhronnymi protsessami v EVM i diskretnykh sistemakh* [Automata control of asynchronous processes in computers and discrete systems]. Moscow: Nauka. 400 p.

Received June 13, 2022

Contributors

Stepchenkov Yuri A. (b. 1951) — Candidate of Science (PhD) in technology, leading scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; YStepchenkov@ipiran.ru

Diachenko Yuri G. (b. 1958) — 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 119133, Russian Federation; diaura@mail.ru

Rogdestvenski Yuri V. (b. 1952) — Candidate of Science (PhD) in technology, leading scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; YRogdest@ipiran.ru

Morozov Nikolay V. (b. 1956) — senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; NMorozov@ipiran.ru

Stepchenkov Dmitri Y. (b. 1973) — senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; Stepchenkov@mail.ru

Diachenko Denis Y. (b. 1987) — engineer-researcher, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; diaden87@gmail.com

МЕТОДЫ ПОИСКА ИМПЛИЦИТНЫХ ЛОГИКО-СЕМАНТИЧЕСКИХ ОТНОШЕНИЙ В МОНОЯЗЫЧНЫХ ТЕКСТАХ*

А. А. Гончаров¹

Аннотация: Анализируется проблема целенаправленного поиска фрагментов с имплицитными логико-семантическими отношениями (ЛСО) в монологичных текстах на естественном языке. Чтобы ответить на вопрос о том, возможен ли такой поиск, был проведен обзор работ с результатами исследований имплицитных ЛСО. Цель обзора — выявление методов, которые применяются в этих работах для создания массивов текстовых фрагментов с имплицитными ЛСО. Показано, что используются в основном два метода: (1) автоматическая генерация фрагментов с имплицитными ЛСО на основе фрагментов с эксплицитными ЛСО; (2) выгрузка фрагментов с имплицитными ЛСО из текстовых корпусов, размеченных вручную. Ни один из этих методов не может считаться методом собственно поиска, так как не подразумевает выявления фрагментов с имплицитными ЛСО в естественном тексте и их последующего извлечения. Таким образом, можно констатировать, что для монологичных текстов задача целенаправленного поиска таких фрагментов на данный момент, в сущности, не решена.

Ключевые слова: имплицитность; имплицитные логико-семантические отношения; целенаправленный поиск; обработка естественного языка; извлечение информации

DOI: 10.14357/08696527220309

1 Введение

В 2021 г. под эгидой Ассоциации по компьютерной лингвистике (*Association for Computational Linguistics, ACL*) впервые прошел семинар, полностью посвященный явлениям имплицитности и недоспецификации в языке (*implicit and underspecified phenomena in language*): “UnImplicit 2021: 1st Workshop on Understanding Implicit and Underspecified Language” [1]. Мероприятие, по всей видимости, имеет шансы стать ежегодным: в 2022 г. семинар состоялся во второй раз [2]. Во введении к сборнику статей, представленных на семинаре, отмечается, что сложность имплицитности и недоспецификации обусловлена тем, что

* Исследование выполнено с использованием ЦКП «Информатика» ФИЦ ИУ РАН.

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, a.gonch48@gmail.com

для их обработки может быть необходимо: (1) анализировать широкий контекст; (2) использовать логический вывод (*symbolic inference*); (3) использовать рассуждения на основе здравого смысла (*common-sense reasoning*); (4) анализировать некомпозиционные и не выраженные на уровне лексики значения [1, с. III]. Появление подобного семинара свидетельствует о том, что создание методов работы с имплицитной информацией становится все более актуальным направлением исследований в сфере обработки текстов на естественном языке и компьютерной лингвистики.

Настоящая статья посвящена одной из форм имплицитности — имплицитным ЛСО в тексте (подробнее об этом понятии см. [3]). Под ЛСО понимаются отношения связности между небольшими фрагментами текста, потенциально способные иметь показатель (прежде всего коннектор)¹. Как правило, при помощи ЛСО связываются фрагменты с пропозициональным² содержанием; для упрощения в качестве таких фрагментов могут рассматриваться предложения. Так, пример (1) иллюстрирует эксплицитное ЛСО причины: между фрагментами «он опоздал» и «автобуса долго не было», для наглядности заключенными в квадратные скобки, есть коннектор «потому что», выражающий это ЛСО:

(1) [Он опоздал,] **потому что** [автобуса долго не было.]

Пример (2) иллюстрирует **имплицитное ЛСО** причины: связываемые фрагменты совпадают с теми, что приведены в (1), и, хотя коннектор между ними отсутствует, ЛСО причины может быть выведено читающим в ходе интерпретации смыслового соотношения фрагментов:

(2) [Он опоздал,] [автобуса долго не было.]

Фрагменты текста, связанные некоторым ЛСО, будем называть **аргументами ЛСО**; отрезок текста, включающий аргументы ЛСО и его показатель (если таковой имеется), будем называть **фрагментом с ЛСО**. В зависимости от наличия показателя фрагменты с ЛСО делятся на два типа: **фрагменты с эксплицитными ЛСО** и **фрагменты с имплицитными ЛСО**.

Если при поиске и определении семантики эксплицитных ЛСО исследователи имеют возможность использовать для решения этих задач показатели ЛСО, то при работе с имплицитными ЛСО эта возможность очевидным образом недоступна. Между тем доля фрагментов с имплицитными ЛСО в общей массе фрагментов с ЛСО весьма высока, хотя и варьирует в зависимости от типа анализируемых текстов и методов анализа: так, по данным [7, с. 587], она составляет $\approx 69\%$ в корпусе устных текстов и $\approx 57\%$ в корпусе статей из газеты “Wall Street

¹ В работах, где рассматриваются такие отношения, для их обозначения используются разнообразные термины: отношения связности, риторические, дискурсивные или риторико-семантические отношения; также друг от друга сильно отличаются наборы этих отношений (см. подробнее [4]). Основания классификаций отношений, которые в данной статье называются логико-семантическими, подробно проанализированы в [5].

² См. об этом понятии [6, с. 401].

Journal”; по данным [8, с. 5], она равна $\approx 45\%$ ¹ (в корпусе, также включающем статьи из газеты “Wall Street Journal”).

Цель настоящей статьи состоит в поиске ответа на вопрос о том, возможно ли целенаправленно искать фрагменты с имплицитными ЛСО в моноязычных текстах. Для ответа на этот вопрос проводится обзор работ, в которых описываются различные эксперименты по обработке имплицитных ЛСО средствами информатики. Особое внимание уделяется методам формирования массивов фрагментов с имплицитными ЛСО.

2 Предыстория

Одной из наиболее ранних работ, оказавших значительное влияние на последующие исследования имплицитных ЛСО, стала опубликованная в 2002 г. работа [9], написанная Д. Марку и А. Эчихаби, исследователями из Института информационных наук (*Information Sciences Institute*) Университета Южной Калифорнии. Целью авторов была разработка подхода к распознаванию четырех ЛСО (как эксплицитных, так и имплицитных) в тексте, т. е. имплицитные ЛСО отдельно ими не рассматривались.

Исходя из того что, с одной стороны, не все ЛСО в тексте имеют показатели, а с другой стороны, не существует ни надежных средств семантической интерпретации, ни баз знаний общего назначения (*general purpose knowledge bases*), которые позволили бы сделать нужные умозаключения и распознать некоторое ЛСО в тексте в отсутствие показателя, Д. Марку и А. Эчихаби решили разработать такую технологию, которая давала бы возможность определить семантику ЛСО, но не опиралась бы исключительно на их показатели. Входными данными должны были служить фрагменты с ЛСО (эксплицитными или имплицитными); выходными — эти фрагменты, снабженные меткой, указывающей на семантику ЛСО, которым связаны аргументы, составляющие каждый из фрагментов, или меткой, указывающей на отсутствие какого-либо ЛСО.

Авторы выдвинули гипотезу о том, что существуют «пары лексических единиц, которые могут давать подсказки относительно ЛСО между фрагментами, в которых употреблены эти лексические единицы»² [9, с. 372]. В качестве иллюстрации они привели пример (3), где ЛСО противопоставления (*Contrast*) имплицитно, но о его наличии свидетельствует тот факт, что первый аргумент ЛСО содержит слово *good*, а второй — *fails*; знание о таких парах и должно было бы позволить определять семантику ЛСО даже в отсутствие показателя:

(3) John is **good** in math and sciences. Paul **fails** almost every class he takes.

¹Доля вычислялась на основе данных из таблицы, приведенной в [8, с. 5], где общим числом фрагментов с ЛСО считалась сумма чисел из строк Explicit, Implicit, AltLex и AltLexC (столбцы Intra-S PDTB-3 и Inter-S PDTB-3), а общим числом фрагментов с имплицитными ЛСО — сумма чисел из строки Implicit (те же столбцы).

²Здесь и далее перевод мой. — А. Г.

‘Джон **силен** в математике и естественных науках. Пол **не успевает** почти по всем предметам, которые он изучает’.

Как следствие, было необходимо:

- (1) получить массив данных, где можно будет искать такие пары;
- (2) разработать средства, которые позволят определять, какие пары лексических единиц с высокой вероятностью встречаются во фрагментах с некоторым ЛСО;
- (3) разработать средства, которые позволят определять семантику ЛСО между аргументами с использованием выявленных пар.

Для получения массива данных авторы исходили из предположения, согласно которому между показателем и ЛСО можно устанавливать прямую связь (хотя и признавали, что при таком подходе часть примеров неизбежно окажутся ошибочными), что во многом обуславливалось широкой семантикой рассматриваемых ЛСО¹. Действительно, сами авторы отмечали: «Преимущество работы с грубо определенными дискурсивными отношениями [= ЛСО] состоит в том, что они позволяют нам автоматически создавать наборы данных, содержащие довольно мало шума и подходящие для обучения» [9, с. 370]. Используемый список ЛСО включал:

- (1) *Contrast* (противопоставление);
- (2) *Cause–Explanation–Evidence* (причина–объяснение–обоснование);
- (3) *Condition* (условие);
- (4) *Elaboration* (детализация) [9, с. 370–371].

Обратная сторона использования ЛСО со столь широкой семантикой — низкая информативность получаемых данных. Так, с позиции других теорий связности текста, множество примеров, относимых в описываемом подходе к ЛСО *Contrast*, может рассматриваться как несколько самостоятельных множеств: множество примеров на ЛСО противопоставления, на ЛСО уступки, на ЛСО сопоставления и т. п.; то же и для других ЛСО из [9].

Итак, для четырех вышеперечисленных отношений авторы сформировали список «однозначных» показателей и создали с их использованием шаблоны для извлечения фрагментов с ЛСО. К примеру, приняв *but* (‘но’) за однозначный

¹В том же ключе действовали и авторы работы [10]. Используя набор из четырех ЛСО с широкой семантикой, они установили, что многие показатели ЛСО могут считаться однозначными, однако данных о том, насколько однозначными окажутся эти показатели, если использовать большее число ЛСО с более узкой семантикой, в работе не приводится. Проблеме многозначности/однозначности показателей ЛСО посвящена обширная литература, однако ее обсуждение выходит за рамки настоящей статьи. В качестве примера лингвистического исследования, где проводится семантический анализ шести русских и одного французского показателя ЛСО, можно привести монографию [11]. С теоретических позиций, которых придерживаются авторы этой монографии, все рассматриваемые ими показатели оказываются многозначными.

показатель отношения противопоставления (*Contrast*), с использованием шаблона «[Начало_Предложения... Конец_Предложения] [Начало_Предложения *But*... Конец_Предложения]» они извлекали фрагменты с этим ЛСО. Аналогичным образом материал отбирался для всех рассматриваемых отношений.

Для выявления пар лексических единиц, значимых для определения семантики ЛСО, был разработан набор наивных байесовских классификаторов (*Naive Bayes classifiers*), которые тестировались на каждой паре ЛСО (т. е. использовался метод бинарной классификации), а массивы включали равное число примеров (т. е. использовались одинаково распределенные данные). Для того чтобы полученные пары лексических единиц не включали сами показатели, с помощью которых формировался массив данных (*but* ‘но’, *although* ‘хотя’, *because* ‘потому что’, *thus* ‘следовательно’, *if* ‘если’, *if... then* ‘если... то’, *for example* ‘например’ и др.), все эти показатели были удалены авторами из пар фрагментов, входящих в обучающую выборку. Особенно важно было сделать это в случае двухкомпонентных показателей, таких как *if... then*.

Именно последний пункт — удаление из фрагментов с ЛСО тех показателей, которые использовались для извлечения этих фрагментов из корпуса, — имеет ключевое значение для понимания того, как развивались методы исследования имплицитных ЛСО в последующие годы.

3 Современное состояние

Прежде всего необходимо подчеркнуть, что, принимая во внимание цель работы [9] — поиск пар лексических единиц, которые могут использоваться для определения семантики ЛСО, если оно не имеет однозначного показателя, удаление показателей ЛСО из обрабатываемого массива фрагментов с этими ЛСО было вполне оправданно. Однако впоследствии в целом ряде работ этот метод стал использоваться для формирования массивов фрагментов с имплицитными ЛСО (см., например, [12–15]), так как, по мнению авторов этих работ, он позволял решать проблему недостаточного объема массивов фрагментов с имплицитными ЛСО, размеченных вручную. Эти массивы данных, названных «синтетическими» [16, с. 431], как правило, использовались в сочетании с массивами «естественных» данных — найденных и размеченных вручную (прежде всего это были данные Пенсильванского дискурсивно аннотированного корпуса, *Penn Discourse Treebank* [17]). Тем не менее сама идея отбирать фрагменты с «имплицитными» ЛСО из монопольных текстов именно таким способом в итоге стала одной из доминирующих в исследованиях имплицитных ЛСО.

Примечательно, что идея использования синтетических данных в качестве обучающих при разработке средств, позволяющих определить семантику ЛСО, была детально рассмотрена в опубликованной еще в 2008 г. работе [18], написанной К. Спорледер и А. Ласкаридес, специалистами по информатике из

Тилбургского и Эдинбургского университетов. Эти исследователи пришли к выводу, что использование синтетических данных — непродуктивная стратегия, причем как минимум по двум причинам.

1. Модели, обученные на таких данных, плохо переносятся на естественные данные, так как примеры, где ЛСО было эксплицитным, но показатель впоследствии был удален, с лингвистической точки зрения сильно отличаются от примеров, где показателя изначально не было¹.
2. Часть генерируемых примеров недопустимо использовать еще на этапе создания модели, так как в ряде случаев после удаления показателя семантика ЛСО меняется (тогда как, согласно методике подготовки данных, метка, указывающая на семантику ЛСО, полностью обусловлена тем, какой показатель был использован во фрагменте).

Все это негативно сказывается на результатах, получаемых при проведении экспериментов с использованием синтетических фрагментов с имплицитными ЛСО. Более того, К. Спорледер и А. Ласкаридес пришли к выводу, что используя значительно меньший объем размеченных вручную данных, можно получить куда более высокие результаты, чем при использовании большого массива данных, сгенерированных автоматически: «возможно, лучше вложиться в ручную разметку относительно небольшого набора данных и использовать этот набор для обучения классификатора. . .» [18, с. 407].

Действительно, в целом ряде работ синтетические данные при исследовании имплицитных ЛСО не привлекались: эксперименты проводились исключительно на данных, размеченных вручную, прежде всего на данных, полученных из корпуса PDTB [19–22].

Таким образом, можно констатировать, что в рассмотренных исследованиях использовались два метода отбора фрагментов с имплицитными ЛСО:

- (1) генерация синтетических примеров (определение «однозначных» показателей ЛСО; поиск фрагментов с такими ЛСО на основе этих показателей; удаление из найденных фрагментов всех показателей, использовавшихся для их поиска; проставление меток, указывающих на семантику ЛСО, основанное на том, к какому ЛСО был приписан показатель, использовавшийся для поиска);
- (2) выгрузка примеров из корпусов, размеченных вручную.

Ни один из них не может считаться действительно методом поиска, поскольку новые фрагменты с имплицитными ЛСО из текстов не извлекаются.

¹Именно проблема переноса модели, обученной на одном типе данных, на данные другого типа оказалась в центре внимания в работах [13–15], где синтетические и естественные данные использовались при исследовании имплицитных ЛСО во французском языке.

4 Заключение

Вывод, который можно сделать по итогам проведенного в статье анализа: в исследованиях, где имплицитные ЛСО рассматриваются на мооязычном материале, ставится прежде всего цель обработать массив фрагментов с имплицитными ЛСО (как правило, методами машинного обучения), чтобы на выходе получить массив, где каждому фрагменту из этого массива будет приписана метка, содержащая указание на то, какова семантика ЛСО в данном фрагменте. Задача целенаправленного поиска фрагментов с имплицитными ЛСО в мооязычных текстах на сегодняшний день не решена.

Закономерным образом возникает вопрос о том, какой еще материал может быть использован для целенаправленного поиска фрагментов с имплицитными ЛСО. Таким материалом оказались параллельные тексты. Методы поиска имплицитных ЛСО в параллельных текстах планируется проанализировать в рамках отдельной статьи.

Литература

1. 1st Workshop on Understanding Implicit and Underspecified Language Proceedings / Eds. M. Roth, R. Tsarfaty, Y. Goldberg. — Association for Computational Linguistics, 2021. 83 p.
2. 2nd Workshop on Understanding Implicit and Underspecified Language Proceedings / Eds. V. Pyatkin, D. Fried, T. Anthonio. — Seattle, WA, USA: Association for Computational Linguistics, 2022. 41 p.
3. Гончаров А. А., Инькова О. Ю. Имплицитные логико-семантические отношения и метод их поиска в параллельных текстах // Компьютерная лингвистика и интеллектуальные технологии: По мат-лам ежегодной Междунар. конф. «Диалог». — М.: РГГУ, 2020. Вып. 19(26). С. 310–320.
4. Инькова О. Ю. Логико-семантические отношения: проблемы классификации // Связность текста: мерологические логико-семантические отношения. — М.: ЯСК, 2019. С. 11–98.
5. Гончаров А. А. Классификации внутритекстовых отношений: основания и принципы структурирования // Вопросы языкознания, 2021. № 3. С. 97–119.
6. Языкознание: Большой энциклопедический словарь / Под ред. В. Н. Ярцевой. — 2-е изд. — М.: Большая Российская энциклопедия, 1998. 685 с.
7. Taboada M. Discourse markers as signals (or not) of rhetorical relations // J. Pragmatics, 2006. Vol. 38. No. 4. P. 567–592.
8. Webber B., Prasad R., Lee A., Joshi A. The Penn Discourse Treebank 3.0 annotation manual. 2019. <https://catalog.ldc.upenn.edu/docs/LDC2019T05/PDTB3-Annotation-Manual.pdf>.
9. Marcu D., Eshihabi A. An unsupervised approach to recognizing discourse relations // 40th Annual Meeting of the Association for Computational Linguistics Proceedings. — Philadelphia, PA, USA: Association for Computational Linguistics, 2002. P. 368–375.

10. *Pitler E., Raghupathy M., Mehta H., Nenkova A., Lee A., Joshi A.* Easily identifiable discourse relations // 22nd Conference (International) on Computational Linguistics Proceedings. — Manchester, U.K., 2008. P. 87–90.
11. Семантика коннекторов: контрастивное исследование / Под ред. О. Ю. Иньковой. — М.: ТОРУС ПРЕСС, 2018. 368 с.
12. *Pitler E., Louis A., Nenkova A.* Automatic sense prediction for implicit discourse relations in text // Joint Conference of the 47th Annual Meeting of the ACL and the 4th Joint Conference (International) on Natural Language Processing of the AFNLP Proceedings. — Singapore, Suntec: Association for Computational Linguistics, 2009. P. 683–691.
13. *Braud Chl., Denis P.* Identification automatique des relations discursives “implicites” à partir de données annotées et de corpus bruts // Actes de la 20e conférence sur le Traitement Automatique des Langues Naturelles. — Les Sables d’Olonne, France: ATALA, 2013. Vol. 1. P. 104–117.
14. *Braud Chl., Denis P.* Identifier les relations discursives implicites en combinant données naturelles et données artificielles // Traitement Automatique Langues, 2014. Vol. 55. No. 1. P. 135–165.
15. *Braud Chl.* Identification automatique des relations discursives implicites à partir de corpus annotés et de données brutes. — Paris: Université Paris Diderot, 2015. 221 p.
16. *Blair-Goldensohn S., McKeown K. R., Rambow O. C.* Building and refining rhetorical-semantic relation models // Human Language Technologies Conference Proceedings. — Rochester, New York: Association for Computational Linguistics, 2007. P. 428–435.
17. Penn Discourse Treebank Project. <https://www.seas.upenn.edu/~pdtb>.
18. *Sporleder C., Lascarides A.* Using automatically labelled examples to classify rhetorical relations: An assessment // Nat. Lang. Eng., 2008. Vol. 14. No. 3. P. 369–416.
19. *Lin Z., Kan M.-Y., Ng H. T.* Recognizing implicit discourse relations in the Penn Discourse Treebank // Conference on Empirical Methods in Natural Language Processing Proceedings. — Singapore: Association for Computational Linguistics, 2009. P. 343–351.
20. *Park J., Cardie C.* Improving implicit discourse relation recognition through feature set optimization // 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue Proceedings. — Seoul, South Korea: Association for Computational Linguistics, 2012. P. 108–112.
21. *Rutherford A. T., Xue N.* Discovering implicit discourse relations through brown cluster pair representation and coreference patterns // 14th Conference of the European Chapter of the Association for Computational Linguistics Proceedings. — Gothenburg, Sweden: Association for Computational Linguistics, 2014. P. 645–654.
22. *Liang L., Zhao Zh., Webber B.* Extending implicit discourse relation recognition to the PDTB-3 // 1st Workshop on Computational Approaches to Discourse Proceedings. — Association for Computational Linguistics, 2020. P. 135–147.

Поступила в редакцию 15.08.22

METHODS FOR RETRIEVAL OF IMPLICIT LOGICAL-SEMANTIC RELATIONS FROM MONOLINGUAL TEXTS

A. A. Goncharov

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

Abstract: The paper focuses on the problem of targeted retrieval of fragments with implicit logical-semantic relations (LSRs) from monolingual natural language texts. In order to answer the question whether such retrieval is possible, a survey of works with the results of implicit LSRs studies has been conducted. The aim of the survey is to identify the methods used in these works to create arrays of text fragments with implicit LSRs. It was revealed that mainly two methods are used: (i) automatic generation of fragments with implicit LSRs using fragments with explicit LSRs; and (ii) exporting of fragments with implicit LSRs from manually annotated text corpora. Neither of these techniques can be considered as a retrieval method because they do not imply detection of fragments with implicit LSRs in the natural language text and their subsequent extraction. Therefore, it can be stated that for monolingual texts, the task of targeted retrieval of such fragments currently remains unsolved.

Keywords: implicitness; implicit logical-semantic relations; targeted retrieval; natural language processing; information extraction

DOI: 10.14357/08696527220309

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).

References

1. Roth, M., R. Tsarfaty, and Y. Goldberg, eds. 2021. *1st Workshop on Understanding Implicit and Underspecified Language Proceedings*. Association for Computational Linguistics. 83 p.
2. Pyatkin, V., D. Fried, and T. Anthonio, eds. 2022. *2nd Workshop on Understanding Implicit and Underspecified Language Proceedings*. Seattle, WA: Association for Computational Linguistics. 41 p.
3. Goncharov, A. A., and O. Yu. Inkova. 2020. Implitsitnye logiko-semanticheskie ot-nosheniya i metod ikh poiska v parallel'nykh tekstakh [Implicit logical-semantic relations and a method of their identification in parallel texts]. *Computational Linguistics and Intellectual Technologies: Papers from the Annual Conference (International) “Dialogue”*. Moscow: RSHI. 19(26):310–320.

4. Inkova, O. Yu. 2019. Logiko-semanticheskie otnosheniya: Problemy klassifikatsii [Logical-semantic relations: Classification problems]. *Svyaznost' teksta: mereologicheskie logiko-semanticheskie otnosheniya* [Text coherence: Mereological logical semantic relations]. Moscow: LRC Publishing House. 11–98.
5. Goncharov, A. A. 2021. Klassifikatsii vnutritekstovyykh otnosheniy: Osnovaniya i printsipy strukturirovaniya [Classifications of intratextual relations: Bases and structuring principles]. *Voprosy yazykoznavaniya* [Topics in the Study of Language] 3:97–119.
6. Yartseva, V. N., ed. 1998. *Yazykoznanie: Bol'shoy entsiklopedicheskiy slovar'* [Linguistics. Great encyclopedic dictionary]. 2nd ed. Moscow: Bol'shaya Rossiyskaya entsiklopediya. 685 p.
7. Taboada, M. 2006. Discourse markers as signals (or not) of rhetorical relations. *J. Pragmatics* 38(4):567–592.
8. Webber, B., R. Prasad, A. Lee, and A. Joshi. 2019. The Penn Discourse Treebank 3.0 annotation manual. Available at: <https://catalog.ldc.upenn.edu/docs/LDC2019T05/PDTB3-Annotation-Manual.pdf> (accessed August 24, 2022).
9. Marcu, D., and A. Echihabi. 2002. An unsupervised approach to recognizing discourse relations. *40th Annual Meeting of the Association for Computational Linguistics Proceedings*. Philadelphia, PA: Association for Computational Linguistics. 368–375.
10. Pitler, E., M. Raghupathy, H. Mehta, A. Nenkova, A. Lee, and A. Joshi. 2008. Easily identifiable discourse relations. *22nd Conference (International) on Computational Linguistics Proceedings*. Manchester, U.K. 87–90.
11. In'kova, O. Yu., ed. 2018. *Semantika konnektorov: kontrastivnoe issledovanie* [Semantics of connectives: Contrastive study]. Moscow: TORUS PRESS. 368 p.
12. Pitler, E., A. Louis, and A. Nenkova. 2009. Automatic sense prediction for implicit discourse relations in text. *Joint Conference of the 47th Annual Meeting of the ACL and the 4th Joint Conference (International) on Natural Language Processing of the AFNLP Proceedings*. Singapore, Suntec: Association for Computational Linguistics. 683–691.
13. Braud, Chl., and P. Denis. 2013. Identification automatique des relations discursives “implicites” à partir de données annotées et de corpus bruts. *Actes de la 20e conférence sur le Traitement Automatique des Langues Naturelles*. Les Sables d'Olonne, France: ATALA. 1:104–117.
14. Braud, Chl., and P. Denis. 2014. Identifier les relations discursives implicites en combinant données naturelles et données artificielles. *Traitement Automatique Langues* 55(1):135–165.
15. Braud, Chl. 2015. *Identification automatique des relations discursives implicites à partir de corpus annotés et de données brutes*. Paris: Université Paris Diderot. 221 p.
16. Blair-Goldensohn, S., K. R. McKeown, and O. C. Rambow. 2007. Building and refining rhetorical-semantic relation models. *Human Language Technologies Conference Proceedings*. Rochester, New York: Association for Computational Linguistics. 428–435.
17. Penn Discourse Treebank Project. Available at: <https://www.seas.upenn.edu/~pdtb> (accessed August 24, 2022).
18. Sporleder, C., and A. Lascarides. 2008. Using automatically labelled examples to classify rhetorical relations: An assessment. *Nat. Lang. Eng.* 14(3):369–416.

19. Lin, Z, M.-Y. Kan, and H. T. Ng. 2009. Recognizing implicit discourse relations in the Penn Discourse Treebank. *Conference on Empirical Methods in Natural Language Processing Proceedings*. Singapore: Association for Computational Linguistics. 343–351.
20. Park, J., and C. Cardie. 2012. Improving implicit discourse relation recognition through feature set optimization. *13th Annual Meeting of the Special Interest Group on Discourse and Dialogue Proceedings*. Seoul, South Korea : Association for Computational Linguistics. 108–112.
21. Rutherford, A. T., and N. Xue. 2014. Discovering implicit discourse relations through brown cluster pair representation and coreference patterns. *14th Conference of the European Chapter of the Association for Computational Linguistics Proceedings*. Gothenburg, Sweden: Association for Computational Linguistics. 645–654.
22. Liang, L., Zh. Zhao, and B. Webber. 2020. Extending implicit discourse relation recognition to the PDTB-3. *1st Workshop on Computational Approaches to Discourse Proceedings*. Association for Computational Linguistics. 135–147.

Received August 15, 2022

Contributor

Goncharov Alexander A. (b. 1994) — 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; a.gonch48@gmail.com

ИНТЕРФЕЙС ПОЛЬЗОВАТЕЛЯ НАДКОРПУСНОЙ БАЗЫ ДАННЫХ ИЕРАРХИЙ ЛОГИКО-СЕМАНТИЧЕСКИХ ОТНОШЕНИЙ*

А. А. Дурново¹, О. Ю. Инькова², Н. А. Попкова³

Аннотация: Описываются пользовательские возможности Надкорпусной базы данных иерархии (НБДИ) логико-семантических отношений (ЛСО), нового инструмента для аннотации дискурсивных отношений, разработанного в ФИЦ ИУ РАН. Перечислены этапы построения аннотаций в НБДИ ЛСО: выбор контекста для аннотирования, его возможные модификации, аннотирование показателя ЛСО и самого ЛСО. Сформированные аннотации не только отражают в единой аннотации три важных теоретических аспекта дискурсивного анализа: показатель ЛСО (или его отсутствие), иерархию ЛСО и границы фрагментов текста, непосредственно связанных ЛСО. Интерфейс НБДИ ЛСО позволяет редактировать, дополнять и фильтровать аннотированную информацию: сохранять предыдущие состояния аннотаций и при необходимости возвращаться к одному из них, восстанавливать удаленные аннотации, использовать созданные аннотации для создания новых. Полученные с помощью НБДИ ЛСО результаты уже позволили уточнить некоторые теоретические положения, в частности сформулированные в наиболее популярной теории риторической структуры. Дальнейший анализ материала и статистические возможности НБДИ ЛСО позволят установить закономерности в глубине иерархии ЛСО. Эта информация представляет интерес для когнитивных исследований — как моноязычных, так и контрастивных, а также для автоматической обработки текста и разработки искусственного интеллекта.

Ключевые слова: надкорпусная база данных; аннотирование корпусов текстов; дискурсивные отношения; коннектор

DOI: 10.14357/08696527220310

*Работа выполнялась с использованием инфраструктуры Центра коллективного пользования «Высокопроизводительные вычисления и большие данные» (ЦКП «Информатика») ФИЦ ИУ РАН (г. Москва) в рамках швейцарско-российского проекта «Методология аннотирования в надкорпусной базе данных коннекторов» по гранту Швейцарского государственного секретариата по образованию, исследованиям и инновациям.

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, duralex49@mail.ru

²Федеральный исследовательский центр «Информатика и управление» Российской академии наук; Женевский университет, olyainkova@yandex.ru

³Федеральный исследовательский центр «Информатика и управление» Российской академии наук, natasha_popkova@mail.ru

1 Вводные замечания

Перед создателями лингвистических ресурсов, позволяющих исследовать дискурсивные, или, в рассматриваемом подходе, логико-семантические, отношения, играющие важную роль в обеспечении связности текста [1], стоят следующие основные задачи.

1. Разметка в единой аннотации как самих ЛСО, так и их показателей с учетом формального и функционального разнообразия последних [2]¹.
2. Разметка ЛСО, не имеющих показателя, так называемых имплицитных ЛСО [3].
3. Отображение границ фрагментов текста, непосредственно связанных ЛСО; см. пример:
[(Идеально сделано)*p*, *клянусь богом*, — говорил Лариосик. (Как не идеально!)]*q*]*t* [Вещь (под руками)]*г* и в то же время (вне квартиры)]*s*]*u* [Михаил Булгаков. Белая гвардия (1924)].
4. Отображение иерархии ЛСО, т. е. тех случаев, когда одно ЛСО как бы вкладывается в другое, нарушая линейность интерпретации текста; см. приведенный пример.
5. Хранение и визуализация аннотаций в удобном для пользователя формате с возможностью поиска по заданным исследователями параметрам и получения количественных данных по каждому из них.

В приведенном примере ЛСО иллокутивной причины, с помощью которого говорящий обосновывает истинность сказанного им [1], связывает без использования какого-либо показателя фрагменты текста *t* и *u* (в квадратных скобках). Фрагмент *t* содержит, в свою очередь, фрагменты *p* и *q*, связанные ЛСО переформулирования (говорящий говорит одно и то же, но разными словами). Границы *p* и *q* обозначены при помощи круглых скобок, а текст, который непосредственно не задействован в данном ЛСО, выделен серым цветом. Во фрагменте *u* также есть ЛСО (временное), выраженное показателем *и в то же время* и соединяющее предикаты (*быть*) *под руками* и (*быть*) *вне квартиры* с общим подлежащим *вещь*. По отношению к ЛСО иллокутивной причины ЛСО переформулирования и времени являются встроеными, поскольку они находятся во фрагментах текста,

¹ Не вдаваясь в детали, можно сказать, что ЛСО устанавливаются в тексте между дискурсивными единицами, имеющими, как правило, пропозициональное содержание и следующими друг за другом. Например, *Петя заболел, поэтому он не пришел*. С формальной точки зрения, показатели ЛСО могут быть однословными (поэтому, следовательно) и неоднословными (*так как, несмотря на то что, хотя... но*). С функциональной точки зрения, выделяются показатели, устанавливающие ЛСО на пропозициональном уровне, на уровне высказывания (например, при обосновании сказанного, как в *Замолчи, или я тебя накажу*) или на уровне организации текста (*во-первых, во-вторых* и др.).

связанных этим ЛСО. Иерархия семантической структуры приведенного примера, границы соединяемых ЛСО фрагментов, а также наличие или отсутствие его показателя должны быть в идеале отражены в аннотации.

Существующие ресурсы решают эти проблемы лишь отчасти (см. Ru-RSTreebank [4], Пенсильванский дискурсивно аннотированный корпус [5]) или предоставляют ограниченные возможности для дальнейшего использования аннотаций (см. RST Signaling Corpus [6] или GUM [7]); подробнее об аннотировании текстовых отношений с помощью этих ресурсов см. [8–10].

Надкорпусная база данных иерархий ЛСО создана в Институте проблем информатики ИПИ РАН с целью расширения возможностей аннотирования ЛСО и их показателей, предоставляемых Надкорпусной базой данных коннекторов (НБДК), также созданной в ИПИ РАН [11]. Как и последняя, НБДИ ЛСО использует тексты параллельных французских и итальянских подкорпусов Национального корпуса русского языка [12] общим объемом свыше 11 млн словоупотреблений, а также аннотации НБДК. В работе [8] описана архитектура НБДИ ЛСО, в этом исследовании показано, как в ней решаются сформулированные выше задачи.

2 Аннотирование: построение графа

2.1 Выбор текста для аннотирования и работа с ним

Логико-семантическими отношениями считаются такие дискурсивные отношения, которые могут быть потенциально выражены коннекторами, выступающими, следовательно, их показателями [3, с. 311]. К коннекторам относятся сочинительные и подчинительные союзы (*или, хотя, потому что* и др.), наречия, называемые в русской грамматической традиции «аналогами союзов» (*например, однако, между тем, впрочем* и др.), и некоторые предлоги (*несмотря на, после* и др.) [13]. К показателям ЛСО следует добавить два семантически значимых знака препинания: двоеточие и тире, совместимые, в отличие от других знаков препинания, с ограниченным набором ЛСО (в основном причины и спецификации). Исходя из этой теоретической предпосылки, предметом аннотирования в НБДИ ЛСО прежде всего следует считать показатель ЛСО в тексте оригинала или в тексте перевода¹.

¹В тексте перевода коннектор может иметь нулевой эквивалент; ср.: Это благородно, *то есть* я хотел сказать, гуманно! [Ф. М. Достоевский. Преступление и наказание (1866)] — *C'est noble, je voulais dire humain* [Tr. par É. Guertik]. Русскому коннектору *то есть* в переводе не соответствует никакой показатель. Возможна и обратная ситуация: коннектор в переводе имеет нулевой стимул: Я ему советовала: «Сашка, найди хорошую женщину. Сопьешься» [С. Алексиевич. Время секунд хэнд (2013)] — *Moi, Sacha, je lui donnais des conseils: «Trouve-toi une gentille femme, sinon tu vas te mettre à boire»* [Tr. par S. Benech]. Имплицитному ЛСО отрицательной альтернативы в оригинале соответствует показатель этого ЛСО *sinon* в переводе. Такие случаи НБДИ ЛСО позволяет аннотировать. На этом вопросе не будем останавливаться подробно; см. [3].

Список построенных деревьев

Фильтровать	Создать новое дерево	Список удаленных деревьев	Выход		
	№	Id пары	Автор	Название	Текст
Показать	1	224710	С. Д. Довлатов	Заповедник	Не то чтобы меня выбросили из ресторана. Я выполз сам, окутанный драпировочной тканью.
Показать	2	492230	А. А. и Г. А. Вайнеры	Эра милосердия	Слушай, Глеб, мне как-то Пасюк говорил, что если к нам, например, позвонят, скажут чего-нибудь, а потом бросят трубку, а ты хочешь узнать, откуда звонили, то это можно. Так это?
Показать	3	487417	А. А. и Г. А. Вайнеры	Эра милосердия	А главный человек в розыске — свидетель, потому что в самом тайном делишке всегда отыщется человечек, который или что-то видел, или слышал, или знает, или помнит, или догадывается.

Рис. 1 Фрагмент страницы главного меню НБДИ ЛСО

Показать	28	49483	Ф. М. Достоевский	Преступление и наказание (ч. 1)	Силы его тут как бы не было. Но как только он раз опустил топор, тут и родилась в нем сила.
Показать	29	36972	Н. В. Гоголь	Шинель	Ни один раз в жизни не обратил он внимания на то, что делается и происходит всякий день на улице, на что, как известно, всегда посмотрит его же брат, молодой чиновник, простирающий до того пронзительность своего бойкого взгляда, что заметит даже, у кого на другой стороне тротуара отпоролась внизу панталон стремешка, — что вызывает всегда лукавую усмешку на лице его. Но Акакий Акакиевич если и глядел на что, то видел на всем свои чистые, ровным почерком выписанные строки, и только разве если, неизвестно откуда взявшись, лошадиная морда помешалась ему на плечо и напускала ноздрями целый ветер в щеку, тогда только замечал он, что он не на середине строки, а скорее на середине улицы.

Рис. 2 Примеры деревьев с расширенным контекстом

Единицей хранения информации в НБДИ ЛСО служит аннотация, имеющая форму дерева, т. е. ориентированного связного ациклического графа, позволяющего отразить иерархическую структуру текста. Создание нового дерева начинается на странице главного меню НБДИ ЛСО нажатием кнопки «Создать новое дерево» (рис. 1).

Существуют две возможности: использовать аннотацию, уже созданную в НБДК для выбранного коннектора, указав ее номер, или аннотировать «с нуля», выбрав контекст употребления коннектора из корпуса текстов по идентификационному номеру. После нажатия на кнопку «Искать» в окне «Выбранный текст» появляется соответствующий фрагмент.

На следующем этапе аннотирования лингвист определяет границы фрагментов текста, связанных коннектором. Для этого к тексту в окне «Выбранный текст» можно добавить текст из левого или правого контекста коннектора. Так, на рис. 2

Окружение:

Аннотации:

№	Текст
3600	Не то чтобы меня выбросили из ресторана. Я выполз сам, окутанный драпировочной тканью.
8284	Не то чтобы меня выбросили из ресторана. Я выполз сам, окутанный драпировочной тканью.

Деревья:

	№	Id пары	Автор	Название	Текст
<input type="button" value="Выбор"/>	1	224710	С. Д. Довлатов	Заповедник	Не то чтобы меня выбросили из ресторана. Я выполз сам, окутанный драпировочной тканью.

Рис. 3 Информация, получаемая для контекста № 1 на рис. 1 нажатием кнопки «Окружение»

в дереве 29 для корректной разметки *но* добавлен текст, предшествующий *Но Акакий Акакиевич...*

С помощью кнопки «Окружение» можно вызвать информацию об уже созданных в НБДК аннотаций с данным коннектором и созданными для них графами в НБДИ ЛСО с гиперссылками на каждый из них (рис. 3).

2.2 Аннотирование коннектора и создание узлов графа

После того как текст выбран и определены левый и правый контексты коннектора, лингвист приступает непосредственно к созданию графа (рис. 4).

Нажав на текст дерева, разметчик может проводить следующие операции: «Отменить», «Редактировать», «Удалить», «Разделить на два» и «Разделить на несколько», «Вставить промежуточный», «Разметить», «Отдельное дерево» (рис. 5).

Кнопки «Разделить на два», «Разделить на несколько» и «Разметить» служат для создания структуры дерева, а их использование зависит от того, к какому классу принадлежит коннектор [2]: однокомпонентный, т. е. находящийся в одном из фрагментов текста, двух- или многокомпонентный, т. е. находящийся в каж-

Дерево контекстов

№ 74

Для операции с узлом необходимо кликнуть мышью на тексте внутри этого узла

Поэтому я спросил: — Слушай, Глеб, мне как-то Пасюк говорил, что если к нам, например, позвонят, скажут чего-нибудь, а потом бросят трубку, а ты хочешь узнать, откуда звонили, то это можно. Так это?

Окружение:

Рис. 4 Начальный этап построения графа-дерева

Операции с узлом дерева

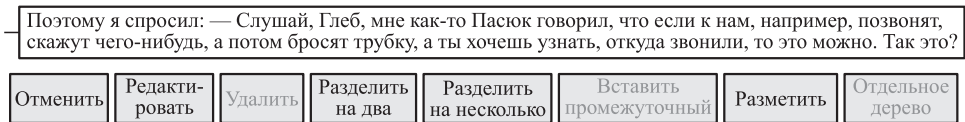


Рис. 5 Операции с узлом дерева

дом из соединяемых фрагментов текста, как, например, *если... то* в № 2 на рис. 1. Для него будет использована кнопка «Разделить на два».

Дополнительно существуют следующие функции:

- кнопка «Отменить» возвращает обратно в окно «Дерево контекстов»;
- кнопка «Редактировать» позволяет редактировать текст узла, удалять контекст, непосредственно не задействованный в установлении ЛСО;
- кнопка «Удалить» вызывает удаление соответствующего узла (и всех нижестоящих) и возврат к режиму просмотра дерева;
- «Вставить промежуточный» служит для вставки промежуточного иерархического уровня;
- «Отдельное дерево» позволяет создать отдельное дерево, где текущий узел станет главным.

После того как определено, на какое число фрагментов должен быть разбит текст, лингвист размечает сам коннектор. Для этого ему нужно определить, есть ли другие коннекторы в аннотируемом тексте, и, в этом случае, является ли аннотируемый коннектор встраивающим или встроенным¹. В № 2 на рис. 1 коннектор *если... то* — встраивающий, поскольку в первом из соединяемых им фрагментов текста находятся другие коннекторы: *а потом* и *а*, которые, соответственно, являются по отношению к нему встроенными, причем *а потом* оказывается, в свою очередь, встраивающим по отношению к *а*. Это будет отражено в структуре дерева, а именно в иерархии его узлов.

Для аннотирования коннектора и выражаемого им ЛСО лингвист использует кнопку «Выбор коннектора и отношения», которая вызывает соответствующий список. Для выбора коннектора в окне «Найти существующий тип» следует ввести коннектор или его часть и нажать кнопку «Искать». Сформируется список в алфавитном порядке из 30 показателей, которые включают введенное значение. После выбора коннектора в поле «Результат» автоматически добавляется то

¹Встраивающим называется коннектор, в сфере действия которого находится еще один коннектор, называемый встроенным. Ср.: Между нами уже было не так, как раньше, *то есть* мы дружили, *но* все больше я на нее покрикивала [С. Алексиевич. *Время секонд хэнд* (2013)]. В сфере действия коннектора *то есть* (в его правом контексте, включающем фрагмент «мы дружили, *но* все больше я на нее покрикивала») находится коннектор *но*. *То есть* является здесь встраивающим коннектором, *но* — встроенным. См. также приведенный пример.

Исходный текст:

Поэтому я спросил: — Слушай, Глеб, мне как-то Пасюк говорил, что если к нам, например, позвонят, скажут чего-нибудь, а потом бросят трубку, а ты хочешь узнать, откуда звонили, то это можно. Так это?

Выбрать тип коннектора и отношения

если||то условные

Разбить автоматически

Левый контекст: пустойПравый контекст: пустой

если к нам, например, позвонят, скажут
чего-нибудь, а потом бросят трубку,
а ты хочешь узнать, откуда звонили,

то это можно.

Разбить вручную

Сохранить

Отменить

Рис. 6 Разбивка фрагмента и удаление текста

ЛСО, которое наиболее часто выражается данным коннектором в НБДК. Таким образом, данные о коннекторах и о частотности выражаемого им ЛСО в одной надкорпусной базе данных используются для автоматизации одной из операций в другой. При отсутствии необходимого коннектора или выражаемого им ЛСО их можно добавить в соответствующие списки.

После аннотирования коннектора и ЛСО текст разбивается на фрагменты автоматически или вручную с помощью соответствующих кнопок. На этом этапе лингвист может удалить текст, который непосредственно не участвует в установлении ЛСО; см. приведенный ранее пример и рис. 6 для пары № 2 с рис. 1, где при построении графа для коннектора *если... то* предшествующий *если* текст *Слушай, Глеб, мне как-то Пасюк говорил, что будет убран*¹.

Если один из контекстов «пустой» (содержит только коннектор и, например, многоточие), в нужном окне есть возможность нажать галочку «пустой». Эта функция недоступна в существующих ресурсах.

После сохранения результатов аннотирования на странице «Дерево контекстов» можно увидеть корневой узел и узлы первого уровня (рис. 7). Аннотирование встроенных коннекторов приводит к созданию узлов второго и третьего уровня (рис. 8).

¹Насколько известно авторам, ни один из существующих ресурсов дискурсивного аннотирования не позволяет видоизменять фрагменты текста, связанные ЛСО. Единственный ресурс, а именно РДТВ, точно указывает границы фрагментов текста при помощи цветовой разметки. Подробнее см. [9].

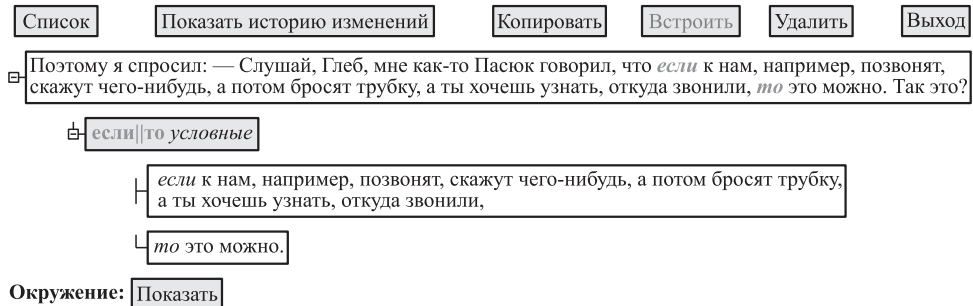


Рис. 7 Создание корневого узла

Для операции с узлом необходимо кликнуть мышью на тексте внутри этого узла

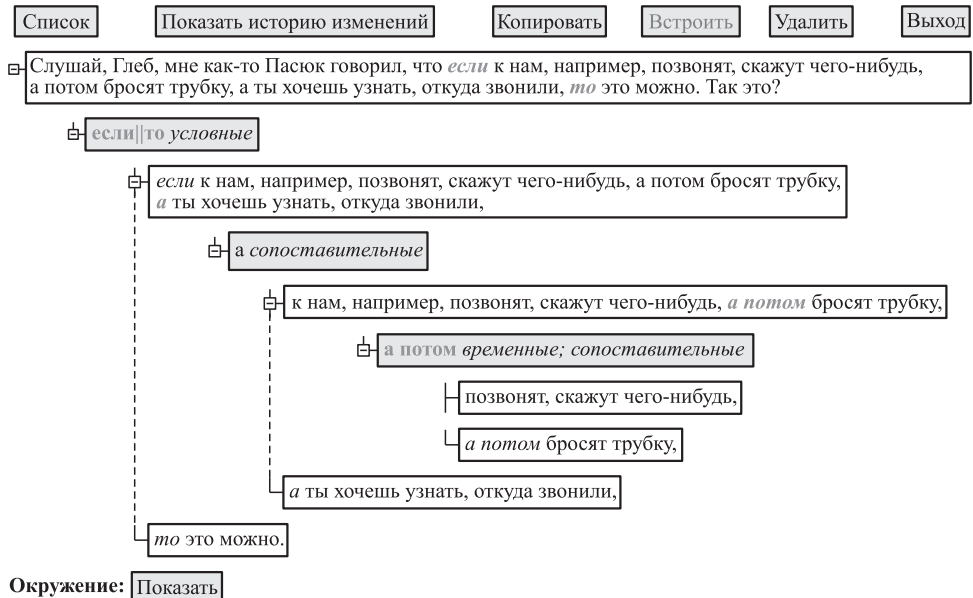


Рис. 8 Аннотирование иерархии ЛСО

Редактировать узел можно на любом этапе с помощью кнопки «Редактировать», но нельзя удалить текст, который используется в узле на уровне ниже. Удаление узла приводит к удалению всех узлов на уровнях ниже.

3 Аннотирование: переключение между состояниями, функция «копировать» и «встроить»

После создания дерева интерфейс НБДИ ЛСО позволяет решить несколько задач по редактированию и дополнению аннотированной информации.

3.1 Необходимо вернуться к предыдущему состоянию дерева или переключиться на следующее

Каждое состояние дерева привязано ко времени внесенного изменения и к имени пользователя, который его внес. Можно выбрать нужное состояние из списка гиперссылок, вызываемого с помощью кнопки «История дерева». Выбрав одно из состояний, можно «Вернуться к текущему состоянию», «Вернуться к истории», «Показать предыдущее состояние», «Сравнить с предыдущим», «Показать последующее», а также «Сделать текущим» (текущее состояние дерева зафиксировано как постоянное, и с ним можно работать далее). Эти функции позволяют учитывать эволюцию разметки, связанную, в частности, с получением лингвистами новых знаний об ЛСО и их показателях (рис. 9).

Дерево №2, состояние 4 с 21-06-2021 18:39:17 по 21-06-2021 18:40:34 (Попкова)



Рис. 9 Одна из версий дерева с указанием времени и даты внесенных изменений

3.2 Разметка узла более высокого или более низкого иерархического уровня

Если сначала был размечен встроенный коннектор, то для встраиваемого коннектора используется функция «Копировать» — «Встроить», позволяющая избежать дублирования при аннотировании. После копирования уже созданного дерева необходимо перейти в окно работы над деревом для встраиваемого коннектора и нажать «Встроить». И наоборот: может быть создано «Отдельное дерево» для встроенного коннектора (такая функция может быть необходима для получения статистики). В этом случае узел, с которым работал разметчик, в новом дереве станет главным.

3.3 Восстановить удаленное дерево и найти существующее

На рис. 1 показано главное меню. При нажатии на кнопку «Список удаленных деревьев» открывается список деревьев, которые были построены, а затем

удалены, — любое из них может быть восстановлено. Кнопка «Фильтровать» используется для отбора деревьев по тексту, на основе которого они были сформированы. Фильтр работает двумя способами:

- (1) если ввести часть текста, заключенную в кавычки, например «*ничто не напоминало старику*», будут найдены деревья, содержащие фрагмент текста *ничто не напоминало старику*;
- (2) если ввести текст без кавычек, например *ничто не напоминало старику*, будут найдены деревья, каждое из которых содержит все указанные словоформы (*ничто, не, напоминало и старику*), причем в любом порядке.

4 Заключение

Интерфейс НБДИ ЛСО позволяет пользователю-лингвисту создавать графы для ЛСО и их показателей. Сформированные графы отражают в единой аннотации три важных теоретических аспекта дискурсивного анализа: показатель ЛСО (или его отсутствие), иерархию ЛСО и границы фрагментов текста, непосредственно связанных ЛСО. Кроме того, поскольку НБДИ ЛСО использует параллельные корпуса текстов, с помощью описанного интерфейса лингвист может создать двуязычную аннотацию. Полученные с помощью НБДИ ЛСО результаты уже позволили уточнить некоторые теоретические положения, в частности сформулированные в наиболее популярной теории риторической структуры [14]; подробнее см. [9, 10]. Дальнейший анализ материала и статистические возможности НБДИ ЛСО позволяют установить закономерности в глубине иерархии ЛСО: сколько максимально уровней может быть в семантической структуре текста при наличии и при отсутствии показателя ЛСО. Эта информация представляет интерес для когнитивных исследований — как моноязычных, так и контрастивных, а также для автоматической обработки текста и разработки искусственного интеллекта.

Литература

1. Инькова О. Ю. Логико-семантические отношения: проблемы классификации // Связность текста: Мереологические логико-семантические отношения / Под ред. О. Иньковой, Э. Манзотти. — М.: Языки славянских культур, 2019. С. 11–98.
2. Инькова О. Ю. Структура коннекторов: лингвистические методы описания // Структура коннекторов и методы ее описания / Под ред. О. Ю. Иньковой. — М.: ТОРУС ПРЕСС, 2019. С. 5–47.
3. Гончаров А. А., Инькова О. Ю. Имплицитные логико-семантические отношения и метод их поиска в параллельных текстах // Компьютерная лингвистика и интеллектуальные технологии: По мат-лам ежегодной Междунар. конф. «Диалог». — М.: РГГУ, 2021. Вып. 20. С. 310–320.
4. Ru-RSTreebank: Русскоязычный дискурсивный корпус. <https://rstreebank.ru>.
5. Penn Discourse Treebank Project. <https://www.seas.upenn.edu/~pdtb>.

6. *Das D., Taboada M.* RST signalling corpus Annotation manual. 2014. https://www.sfu.ca/~mtaboada/docs/publications/RST_Signalling_Corpus_Annotation_Manual.pdf.
7. GUM: The Georgetown University Multilayer Corpus. <https://corpling.uis.georgetown.edu/gum/annotations.html>.
8. *Дурново А. А., Инькова О. Ю., Попкова Н. А.* Архитектура базы данных иерархии логико-семантических отношений // Системы и средства информатики, 2022. Т. 32. № 1. С. 114–125.
9. *Дурново А. А., Инькова О. Ю., Попкова Н. А.* Принципы описания показателей логико-семантических отношений и их иерархии // Информатика и её применения, 2022. Т. 16. Вып. 2. С. 52–59.
10. *Inkova O. Yu., Nuriev V. A., Popkova N. A.* The role of paragraph in the corpora of annotated texts // Компьютерная лингвистика и интеллектуальные технологии: По мат-лам ежегодной Междунар. конф. «Диалог». — М.: РГГУ, 2022. Вып. 21. С. 273–283.
11. *Inkova O., Popkova N.* Statistical data as information source for linguistic analysis of Russian connectors // Информатика и её применения, 2017. Т. 11. Вып. 3. С. 123–131.
12. Национальный корпус русского языка (НКРЯ). <https://ruscorpora.ru/new>.
13. *Инькова-Манзотти О. Ю.* Коннекторы противопоставления во французском и русском языках: сопоставительное исследование. — М.: МГУ-Информэлектро, 2001. 434 с.
14. *Mann W., Thompson S.* Rhetorical structure theory: Towards a functional theory of text organization // Text, 1988. Vol. 8. Iss. 3. P. 243–281. doi: 10.1515/text.1.1988.8.3.243.

Поступила в редакцию 08.08.22

DATABASE OF HIERARCHIES OF LOGICAL-SEMANTIC RELATIONS: USER INTERFACE

A. A. Durnovo¹, O. Yu. Inkova^{1,2}, and N. A. Popkova¹

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

²University of Geneva, 22 Bd des Philosophes, CH-1205 Geneva 4, Switzerland

Abstract: The paper describes user features of the Database of Hierarchies of Logical-Semantic Relations (SBDH LSR), a new tool for annotating discourse relations, developed at FRC CSC RAS. The article lists annotations stages in the SBDH LSR: selection of the context for annotation, possible modifications, annotation of the LSR marker and of LSR itself. In a single annotation, the formed annotations reflect three important theoretical aspects of discourse analysis: the LSR marker (or its absence); the hierarchy of LSR; and the boundaries of text fragments directly connected by LSR. The database’s interface allows one to edit, complete, or filter the annotated information: to save annotations’ previous states and, if necessary, go back to one of them, to recover deleted

annotations, to use already created annotations to form the new ones. The results obtained with the SBDH LSR have already allowed us to clarify some theoretical points, in particular, those formulated by the Rhetorical Structure Theory. Further analysis of the examples and statistical capabilities of the SBDH LSR will make it possible to establish regularities in the deep LSR hierarchies. This information might be used in cognitive research, both monolingual and contrastive, as well as for automatic text processing and artificial intelligence development.

Keywords: supracorpora database; corpus of text annotations; discourse relations; connector

DOI: 10.14357/08696527220310

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) and supported by the Switzerland–Russia Research Preparation Grant “Annotation methodology in a supracorpora database of connectives” of the State Secretariat for Education, Research, and Innovation.

References

1. Inkova, O. Yu. 2019. Logiko-semanticheskie otnosheniya: Problemy klassifikatsii [Logical-semantic relations: Classification problems]. *Svyaznost' teksta: mereologicheskie logiko-semanticheskie otnosheniya* [Text coherence: Mereological logical semantic relations]. Eds. O. Inkova and E. Manzotti. Moscow: LRC Publishing House. 11–98.
2. Inkova, O. Yu. 2019. *Struktura konnektorov: lingvisticheskie metody opisaniya* [Connectors' structure and linguistic means of their description]. *Struktura konnektorov i metody ee opisaniya* [Connectives structure and methods of its description]. Ed. O. Yu. Inkova. Moscow: TORUS PRESS. 5–47.
3. Goncharov, A. A., and O. Yu. Inkova. 2020. Implitsitnye logiko-semanticheskie otnosheniya i metod ikh poiska v parallel'nykh tekstakh [Implicit logical-semantic relations and a method of their identification in parallel texts]. *Computational Linguistics and Intellectual Technologies: Papers from the Annual Conference (International) “Dialogue.”* Moscow: RSHI. 19(26):310–320.
4. Ru-RSTreebank: Russkoyazychnyy diskursivnyy korpus [Ru-RSTreebank: Russian discourse corpus]. Available at: <https://rstreebank.ru/> (accessed August 19, 2022).
5. Penn Discourse Treebank Project. Available at: <https://www.seas.upenn.edu/~pdtb/> (accessed August 19, 2022).
6. Das, D., and M. Taboada. 2014. RST signalling corpus: Annotation manual. Available at: https://www.sfu.ca/~mtaboada/docs/publications/RST_Signalling_Corpus_Annotation_Manual.pdf (accessed August 19, 2022).
7. GUM: The Georgetown university multilayer corpus. Available at: <https://corpling.uis.georgetown.edu/gum/annotations.html> (accessed August 19, 2022)

8. Durnovo, A. A., O. Yu. Inkova, and N. A. Popkova. 2022. Arkhitektura bazy dannykh iyerarkhii logiko-semanticheskikh otnosheniy [Database of hierarchies of logical-semantic relations: Architecture]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* (32)1:114–125.
9. Durnovo, A. A., O. Yu. Inkova, and N. A. Popkova. 2022. Printsipy opisaniya pokazateley logiko-semanticheskikh otnosheniy i ikh ierarkhii [Logical-semantic relations and their hierarchies: Description principles]. *Informatika i ee Primeneniya — Inform. Appl.* (16)2:52–59.
10. Inkova O. Yu., V. A. Nuriev, and N. A. Popkova. 2022. The role of paragraph in the corpora of annotated texts. *Computational Linguistics and Intellectual Technologies: Papers from the Annual Conference (International) “Dialogue.”* Moscow: RSHI. 21:273–283.
11. Inkova, O., and N. Popkova. 2017. Statistical data as information source for linguistic analysis of Russian connectors. *Informatika i ee Primeneniya — Inform. Appl.* 11(3):123–131.
12. Natsional’nyy korpus russkogo yazyka (NKRYa) [Russian National Corpus (RNC)]. Available at: <http://www.ruscorpora.ru/new> (accessed August 19, 2022)
13. Inkova-Manzotti, O. Yu. 2001. *Konnektory protivopostavleniya vo frantsuzskom i russkom yazykakh. Sopostavitel’noe issledovanie* [Connectives of opposition in French and Russian. A comparative study]. Moscow: MSU-Informelektro. 434 p.
14. Mann, W., and S. Thompson. 1988. Rhetorical structure theory: Towards a functional theory of text organization. *Text* 8(3):243–281. doi: 10.1515/text.1.1988.8.3.243.

Received August 8, 2022

Contributors

Durnovo Aleksandr A. (b. 1949) — leading programmer, 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; duralex49@mail.ru

Inkova Olga Yu. (b. 1965) — Doctor of Science in philology, senior 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; faculty member, University of Geneva, 22 Bd des Philosophes, CH-1205 Geneva 4, Switzerland; olyainkova@yandex.ru

Popkova Nataliia A. (b. 1992) — junior 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; natasha_popkova@mail.ru

СТРАТЕГИЯ ЕВРОПЕЙСКОГО КОМПЬЮТЕРНОГО ОБРАЗОВАНИЯ: ТЕОРЕТИЧЕСКИЕ ОСНОВАНИЯ ПОСТРОЕНИЯ*

И. М. Зацман¹

Аннотация: Анализируется стратегия Европейского компьютерного образования (ЕКО), при формировании которой использовались результаты двух-летнего мониторинга состояния систем преподавания информатики в регионах и странах Европы, включая РФ, а также в Израиле. Поскольку авторы стратегии на основе данных мониторинга анализировали состояние преподавания информатики, им потребовалось дать современное описание этой дисциплины. В стратегии используется подход к позиционированию информатики в системе современного научного познания, учитывающий отсутствие конвенционального определения ее предметной области. Цель статьи состоит в описании этого подхода, включающего две стадии. На первой стадии приводится список положений, которые дают приблизительное описание информатики как образовательной дисциплины, не претендующее на ее строгое определение. На второй стадии перечисляются три публикации, которые дают комплементарные характеристики информатики. Эти публикации рассматриваются в статье как три источника теоретических оснований анализируемой стратегии.

Ключевые слова: компьютерное образование; мониторинг систем преподавания; информатика как фундаментальная наука; информатика как образовательная дисциплина; стратегия преподавания информатики

DOI: 10.14357/08696527220311

1 Введение

Апрельский номер журнала *Communications of the ACM* за 2019 г. включает специальный раздел «Информатика в Европе», состоящий из статей, посвященных актуальным темам в сфере информатики как науки, вопросам ее преподавания и средствам разработки информационных технологий (ИТ) [1]. Раздел состоит из 14 статей, в которых, в частности, дано описание:

- факторов, определяющих направления перспективных европейских исследований и разработок [2];
- кадрового ландшафта европейского сектора ИТ [3];

* Исследование выполнено с использованием ЦКП «Информатика» ФИЦ ИУ РАН.

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

- подхода к позиционированию информатики как фундаментальной науки в системе современного научного познания и построению стратегии Европейского компьютерного образования [4].

По мнению авторов последней статьи, информатику следует рассматривать как дисциплину, не менее важную в системе образования, чем математика, естественные науки и иностранные языки. В ней анализируется «Первый план действий в области цифрового образования», одобренный Еврокомиссией (ЕК) в 2018 г. [5, 6]¹, а также отчет о преподавании информатики в европейском регионе, включая РФ, а также в Израиле [8].

В этом отчете сформулирован ряд рекомендаций, которые ориентированы на улучшение ситуации с изучением информатики в Европе на всех уровнях образования. На основе материалов отчета был подготовлен доклад с описанием предлагаемой стратегии ЕКО (European digital education). Доклад одобрен Европейским комитетом по информатике (The Informatics Europe Board) и Европейским советом АСМ (The ACM Europe Council). В нем информатика позиционируется как фундаментальная наука с собственными научными моделями, методами и подходами [9].

Во многих отношениях основные позиции этого доклада совпадают с инициативой Президента США от 2016 г. [10]. Принципиально иная позиция Европейского доклада [9], которая отличает его от этой инициативы, — это диадический подход к построению стратегии ЕКО (two-tier strategy) и преподаванию, с одной стороны, информатики как самостоятельного предмета, с другой стороны, методов и средств информатики, интегрированных в предметные области других наук.

В материалах отчета [8], послуживших исходными данными для подготовки доклада [9], говорится, что информатика кардинально изменилась за последние 50 лет, что современную информатику сложно уложить в «прокрустово ложе» естественных или технических наук: «В то время как естественные науки определяются применительно к миру (world), в котором мы живем, информатику как научную дисциплину определить сложнее; у нее нет эмпирических основ как у естественных наук; это нечто большее, чем использование только логических выводов (как в математике); и это далеко не только сочетание инженерных принципов и технологий. По этим причинам мы говорим об описании информатики как дисциплины, но не о ее определении» [8, с. 7]. Таким образом, авторы этого отчета не претендуют на строгое определение предметной области информатики и предлагают только приблизительное ее описание. Кроме того, в отчете они перечисляют три публикации [11–13], которые содержат комплементарные характеристики информатики. Эти публикации рассматриваются в статье как три источника построения теоретических оснований стратегии ЕКО.

Цель статьи состоит в кратком описании трех публикаций (разд. 2) и диадического подхода к построению стратегии ЕКО (разд. 3).

¹Есть новый вариант плана действий ЕК в области цифрового образования (2021–2027) [7].

2 Три источника

На первых этапах становления информатики объектами ее исследований были в основном компьютеры и окружающие их явления. Например, в 1967 г. А. Ньюэлл, А. Дж. Перлис и Х. А. Саймон писали: «Объектам и явлениям соответствуют те науки, которые их изучают. Появились компьютеры. Следовательно, назначение информатики — это изучение компьютеров» [14]. Такой подход к определению информатики как компьютерной науки долгие годы был доминирующим в системе компьютерного образования.

В работе [11]¹ был предложен концептуально новый подход к определению информатики как фундаментальной науки. Второй раздел *The definition of the term “informatics”* этой работы посвящен определению термина «информатика». В этом разделе Кристен Ньюгор пишет: «Термин “computer science” следует заменить на “informatics”. Несколько лет назад выбор между этими двумя терминами казался, скорее всего, несущественным. Обсуждения терминологии часто считают праздными, но иногда они могут отражать ключевые различия во мнениях... Сегодня, к сожалению, используется термин “computer science”. Этот термин имеет тенденцию поддерживать слишком узкое представление об информационных системах, которые в настоящее время связывают коллективы людей и интегрируют разнообразные средства обработки информации. При этом коллективы взаимодействуют как посредством межличностных связей людей, так и с использованием электронных каналов связи» [11].

После констатации важности четкого определения терминов «информатика» и «информация» для формирования адекватного представления об информационных системах К. Ньюгор предлагает свою дефиницию первого термина: «Информатика — это наука, которая имеет своей областью [исследований] информационные процессы и *связанные с ними феномены в артефактах, обществе и природе* (курсив добавлен. — И. З.)» [11]. Далее дается ссылка на определение понятия «феномен» в словаре Webster 1960 г. («любой факт, обстоятельство или событие, которые сенсорно воспринимаются и которые могут быть научно описаны или оценены» [15]).

Затем К. Ньюгор дает расширенное толкование этого понятия, которое кардинально отличается от определения в словаре Webster 1960 г. [15]: «Важными примерами феноменов являются: живые организмы, неодушевленные объекты (включая артефакты, такие как, например, машины), события и процессы (например, выполнение компьютерных программ). Мы также можем говорить о *когнитивных феноменах, происходящих в сознании людей*, в отличие от явных [сенсорно воспринимаемых] *феноменов, находящихся вне сознания* (курсив добавлен. — И. З.)» [11].

¹Первый из трех источников — это доклад Кристена Ньюгора (Университет Осло), который был прочитан им на Всемирном конгрессе IFIP (International Federation of Information Processing — Международная федерация по обработке информации) в Дублине в 1986 г.

Толкование К. Ньюгора с делением феноменов, *связанных с информационными процессами*, на когнитивные, происходящие в сознании людей и формирующие концепты знания человека, и сенсорно воспринимаемые (например, тексты), существующие вне сознания, является концептуально новым подходом к определению информатики как науки. Отметим, что толкование понятия «феномен», расширенное по объему его значения, позже было включено и в современный онлайн-словарь Merriam-Webster¹ (в новом определении удалена *сенсорная восприимчивость* для феноменов, поддающихся научному описанию).

Можно предположить, что выбор авторами отчета [8] работы [11], опубликованной более 35 лет назад, был обусловлен включением в предметную область информатики широкого спектра феноменов, связанных с информационными процессами в *технических, живых и социальных системах*.

Из трех публикаций, выбранных авторами отчета, второй источник — это книга Дэвида Харела (Научный институт Вейцмана), опубликованная в 1987 г., с изложением основ информатики, ряда важных и базовых ее тем с алгоритмической точки зрения. В книге подчеркивается фундаментальный характер информатики и анализируются три вида сложности проблем этой науки: вычислительная, системная и когнитивная [12].

Для определения предметной области информатики существенный интерес представляет сложность третьего вида — когнитивная, которая порождается использованием в ИТ сущностей, принадлежащих ментальной среде (концепты знания в сознании людей), или сущностей двойной природы, принадлежащих ее границам с другими средами предметной области информатики [17, 18]. Этот вид сложности связан с задачами, которые не удается решить с применением формальных алгоритмов, так как часто они не поддаются точной постановке, а сущностям, затрагиваемым в процессе их постановки и решения, трудно дать формальные определения. Примером такой задачи может служить поиск новых значений языковых единиц в текстах [19].

Дэвид Харел называет такие задачи «псевдоалгоритмическими» проблемами (pseudoalgorithmic problems), подчеркивая в своей книге, что он следует формальной трактовке термина «алгоритм». Главный вопрос при решении подобных проблем: каким образом представить сложное знание человека так, чтобы его представление поддавалось алгоритмической обработке, применению систем и средств информатики для их решения [12, с. 402].

Из трех публикаций, выбранных авторами отчета [8], третья работа, комбинаторная с [11, 12], — это статья Питера Деннинга и Пола Розенблюма [13]. В ней предложен новый вариант позиционирования информатики в системе современного научного познания и обоснована необходимость существенного расширения ее предметной области.

¹В современном онлайн-словаре Merriam-Webster приводится следующая дефиниция термина «феномен»: «Факт или событие, представляющие научный интерес, поддающиеся научному описанию и объяснению» [16]. Она соответствует толкованию К. Ньюгора, но не содержит явного деления феноменов на когнитивные и сенсорно воспринимаемые.

Один из подходов к ее расширению и предлагаемая парадигма деления предметной области на среды различной природы приведены в работах [17, 18, 20, 21]. В первых двух работах по развитию варианта П. Деннинга и П. Розенблюма были описаны пять сред предметной области информатики (ДНК-среда, нейросреда, ментальная, информационная и цифровая среды) и границы между ними. Отметим, что предлагаемое развитие этого варианта охватывает следующие ключевые идеи и К. Ньюгора, и Д. Харела:

- деление феноменов, связанных с информационными процессами, на когнитивные, происходящие в сознании людей (ментальная среда), и сенсорно воспринимаемые (информационная среда), находящиеся вне сознания, что имплицитно уже вводит в предметную область информатики как фундаментальной науки сущности разной природы [11];
- определение класса проблем когнитивной сложности, для решения которых с применением систем и средств информатики необходимо представить сложное знание так, чтобы оно поддавалось алгоритмической обработке в цифровой среде [12].

Рассмотренные работы К. Ньюгора, Д. Харела, П. Деннинга и П. Розенблюма и Касперсена с соавт., обсуждающие вопросы позиционирования информатики в системе современного научного познания, были опубликованы в период с 1986 по 2019 гг. В настоящее время эти вопросы сохраняют свою актуальность и продолжается их активное обсуждение, в том числе с позиции компьютерного образования, качество которого существенно зависит от их решения [22].

В заключение этого раздела отметим, что в упомянутой статье [22] Матти Тедре и Джон Паюнен не только рассмотрели камни преткновения и проблемные места в дискуссиях о предметной области и теоретических основаниях информатики с позиции компьютерного образования. Они также сформулировали несколько конструктивных положений, ориентированных на решение ряда актуальных вопросов ее преподавания.

Во-первых, для компьютерного образования они предлагают использовать в качестве теоретических оснований те ключевые модели, которые характерны именно для информатики.

Во-вторых, поскольку такие работы, как проектирование, разработка систем и средств информатики с учетом социальных аспектов, уже включаются в программы компьютерного образования, то значение той или иной теории в процессе преподавания должно оцениваться по критериям философии проектирования, разработки систем и средств информатики с учетом социальных аспектов, а не с позиций философии естественных наук.

В-третьих, вместо того чтобы навязывать нерелевантные положения исследовательских парадигм из других отраслей знания, философия компьютерного образования должна стремиться к разработке своей парадигмы преподавания и описанию ее связей с информатикой как фундаментальной наукой в собственной системе терминов [22].

3 Диадический подход к построению стратегии Европейского компьютерного образования

Как отмечено во введении, в докладе *Informatics for All. The Strategy* дано описание стратегии ЕКО, одобренной Европейским комитетом по информатике и Европейским советом АСМ, которая выделяет два направления преподавания информатики на всех уровнях образования [9]. В рамках первого направления, получившего название «специализация», информатика изучается как самостоятельный предмет. В рамках второго направления, получившего название «интеграция», методы и средства информатики интегрируются в предметные области других наук через их учебные программы.

В стратегии ЕКО в процессе преподавания информатики предлагается учитывать существенное расширение ее предметной области. На первых этапах становления основными объектами ее предметной области были компьютерные коды, их генерация и преобразования в цифровой среде, а также явления, окружающие компьютеры.

В конце прошлого столетия изучение информационных трансформаций в живых системах стали также относить к предметной области информатики вместе с генерацией и преобразованиями компьютерных кодов в цифровой среде. Использование слова *computing* в англоязычной литературе вместо двух слов *computer science* стало подчеркивать более широкую предметную область информатики. По словам П. Деннинга, «первоначальное определение информатики [как компьютерной науки]... сейчас устарело. Она изучает и естественные, и искусственные информационные процессы. Информатика включает компьютерную науку, разработку программного обеспечения, информационные технологии, информационную науку и информационные системы» [23].

Такое расширение предметной области информатики обуславливает и пересмотр понятия *computation*. Согласно Г. Розенбергу, «нам придется пересмотреть понятие *computation*, которое должно быть в состоянии учитывать также и обработку информации, происходящую в природе... На самом деле исследования информационных трансформаций в живых системах уже привели к пересмотру ряда парадигм, лежащих в основе традиционного понятия *computation*» [24].

Таким образом, предметная область информатики принципиально изменилась с момента публикации в 1967 г. статьи [14]. Характеризуя это изменение в стратегии ЕКО, авторы [14] позиционируют информатику как основополагающую (*foundational*) дисциплину, которая играет ключевую роль в системе образования XXI в. наряду с математикой, естественными науками и иностранными языками.

Одновременно с изменением предметной области самой информатики она стала оказывать существенное влияние и на преподавание других дисциплин, поэтому в стратегию ЕКО включено второе направление преподавания, интегрирующее методы и модели информатики в предметные области других наук через их учебные программы. Важно отметить, что это влияние информатики на

преподавание других дисциплин было обусловлено кардинальным изменением ее роли в системе современного научного познания.

Описанию новой роли информатики был посвящен аналитический отчет по вопросам обеспечения конкурентоспособности США в XXI в., подготовленный в 2005 г. Консультативным комитетом по информационным технологиям при Президенте США [25]. В этом отчете говорится о революционном изменении парадигмы научных исследований: если раньше «древо познания» держалось на «двух китах» (в оригинале pillar — столп, опора, основание) — теория и эксперимент, то сейчас во всех отраслях науки оно держится на «трех китах» — теория, эксперимент, методы и средства информатики, включая ИТ, обеспечивающие проведение эксперимента. Такое изменение парадигмы практически любого научного исследования и обуславливает необходимость во втором направлении стратегии ЕКО.

В заключение этого раздела отметим, что в опубликованной в 2016 г. инициативе Президента США [10] революционное изменение парадигмы научных исследований (зафиксированное еще в 2005 г. Консультативным комитетом по информационным технологиям при Президенте США [25]) не нашло своего отражения. В стратегии ЕКО деление компьютерного образования на два направления [4] обусловлено кардинальным изменением роли информатики в системе современного научного познания.

4 Заключение

Отчет [8] содержит результаты мониторинга состояния систем образования в области информатики¹, а также результаты сопоставительного анализа этих систем для 55 административных единиц (АЕ) Европы, включая РФ, а также Израиль². В Приложении С этого отчета приведены схемы и карты с результатами мониторинга и сопоставительного анализа, позволяющие сравнить состояние систем образования в области информатики в разных АЕ между собой. Каждая карта или схема отражает, как правило, одну характеристику состояния сопоставляемых систем. В Приложении D отчета приведено структурированное описание состояния системы образования в области информатики для каждой из 55 АЕ по комплексу характеристик³.

Поскольку авторы отчета на основе данных мониторинга анализировали состояние преподавания информатики, им потребовалось дать определение этой

¹Подготовка к мониторингу началась в 2013 г. На сбор данных ушло два года. На основе результатов мониторинга в 2017 г. был опубликован отчет [8].

²Если в стране единая система образования в области информатики, то АЕ — это страна. Если в стране регионы с автономными системами образования, то АЕ — регион этой страны, например Бавария. В Бельгии — 3 региона, Соединенном Королевстве — 4, Испании — 15, в Германии — 16.

³Структурированное описание состояния системы образования в области информатики в РФ приведено в [8, с. 196–197].

дисциплины. Очевидно, что сегодня отсутствует конвенциональное определение предметной области информатики, поэтому перед авторами отчета встала весьма трудная задача. Их заслуживающий внимания подход включает две стадии ее предварительного и весьма содержательного решения.

Сначала авторы отчета дают не определение информатики, а приводят ее приблизительное описание, включающее следующие положения:

- информатика кардинально изменилась за последние 50 лет, и поэтому ранее используемые ее определения устарели (о чем в 2007 г., т. е. за 10 лет до публикации отчета [8], написал П. Деннинг [23]);
- современную информатику сложно уложить в «прокрустово ложе» естественных или технических наук;
- по сравнению с естественными науками информатику как научную дисциплину определить сложнее, так как у нее нет эмпирических основ как у естественных наук;
- информатика — это нечто большее, чем использование только логических выводов (как в математике);
- информатика — это далеко не только сочетание инженерных принципов и технологий.

Перечисленные положения используются на первой стадии приблизительного описания информатики как образовательной дисциплины, которое не претендует на ее строгое определение [8]. На второй стадии перечисляются три публикации [11–13], которые содержат комплементарные характеристики информатики. Эти публикации рассматриваются в статье как три источника теоретических оснований стратегии ЕКО.

Литература

1. *Fatourou P., Hankin C.* Welcome to the Europe Region Special Section // Commun. ACM, 2019. Vol. 62. No. 4. P. 30.
2. *Gulliksen J.* Incorporating Europe's values in future research // Commun. ACM, 2019. Vol. 62. No. 4. P. 40–41.
3. *Korrigane L.* A demographic snapshot of the IT workforce in Europe // Commun. ACM, 2019. Vol. 62. No. 4. P. 32–33.
4. *Caspersen M. E., Gal-Ezer J., McGettrick A., Nardelli E.* Informatics as a fundamental discipline for the 21st century // Commun. ACM, 2019. Vol. 62. No. 4. P. 58–63.
5. Communication on the first digital education action plan. — The European Commission, 2018. 12 p. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0022&from=EN>.
6. Communication on the First Digital Education Action Plan: Commission staff working document. — The European Commission, 2018. 31 p. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018SC0012&from=EN>.

7. Digital Education Action Plan (2021–2027). — The European Commission. <https://education.ec.europa.eu/ru/focus-topics/digital-education/action-plan>.
8. The Committee on European Computing Education. Informatics Education in Europe: Are we all in the same boat? — New York, NY, USA: ACM, 2017. Technical Report. 251 p. <https://portalparts.acm.org/hippo/cecereport.pdf>.
9. *Caspersen M. E., Gal-Ezer J., McGettrick A., Nardelli E.* Informatics for all: The strategy. — New York, NY, USA: ACM, 2018. 16 p. <https://www.acm.org/binaries/content/assets/public-policy/acm-europe-ie-i4all-strategy-2018.pdf>.
10. President Obama announces computer science for all initiative. — The White House, Office of the Press Secretary, 2016. 12 p. <https://obamawhitehouse.archives.gov/the-press-office/2016/01/30/fact-sheet-president-obama-announces-computer-science-all-initiative-0>.
11. *Nygaard K.* Program development as a social activity // Information processing / Ed. H.-J. Kugler. — North Holland: Elsevier Science Publs. B.V., 1986. P. 189–198.
12. *Harel D.* Algorithmics — the spirit of computing. — Reading, MA, USA: Addison-Wesley, 1987. 514 p.
13. *Denning P., Rosenbloom P.* Computing: The fourth great domain of science // Commun. ACM, 2009. Vol. 52. No. 9. P. 27–29.
14. *Newell A., Perlis A., Simon H.* Computer science // Science, 1967. Vol. 157. No. 3795. P. 1373–1374.
15. Webster’s New World Dictionary of the American language / Eds. D. B. Gurelnik, J. H. Friend. — New York, NY, USA: The World Publishing Co., 1960. 1760 p.
16. Definition of phenomenon // Merriam-Webster’s Dictionary. <https://www.merriam-webster.com/dictionary/phenomenon>.
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. *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.
20. *Зацман И. М.* Методология обратимой генерализации в контексте классификации информационных трансформаций // Системы и средства информатики, 2018. Т. 28. № 2. С. 128–144.
21. *Зацман И. М.* Модель процесса извлечения новых терминов и тональных слов из текстов // Системы и средства информатики, 2022. Т. 32. № 2. С. 115–127.
22. *Tedre M., Rajunen J.* Grand theories or design guidelines? Perspectives on the role of theory in computing education research // ACM T. Comput. Educ., 2022. doi: 10.1145/3487049.
23. *Denning P.* Computing is a natural science // Commun. ACM, 2007. Vol. 50. No. 7. P. 13–18.
24. *Rozenberg G.* Computer science, informatics, and natural computing — personal reflections // New computational paradigms: Changing conceptions of what is computable. — New York, NY, USA: Springer, 2008. P. 373–379.

25. Computational science: Ensuring America's competitiveness. — Arlington, VA, USA: National Coordination Office for Information Technology Research and Development, 2005. Report to the President. 104 p. http://www.nitrd.gov/pitac/reports/20050609_computational/computational.pdf.

Поступила в редакцию 15.08.22

THE STRATEGY OF EUROPEAN DIGITAL EDUCATION: THEORETICAL FOUNDATIONS FOR ITS CREATING

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 strategy of the European digital education is analyzed, in creating of which the outcomes of two-year monitoring of the state of informatics teaching systems in the regions and countries of Europe, including the Russian Federation, were used. Since the authors of the strategy were analyzing the teaching system state of informatics, they needed to give a modern description of this discipline. The strategy uses an approach to positioning informatics within modern scientific knowledge, considering the absence of a conventional definition of its subject area. The aim of the paper is to describe this approach which includes two stages. At the first stage, a list of propositions is given that propose an approximate description of informatics as an educational discipline but does not pretend to be its strict definition. At the second stage, three publications are listed that give mutually complementary characteristics of informatics. In the paper, these publications are considered as three sources of theoretical foundations of the analyzed strategy.

Keywords: digital education; monitoring of teaching systems; informatics as a fundamental science; informatics as an educational discipline; strategy of teaching informatics

DOI: 10.14357/08696527220311

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).

References

1. Fatourou, P., and C. Hankin. 2019. Welcome to the Europe Region Special Section. *Commun. ACM* 62(4):30.

2. Gulliksen J. 2019. Incorporating Europe's values in future research. *Commun. ACM* 62(4):40–41.
3. Korrigane L. 2019. A demographic snapshot of the IT workforce in Europe. *Commun. ACM* 62(4):32–33.
4. Caspersen, M. E., J. Gal-Ezer, A. McGettrick, and E. Nardelli. 2019. Informatics as a fundamental discipline for the 21st century. *Commun. ACM* 62(4):58–63.
5. The European Commission: Communication on the first digital education action plan. 12 p. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0022&from=EN> (accessed August 31, 2022).
6. The European Commission: Communication on the first digital education action plan: Commission staff working document. 31 p. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018SC0012&from=EN> (accessed August 31, 2022).
7. The European Commission. Digital Education Action Plan (2021–2027). Available at: <https://education.ec.europa.eu/ru/focus-topics/digital-education/action-plan> (accessed August 31, 2022).
8. Informatics education in Europe: Are we all in the same boat? 2017. New York, NY: ACM. Technical Report. Available at: <https://www.informatics-europe.org/component/phocadownload/category/10-reports.html?download=60:cece-report> (accessed August 31, 2022).
9. Caspersen M. E., J. Gal-Ezer, A. McGettrick, and E. Nardelli. 2018. Informatics for all: The strategy. New York, NY: ACM. 16 p. Available at: <https://www.acm.org/binaries/content/assets/public-policy/acm-europe-ie-i4all-strategy-2018.pdf> (accessed August 31, 2022).
10. President Obama announces computer science for all initiative. 2016. The White House, Office of the Press Secretary. 12 p. Available at: <https://obamawhitehouse.archives.gov/the-press-office/2016/01/30/fact-sheet-president-obama-announces-computer-science-all-initiative-0> (accessed August 31, 2022).
11. Nygaard, K. 1986. Program development as a social activity. *Information processing*. Ed. H.-J. Kugler. North Holland: Elsevier Science Publs. B.V. 189–198.
12. Harel, D. 1987. *Algorithmics — the spirit of computing*. Reading, MA: Addison-Wesley. 514 p.
13. Denning, P., and P. Rosenbloom. 2009. Computing: The fourth great domain of science. *Commun. ACM* 52(9):27–29.
14. Newell, A., A. Perlis, and H. Simon. 1967. Computer science. *Science* 157(3795):1373–1374.
15. Guralnik, D. B., and J. H. Friend, eds. 1960. *Webster's New World Dictionary of the American language*. New York, NY: The World Publishing Co. 1760 p.
16. Definition of phenomenon. *Merriam-Webster's Dictionary*. Available at: <https://www.merriam-webster.com/dictionary/phenomenon> (accessed August 31, 2022).
17. Zatsman, I. M. 2019. Interfejsy 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. 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.

20. Zatsman, I. M. 2018. Metodologiya obratimoy generalizatsii v kontekste klassifikatsii informatsionnykh transformatsiy [Methodology of reversible generalization in the context of classification of information transformations]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(2): 128–144.
21. Zatsman, I. 2022. Model' protsessy izvlecheniya novykh terminov i tonal'nykh slov iz tekstov [A model of discovering novel terms and sentiments in texts]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 32(2):115–127.
22. Tedre, M., and J. Pajunen. 2022. Grand theories or design guidelines? Perspectives on the role of theory in computing education research. *ACM T. Comput. Educ.* doi: 10.1145/3487049.
23. Denning, P. 2007. Computing is a natural science. *Commun. ACM* 50(7):13–18.
24. Rozenberg, G. 2008. Computer science, informatics, and natural computing — personal reflections. *New computational paradigms: Changing conceptions of what is computable*. New York, NY: Springer. 373–379.
25. Computational science: Ensuring America's competitiveness. Arlington, VA: National Coordination Office for Information Technology Research and Development, 2005. Report to the President. 104 p. Available at: http://www.nitrd.gov/pitac/reports/20050609_computational/computational.pdf (accessed February 12, 2022).

Received August 15, 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

МУЛЬТИВАЛЮТНАЯ ЦИФРОВАЯ ПЛАТЕЖНАЯ СИСТЕМА: ОСНОВЫ ТЕХНОЛОГИЧЕСКОГО ОБЕСПЕЧЕНИЯ*

*В. Д. Ильин*¹

Аннотация: Символьная модель (s-модель) нормализованной мультивалютной цифровой платежной системы (НМЦП-системы) рассматривается как составляющая s-модели нормализованного экономического механизма (НЭМ). Технологии НМЦП-системы относятся к технологиям товарно-денежного обращения НЭМ. Ядро s-модели НМЦП-системы представлено ситуационно изменяемыми обязательными и ориентирующими требованиями к свойствам и функционированию системы. Участники НМЦП-системы идентифицируются по ИНН и компьютерным устройствам, зарегистрированным в системе для онлайн-взаимодействия при совершении платежей. На каждом отрезке времени для участников определены множества допустимых торговых сделок, валют, применимых для расчетов, и правила торговых сборов. Платеж с помощью НМЦП-системы может быть совершен в любой валюте из списка, являющегося пересечением множеств валют, разрешенных банками-регуляторами государств, к которым относятся участники торговой сделки.

Ключевые слова: мультивалютная платежная система; нормализованный экономический механизм; персональные электронные банки; корпоративные электронные банки; онлайн-сервисы

DOI: 10.14357/08696527220312

1 Введение

В 2009 г. была обоснована необходимость плановой адаптации экономического механизма по мере изменений социально-экономических отношений и технологической среды [1]. В этой же статье [1, с. 131] приведена аргументация необходимости применения национальных валют во внешнеторговых расчетах:

«Внешнеторговые сделки должны осуществляться при выполнении следующих условий:

- применимы деньги только стран, участвующих в сделке;
- выполняются ограничения на перечень товаров, объемы экспорта и импорта и др., которые определены государственными законами и нормативными актами обеих стран».

*Результаты получены при выполнении научно-исследовательской работы «Моделирование социальных, экономических и экологических процессов» (№ 0063-2016-0005), выполняемой в соответствии с государственным заданием ФАНО России для Федерального исследовательского центра «Информатика и управление» РАН.

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, vdilyin@yandex.ru

Санкционная ограниченность применимости для российских пользователей международных платежных систем Visa и Mastercard (с марта 2022 г.) [2] еще раз подтвердила безотлагательность создания государственной *нормализованной мультивалютной цифровой платежной системы (НМЦП-системы)*, пригодной для совершения внутренних и внешних торговых сделок.

Создание этой системы целесообразно реализовать как первоочередной проект *цифровизации экономической деятельности* [3–7].

Выделение фрагментов текста. Для выделения определений, замечаний и примеров используются средства языка TSM-комплекса (TSM: textual symbolic modeling), разработанного для формализованного описания текстовых моделей¹ [8].

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 представлены замечания об основах технологического обеспечения НМЦП-системы. Чтобы без крайней необходимости не отсылать читателя к ранее опубликованным статьям автора, в тексте приведена часть ключевых определений, относящихся к модели *нормализованного экономического механизма* [9, 10].

2 Нормализованная мультивалютная цифровая платежная система: технологическое ядро

□ *Цифровой технологией* экономической деятельности будем называть совокупность методов, средств и правил реализации этой деятельности посредством онлайн-сервисов, функционирующих в среде цифровых двойников (*Digital Twins*²). □

◇ Далее под *технологией* понимается *цифровая технология*. ◇

Прикладной смысл конкретной технологии определен системой правил, направляющих и контролирующих действия *участников экономической деятельности (эд-участников)*.

□ *Цифровой двойник (DT)* — цифровая *s-модель* [8] некоторого объекта (технического устройства, деятельности или организации), предназначенная для анализа и совершенствования *s-моделируемого* объекта. □

¹Ильин В. Д. Символьное моделирование // Большая российская энциклопедия (электронная версия). http://bigenc.ru/technology_and_technique/text/4010980.

²The new age of manufacturing: Digital Twin technology & IIoT. <https://medium.com/@lior.kitain/the-new-age-of-manufacturing-digital-twin-technology-iiot-494acee5572a>.

DT-технология цифрового s-моделирования предполагает необходимость получения данных о поведении объекта (путем измерения значений определенной совокупности параметров). С момента своего создания (как образа объекта) DT накапливает и обрабатывает данные о своем прообразе, обмениваясь с ним сообщениями и обновляясь.

С 2010-х гг. наблюдается интенсивное совершенствование технологической инфраструктуры цифровизации экономической деятельности, включающей технологии *цифровых двойников*, *M2M-технологии (Machine-to-Machine)*, технологии *облачных вычислений (Cloud Computing)* и электронных сервисов, *интернета вещей (Internet of Things, IoT)* [11–16]. Все большее значение приобретает технологии получения и своевременной (исключающей превышение допустимых задержек) передачи больших объемов быстро изменяющихся данных. В наши дни *технология связи 5G* [17] продолжает успешно доказывать свою эффективность в совершенствовании различных видов деятельности, связанных с обработкой *больших данных*, применением систем *искусственного интеллекта*, *промышленного интернета вещей* и др. (○ *интеллектуальные сети энергоснабжения*; синергетические транспортные системы типа «*машина–дорога*»; системы «*умный порт*» и др. ○).

Технологии связи 5G важны и для инфраструктурного технологического обеспечения НМЦП-системы, в которой программно-аппаратными средствами пользователей служат мобильные компьютерные устройства (смартфоны, планшеты и др.). В таких устройствах установлены *корпоративные электронные банки (КЭБы)* для юридических лиц или *персональные электронные банки (ПЭБы)* для физических лиц [9, 10], используемые для совершения операций в НМЦП-системе.

◇ Исходим из того, что технологии НМЦП-системы при любой попытке нарушить заданные и контролируемые ими правила оставляют столько же шансов на успех, сколько имеет попытка сделать недопустимый ход в игре с шахматной программой. ◇

□ *Нормализованный экономический механизм* — человеко-машинная система, связанная с государственным механизмом отношениями *координации* и *подчинения* [18]. Нормализованный экономический механизм работает на основе онлайн-сервисов, функционирующих в среде цифровых двойников. Устройство НЭМ и правила его функционирования стимулируют (посредством налогов, пошлин и других экономических инструментов) ведение хозяйственной деятельности, ориентированной на защиту и развитие *потенциала страны* [18]. □

□ НМЦП-система — составляющая *системы товарно-денежного обращения (ТДО-системы)* НЭМ, где сделки совершаются в соответствии с *ситуационными цифровыми договорами (СЦД)*, выполнение которых контролируется с помощью *ситуационной интеллектуальной системы (СЦД-системы)* [19]. □

Правила исполнения каждого договора контролируются программно во взаимодействии с цифровыми двойниками, отслеживающими события, определенные в договоре. Выполнение каждого СЦД представлено спецификациями событий,

для которых определены *портреты ситуаций*, хранящиеся в памяти СЦД-системы. Портреты ситуаций, возникающих при выполнении договора, строятся СЦД-системой на основе данных, поставляемых *цифровыми двойниками*.

Перемещение товаров и платежей в процессе выполнения ТДО-сделки, сопровождаемое перемещением соответствующих цифровых кодов, отражается в памяти *цифровой системы учета величин имущественного права* участников ТДО-сделок, реализуемой на основе информационно-безопасных банковских онлайн-сервисов [9, 10].

□ Каждому товару поставлена в соответствие электронная унифицированная спецификация, содержащая наименование товара, назначение и другие сведения, необходимые для его идентификации при совершении ТДО-сделок. □

□ Уникальные мультивалютные счета участников ТДО-сделок, открываемые и программно контролируемые с помощью онлайн-сервисов банка-регулятора, размещаются в *ПЭБах* (для физических лиц) и в *КЭБах* (для юридических лиц) [9, 10]. □

□ *Система имущественных статусов (ис-система)* — система цифрового представления имущественных статусов *участников экономической деятельности (эд-участников)*, реализуемая на основе онлайн-сервисов НЭМ. Денежная составляющая имущественных статусов представлена значениями сумм *нормализованных денег* в разделах *уникальных мультивалютных 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 цифровой мультивалютный банковский счет, размещенный в *ПЭБе* физического лица или *КЭБе* юридического лица. □

□ Изменение состояния ис-счета технически невозможно без документированного зашифрованного подтверждения владельца ис-счета (а в неординарных ситуациях — после документированного зашифрованного подтверждения определенного законом государственного органа). □

◇ Изменение состояния ис-счета осуществимо 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 государственного сбора или др. ○). Таким способом исключаются возможности нарушения установленных законом правил товарно-денежного обращения (недопустимые экономические сделки, неуплата установленных законом государственных сборов и др.). ◇

Банковская система НЭМ включает *ПЭБы*, *КЭБы*, *банки-провайдеры* и *банк-регулятор*, под управлением которого функционируют все другие банки [9, 10].

□ *Технология назначенных платежей* — совокупность методов, средств и правил реализации *назначенных платежей* [20] в НМЦП-системе. Эта технология обеспечивает гарантированную оплату заказанных товаров, правила реализации

которой жестко связаны с правилами исполнения заказа, определенными в СЦД. Сумма назначенного платежа переводится на счет получателя сразу после того, как заказчик подтвердил исполнение заказа. В случае неисполнения заказа назначенный платеж отменяется. □

□ Технология назначенных платежей для пользователей НМЦП-системы из стран, реализующих антироссийские экономические санкции: (1) валютные платежи за российские товары аккумулируются на специальном мультивалютном счете Центрального банка России; (2) с этого счета НМЦП-система списывает валютные средства только для расчетов за товары, продаваемые российским покупателям продавцами из упомянутых стран. □

□ *Технология долговой торговли* позволяет отсроченную часть оплаты товара оформить как *долг покупателя продавцу* (не банку!), имеющему КЭБ или ПЭБ. Применительно к долгам, образовавшимся в результате продажи приоритетных товаров, могут действовать правила досрочного возврата долгов продавцам из средств банка-регулятора [21]. □

◇ *Досрочным возвратом долгов продавцам приоритетных товаров* банк-регулятор реализует государственные программы поддержки производства и реализации таких товаров. ◇

3 Заключение

В НМЦП-системе технология платежей реализуется исходя из того, что для эд-участников (идентифицируемых по ИНН и устройствам, зарегистрированным в системе) на каждом отрезке времени определено некоторое множество товаров, множество применимых для расчетов валют, таможенные правила и правила внешнеторговых сборов. Для каждого типа товара эд-участники имеют возможность совершать сделки в любой валюте из списка, являющегося пересечением множеств валют, разрешенных банками-регуляторами государств, к которым относятся участники сделки.

Литература

1. Ильин В. Д. Модель нормализованной экономики (НЭк-модель): основы концепции // Управление большими системами, 2009. Вып. 25. С. 116–138.
2. Системы Mastercard и Visa заявили о приостановке работы в России // РИА Новости, 06.03.2022. <https://ria.ru/20220306/sistemy-1776871743.html>.
3. Tapscott D. The digital economy: Promise and peril in the age of networked intelligence. — New York, NY, USA: McGraw-Hill, 1996. 342 p.
4. Christensen C. M. The innovator's dilemma: When new technologies cause great firms to fail. — Boston, MA, USA: Harvard Business School Press, 1997. 288 p.
5. The New Digital Economy: How it will transform business. — Oxford, U.K.: Oxford Economics, 2015. 34 p. <http://www.pwc.com/mt/en/publications/assets/the-new-digital-economy.pdf>.

6. G20 digital economy development and cooperation initiative // G20 Summit, 2016. <http://en.kremlin.ru/supplement/5111>.
7. Цифровая экономика Российской Федерации: Программа, утвержденная распоряжением Правительства Российской Федерации от 28 июля 2017 г. № 1632-р. <http://d-russia.ru/wpcontent/uploads/2017/07/programma-tsifrov-econ.pdf>.
8. *Ilyin V. D.* Symbolic modeling (s-modeling): An introduction to theory // Artificial intelligence trends in systems / Ed. R. Silhavy. — Lecture notes in networks and systems ser. — Cham: Springer, 2022. Vol. 502. P. 585–591. doi: 10.1007/978-3-031-09076-9_54.
9. *Ilyin A. V., Ilyin V. D.* Towards a normalized economic mechanism based on e-services // Agris on-line Papers in Economics and Informatics, 2014. Vol. 6. Iss. 3. P. 39–49.
10. *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.
11. *Armbrust M., Fox A., Griffith R., et al.* A view of cloud computing // Commun. ACM, 2010. Vol. 53. No. 4. P. 50–58.
12. *Wei Y., Blake M. B.* Service-oriented computing and cloud computing: Challenges and opportunities // IEEE Internet Comput., 2010. Vol. 14. Iss. 6. P. 72–75.
13. *Lien S. Y., Liau T. H., Kao C. Y., et al.* Cooperative access class barring for machine-to-machine communications // IEEE T. Wirel. Commun., 2012. Vol. 11. No. 1. P. 27–32.
14. *Kim R. Y.* Efficient wireless communications schemes for machine to machine communications // Comm. Com. Inf. Sc., 2011. Vol. 181. No. 3. P. 313–323.
15. *Perera C., Liu C. H., Jayawardena S.* The emerging Internet of Things marketplace from an industrial perspective: A survey // IEEE T. Emerging Topics Computing, 2015. Vol. 3. No. 4. P. 585–598.
16. *Jede A., Teuteberg F.* Understanding socio-technical impacts arising from software-as-a-service usage in companies // Bus. Inf. Syst. Eng., 2016. Vol. 58. No. 3. P. 161–176.
17. Почему 5G называют фундаментом цифровой экономики // C-News, 19.06.2020. https://www.cnews.ru/articles/2020-06-16_pochemu_5g_nazyvayut_fundamentom_tsifrovoj.
18. *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.
19. *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.
20. *Ильин В. Д.* Основания ситуационной информатизации. — М.: Наука, Физматлит, 1996. 180 с.
21. *Ильин А. В., Ильин В. Д.* Ситуационная модель универсального договора в цифровой среде // Системы и средства информатики, 2020. Т. 30. № 3. С. 154–162.

Поступила в редакцию 13.07.22

MULTICURRENCY DIGITAL PAYMENT SYSTEM: FUNDAMENTALS OF TECHNOLOGICAL SUPPORT

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 symbolic model (s-model) of the normalized multicurrency digital payment (NMDP) system is considered as a component of the s-model of the normalized economic mechanism (NEM). The technologies of the NMDP-system relate to the technologies of commodity-money circulation of NEM. The core of the s-model of the NMDP-system is represented by situationally changeable mandatory and orienting requirements for the properties and functioning of the system. The participants of the NMDP-system are identified by individual tax numbers and computer devices registered in the system for online interaction when making payments. At each time interval, the sets of acceptable trade transactions, the currencies used for settlements, and the rules of trade fees are defined. The NCSP-payment can be made in any currency from the list which is the intersection of currencies sets allowed by central banks of the states to which the participants in the trade transaction belong.

Keywords: multicurrency payment system; normalized economic mechanism; personal electronic banks; corporate electronic banks; online services

DOI: 10.14357/08696527220312

Acknowledgments

The results were obtained in the course of the research work “Modeling of social, economic, and ecological processes” (No. 0063-2016-0005) carried out in accordance with the state assignment of FANO of Russia for the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences.

References

1. 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.
2. RIA Novosti [RIA News]. 06.03.2022. Sistemy Mastercard i Visa zavayavili o priostanovke raboty v Rossii [Mastercard and Visa systems announced the suspension of work in Russia]. Available at: <https://ria.ru/20220306/sistemy-1776871743.html> (accessed August 12, 2022).
3. Tapscott, D. 1996. *The digital economy: Promise and peril in the age of networked intelligence*. New York, NY: McGraw-Hill. 342 p.
4. Christensen, C. M. 1997. *The innovator's dilemma: When new technologies cause great firms to fail*. Boston, MA: Harvard Business School Press. 288 p.
5. The new digital economy: How it will transform business. 2015. Oxford, U.K.: Oxford Economics. 34 p. Available at: <http://www.pwc.com/mt/en/publications/assets/the-new-digital-economy.pdf> (accessed August 12, 2022).
6. G20 digital economy development and cooperation initiative. 2016. *G20 Summit*. Available at: <http://en.kremlin.ru/supplement/5111> (accessed August 12, 2022).

7. Tsifrovaya ekonomika Rossiyskoy Federatsii: Programma, utverzhdannaya rasporyazheniem Pravitel'stva Rossiyskoy Federatsii ot 28 iyulya 2017 g. No.1632-r [Digital economy of the Russian Federation: Program approved by Order No.1632-r dated July 28, 2017 of the Government of the Russian Federation]. Available at: <http://d-russia.ru/wpcontent/uploads/2017/07/programma-tsifrov-econ.pdf> (accessed August 12, 2022).
8. Ilyin, V. D. 2022. Symbolic modeling (s-modeling): An introduction to theory. *Artificial intelligence trends in systems*. Ed. R. Silhavy. Lecture notes in networks and systems ser. Cham: Springer. 502:585–591. doi: 10.1007/978-3-031-09076-9_54.
9. Ilyin, A. V., and V. D. Ilyin. 2014. Towards a normalized economic mechanism based on e-services. *Agris on-line Papers in Economics and Informatics* 6(3):39–49.
10. 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.
11. Armbrust, M., A. Fox, R. Griffith, et al. 2010. A view of cloud computing. *Commun. ACM* 53(4):50–58.
12. Wei, Y., and M. B. Blake. 2010. Service-oriented computing and cloud computing: Challenges and opportunities. *IEEE Internet Comput.* 14:72–75.
13. Lien, S. Y., T. H. Liao, C. Y. Kao, et al. 2012. Cooperative access class barring for machine-to-machine communications. *IEEE T. Wirel. Commun.* 11(1):27–32.
14. Kim, R. Y. 2011. Efficient wireless communications schemes for machine to machine communications. *Comm. Com. Inf. Sc.* 181 (3):313–323.
15. Pereyra, C., C. H. Liu, and S. Jayawardena. 2015. The emerging Internet of Things marketplace from an industrial perspective: A survey. *IEEE T. Emerging Topics Computing* 3(4):585–598.
16. Jede, A., and F. Teuteberg. 2016. Understanding socio-technical impacts arising from software-as-a-service usage in companies. *Bus. Inf. Syst. Eng.* 58(3):161–176.
17. C-News. 19.06.2020. Pochemu 5G nazyvayut fundamentom tsifrovoy ekonomiki [Why 5G is called the foundation of the digital economy]. Available at: https://www.cnews.ru/articles/2020-06-16_pochemu_5g_nazyvayut_fundamentom_tsifrovoy (accessed August 12, 2022).
18. 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.
19. 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.
20. Ilyin, V. D. 1996. *Osnovaniya situatsionnoy informatizatsii* [Fundamentals of situational informatization]. Moscow: Nauka, Fizmatlit. 180 p.
21. Ilyin, A. V., and V. D. Ilyin. 2020. Situatsionnaya model' universal'nogo dogovora v tsifrovoy srede [Situational model of universal agreement in a digital environment]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(3):154–162.

Received July 13, 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

ПОДХОД К ПОИСКУ АНОМАЛИЙ В КОНКРЕТНО-ИСТОРИЧЕСКИХ ДАННЫХ

И. М. Адамович¹, О. И. Волков²

Аннотация: Статья продолжает серию работ, посвященных технологии поддержки конкретно-исторических исследований (ПКИИ). Технология построена на принципах сотворчества и краудсорсинга и ориентирована на широкий круг не относящихся к профессиональным историкам и биографам пользователей. Статья посвящена дальнейшему развитию технологии за счет интеграции в нее механизма автоматизированного поиска аномалий в конкретно-исторических данных. Проведен анализ существующего подхода к поиску противоречий в историко-биографических фактах, выявлены его ограничения и недостатки. В рамках развития этого подхода обоснована целесообразность перехода от поиска явных противоречий к поиску аномалий. Проанализированы причины возникновения аномалий, выполнена их классификация, определены особенности аномалий в конкретно-исторических данных. Проведен анализ известных методов обнаружения точечных аномалий, обоснована невозможность использования в технологии методов, основанных на распознавании с учителем, а также метрических методов. Был определен наиболее перспективный метод, основанный на кластеризации, и намечены дальнейшие шаги по его реализации. Описаны и обоснованы необходимые изменения объектной модели технологии.

Ключевые слова: конкретно-историческое исследование; распределенная технология; аномалия; историко-биографический факт; автоматизированная процедура

DOI: 10.14357/08696527220313

1 Введение

Поддержка конкретно-исторических исследований стала одной из актуальных задач современности в связи с 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].

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Adam@amsd.com

²Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Volkov@amsd.com

В [2, 3] описана разработанная в ФИЦ ИУ РАН распределенная технология ПКИИ, основанная на принципах краудсорсинга (мобилизации ресурсов широкого круга добровольцев посредством информационных технологий). Данные технологии организованы в форме семантической сети. Узлы сети представляют собой именованные универсальные классы объектов. Факты задаются значениями экземпляров классов и связями между ними. Связи наследуются из сети классов [4].

Специфика конкретно-исторического исследования состоит в том, что факты, которыми оперирует исследователь (историко-биографические факты), обладают свойством фрагментарности и противоречивости. Получение этих фактов — необходимая и наиболее трудоемкая стадия биографического исследования, состоящая в работе в архивах и просмотре огромного числа документов. Работа исследователя на втором этапе исследования состоит в сопоставлении выявленных фактов с регламентирующими правилами, законами, традициями и т. п. (нормами) и между собой с целью взаимоувязывания и разрешения противоречий [5]. При этом противоречие не означает автоматически ошибки в данных, поскольку любая норма может иметь как сроки действия, так и множество исключений и редакций. Поэтому выявление факта противоречия не становится поводом отбросить якобы ошибочные данные, как предполагает традиционный подход к задаче выявления противоречий, а служит поводом сконцентрироваться на поиске свидетельств наличия оснований для исключений. Таким образом, есть выявление противоречия вместо сужения объема информации дает новый импульс исследованию, что иллюстрирует специфичность конкретно-исторического исследования.

Подход к поиску противоречий в историко-биографических фактах, учитывающий специфику конкретно-исторического исследования, описан в [6]. Проведена классификация видов противоречий в конкретно-исторической информации, описаны и обоснованы необходимые изменения объектной модели технологии, приведено подробное описание алгоритма, определены процедуры обработки результатов поиска.

Подход заключается в рассмотрении двух случаев.

1. Поиск потенциального противоречия между группой фактов и связывающих их норм.

Противоречие в этом случае рассматривается как потенциальное, поскольку в результате проверки пользователем данное сочетание фактов может найти какое-то рациональное объяснение. Поэтому вместо противоречия следует говорить об обнаружении аномального сочетания фактов или об аномальном объекте, атрибутами которого выступают рассматриваемые факты. Предложенный подход предполагает выявление таких аномалий за счет автоматического сравнения с так называемыми прецедентами, т. е. с аналогичными сочетаниями фактов и норм, уже просмотренных пользователями и помеченных ими как аномальные. Недостаток этого подхода заключается

в невозможности автоматического выявления уникальных аномальных объектов, а также аномальных объектов, прецедент для которых еще не выявлен пользователями.

2. Поиск явного противоречия между двумя фактами.

Для этого случая в [6] сделан вывод, что в технологии ПКИИ данная задача сводится к сравнению различных значений одного атрибута одного объекта; такое противоречие видно исследователю непосредственно и не нуждается в автоматизации, какая-либо доработка технологии в этом случае не требуется. Данный вывод верен только в ситуации, когда исследователь осуществляет ручное внесение и обработку фактов, что, безусловно, является одной из основных форм формирования исторических данных в связи с большой сложностью их поиска. Но в технологию ПКИИ интегрирована система Т-парсер [7], осуществляющая автоматическое, адаптированное к специфике биографического поиска извлечение фактов из текстов историко-биографической направленности и решающая задачи фактографического индексирования номинативных (содержащих персональную информацию с указанием имени собственного) и нарративных (повествовательных) источников и их семантической разметки (вычленения значимых фрагментов и оснащения их метаданными). Автоматическая обработка текстовых источников конкретно-исторических данных может порождать большое число узлов семантической сети, ручной поиск аномалий в которой может быть очень трудоемок. Также к недостаткам предложенного подхода следует отнести возможность визуального выявления ошибочного значения атрибута только за счет сравнения с корректным значением этого же атрибута. Но значение данного атрибута может быть единственным, т. е. объект, к которому этот атрибут относится, может быть аномальным, но не содержать явного противоречия в значениях атрибутов.

В связи с этим становится актуальной задача автоматического поиска аномалий в конкретно-исторических данных, свободного от указанных выше недостатков и дополняющего предложенный в [6] подход.

2 Аномалии в конкретно-исторических данных

Обычно выделяют следующие возможные источники аномальных наблюдений [8].

1. Неумышленная ошибочная запись значения наблюдаемого показателя.
2. Резкое изменение внешних и внутренних условий, связанных с формированием результата наблюдения.
3. Резкие отличия показаний объектов исследования, неочевидно находящихся в разных условиях (действие скрытых факторов).
4. Умышленное искажение с определенной целью результатов наблюдений.

Применительно к конкретно-историческим данным указанные выше причины возникновения аномальных данных можно сформулировать следующим образом.

1. Неумышленная ошибка, причиной которой могут быть:

- ошибка формирования данных (непреднамеренная ошибка в исходном документе-источнике, что встречается весьма часто);
- ошибка автоматического формирования базы данных (БД) технологии, вызванная несовершенством алгоритма Т-парсера (неверные значения узлов и неверные связи между узлами семантической сети технологии);
- ошибка пользователя технологии при ручном занесении информации в БД.

2. Умышленное искажение информации:

- при создании документа-источника, например: искажение возраста с целью избежать призыва на военную службу, сокрытие неприглядных фактов биографии, использование непроверенной информации;
- при формировании БД технологии с целью фальсификации истории [9].

3. Аномалия, вызванная реальным, но редким сочетанием признаков. Ошибкой не является. В качестве примеров можно привести несовершеннолетнего Героя Советского Союза Валю Котика; академика, не имеющего диплома о высшем образовании, И. М. Гельфанда и т. п.

4. Сильные отличия между объектами, вызванные изменением внешних условий. При отсутствии в БД технологии на каком-то этапе ее заполнения большого объема примеров, относящихся к некоторой специфической ситуации или группе, такие редкие объекты будут восприниматься как явные выбросы, хотя ошибкой не являются. Из этого следует, что для корректного определения аномалий, т. е. выявления объектов с нетипичным сочетанием признаков, очень важно правильное определение контекста, поскольку понятие типичности не является универсальным.

3 Типы аномалий в данных

Согласно [10], аномалии в данных могут быть отнесены к одному из трех основных типов.

1. Точечные аномалии возникают в ситуации, когда отдельный экземпляр данных может рассматриваться как аномальный по отношению к остальным данным. Данный вид аномалий наиболее легко распознаваем, большинство существующих методов создано именно для распознавания точечных аномалий.

2. Контекстуальные аномалии наблюдаются, когда экземпляр данных аномален лишь в определенном контексте (данный вид аномалий также называется условным). Для определения аномалий этого типа основным является выделение контекстуальных и поведенческих атрибутов. Контекстуальные атрибуты используются для определения контекста (или окружения) для каждого экземпляра. Поведенческие атрибуты определяют неконтекстуальные характеристики, относящиеся к конкретному экземпляру данных. Аномальное поведение определяется посредством значений поведенческих атрибутов исходя из конкретного контекста. Таким образом, экземпляр данных может быть контекстуальной аномалией при данных условиях, но при таких же поведенческих атрибутах считаться нормальным в другом контексте.
3. Коллективные аномалии возникают, когда последовательность связанных экземпляров данных оказывается аномальной по отношению к целому набору данных. Отдельный экземпляр данных в такой последовательности может не быть отклонением, однако совместное появление таких экземпляров является коллективной аномалией. Коллективные аномалии наблюдаются только в тех наборах данных, где последние связаны между собой.

Но, как показано в [11], точечные или коллективные аномалии могут в то же время быть и контекстуальными. Поэтому следует рассматривать классификацию типов аномалий по двум независимым бинарным признакам: «точечные–коллективные» и «контекстуальные–неконтекстуальные».

Аномалии в конкретно-исторических данных, безусловно, являются точечными и, как показано выше, контекстуальными.

Разделение атрибутов на контекстуальные и поведенческие в технологии ПКИИ является сложной задачей и не может быть проведено заранее, поскольку состав атрибутов формируется в процессе заполнения семантической сети независимыми пользователями. Только пользователь, вводя новый атрибут, может иметь представление о том, является ли он контекстуальным. Для этого объектная модель технологии должна предусматривать возможность установить признак контекстуальности для атрибута. Примерами контекстуальных атрибутов в технологии ПКИИ могут служить: временной период, сословная принадлежность, национальность в связи с существовавшими ограничениями для «инородцев» и т. п.

4 Методы обнаружения точечных аномалий

В [11] подробно рассмотрены существующие методики поиска аномалий. Распознавание аномалий может выполняться в одном из перечисленных ниже режимов.

1. Режим распознавания с учителем (supervised anomaly detection).

Методики, использующие данный режим распознавания, требуют наличия обучающей выборки, полноценно представляющей систему и включающей

экземпляры данных нормального и аномального классов. Работа алгоритма проходит в два этапа: обучение и распознавание.

2. Режим распознавания частично с учителем (semisupervised anomaly detection).

Обучающая выборка при этом подходе представляет только нормальный класс. Обучившись на одном классе, система может определять принадлежность новых данных к нему, таким образом определяя противоположный. Создание обучающей промаркированной выборки обычно проводится вручную и является трудоемким и дорогостоящим процессом.

3. Режим распознавания без учителя (unsupervised anomaly detection).

Данный режим применяется при отсутствии априорной информации о данных. Алгоритмы распознавания в режиме без учителя базируются на предположении о том, что аномальные экземпляры встречаются гораздо реже нормальных. Данные обрабатываются, наиболее отдаленные определяются как аномалии.

Ниже приведены существующие методики распознавания аномалий.

1. Классификация.

Реализация данного метода основана на предположении о том, что нормальное поведение системы может определяться одним или несколькими классами. Таким образом, экземпляр, не принадлежащий ни к одному из классов, считается отклонением. Поиск аномалий проходит в режиме supervised или semisupervised. К наиболее широко применяемым механизмам реализации распознавания аномалий с помощью классификации относятся: нейронные сети, байесовские сети, метод опорных векторов и метод на основе правил.

2. Кластеризация.

Данная методика предполагает группировку похожих экземпляров в кластеры и не требует знаний о свойствах возможных отклонений. Выявление аномалий осуществляется в режиме unsupervised и строится на одном из следующих предположений:

- нормальные экземпляры данных относятся к кластеру данных, в то время как аномалии не принадлежат ни к одному из кластеров;
- нормальные данные ближе к центру кластера, а аномальные — значительно дальше.

Одна из простейших реализаций подхода на основе кластеризации — алгоритм k-means [12].

3. Статистический анализ.

При использовании этого подхода исследуется процесс, строится его профиль (модель), который затем сравнивается с реальным поведением. Если

разница в реальном и предполагаемом поведении системы, определяемая заданной функцией аномальности, выше установленного порога, делается вывод о наличии отклонений. Применяется предположение том, что нормальное поведение системы будет находиться в зоне высокой вероятности, в то время как выбросы — в зоне низкой. Методы статистического анализа подразделяются на параметрические и непараметрические и реализуются в режиме *semisupervised*.

4. Алгоритм ближайшего соседа.

Для использования данной методики, осуществляемой в режиме *unsupervised*, необходимо определить понятие расстояния (меры похожести) между объектами. В рамках методики существуют два подхода к определению аномалии:

- (а) экземпляр, считающийся выбросом, наиболее отдален от ближайшего соседа;
- (б) экземпляр, считающийся выбросом, находится в окрестности с низкой плотностью.

5. Спектральные и гибридные методы.

Гибридные методики распознавания аномалий позволяют сочетать преимущества различных подходов. При этом разные техники могут применяться как последовательно, так и параллельно для достижения усредненных результатов. Совместно с другими алгоритмами для предобработки данных часто применяются спектральные методы, основанные на следующем предположении: данные могут быть вложены в подпространство меньшей размерности, в котором нормальное состояние и аномалии проявляются иначе.

Специфика историко-биографического факта состоит в том, что выявляемые в процессе конкретно-исторического исследования данные отличаются большим многообразием как по составу характеристик, так и по связям между ними [5]. Поскольку к тому же технология ПКИИ является распределенной, а ее данные пополняются независимо не связанными между собою исследователями, можно сделать вывод, что априорные знания о характере аномалий в этих данных принципиально невозможны. Этот вывод отсекает использование методик, осуществляемых в режимах *supervised* или *semisupervised*, поскольку и без того весьма трудоемкая задача создания обучающей выборки становится в этом случае нерешаемой. Учитывая также невозможность определения единой метрики (меры похожести) на столь разнородных данных, методы с опорой на алгоритм ближайшего соседа также в этом случае не применимы. Поэтому единственным подходом к поиску аномалий в конкретно-исторических данных, заслуживающим особого внимания, остается использование методов кластерного анализа, свободных от выявленных выше недостатков.

5 Выводы

Интеграция в технологию ПККИИ механизма поиска аномалий в конкретно-исторических данных может существенно упростить и повысить эффективность такой плохо формализуемой и очень трудоемкой фазы конкретно-исторического исследования, как взаимоувязывание найденной информации и поиск в ней противоречий. Показана принципиальная возможность и перспективность построения этого механизма на базе метода кластеризации. Реализация данного подхода требует дальнейшей проработки как в части выбора конкретного алгоритма и его адаптации под специфику данных технологии ПККИИ, так и в части использования контекстуальных атрибутов с соответствующей модификацией объектной модели технологии.

Актуальность дальнейшего развития в рассмотренном направлении технологии ПККИИ, ориентированной на широкий круг не относящихся к профессиональным историкам и биографам пользователей, вытекает из все возрастающего общественного интереса к частной, семейной истории.

Литература

1. Грибач С. В. Исследование семейных кризисов посредством психолингвистического эксперимента // Сборники конференций НИЦ Социосфера, 2010. № 6. С. 45–54.
2. Адамович И. М., Волков О. И. Технология распределенного автоматизированного анализа исторических текстов // Системы и средства информатики, 2016. Т. 26. № 3. С. 148–161. doi: 10.14357/08696527160311.
3. Адамович И. М., Волков О. И. Единая технология поддержки конкретно-исторических исследований // Системы и средства информатики, 2019. Т. 29. № 1. С. 194–205. doi: 10.14357/08696527190116.
4. Адамович И. М., Волков О. И. Принципы организации данных для технологии поддержки конкретно-исторических исследований // Системы и средства информатики, 2019. Т. 29. № 2. С. 161–171. doi: 10.14357/08696527190116.
5. Адамович И. М., Волков О. И. Иерархическая форма представления биографического факта // Системы и средства информатики, 2016. Т. 26. № 2. С. 108–122. doi: 10.14357/08696527160207.
6. Адамович И. М., Волков О. И. Автоматизированный поиск противоречий в конкретно-исторической информации // Системы и средства информатики, 2020. Т. 30. № 3. С. 145–153. doi: 10.14357/08696527200313.
7. Адамович И. М., Волков О. И. Система извлечения биографических фактов из текстов исторической направленности // Системы и средства информатики, 2015. Т. 25. № 3. С. 235–250. doi: 10.14357/08696527150315.
8. Кисляков А. Н., Поляков С. В. Иерархические методы кластеризации в задаче поиска аномальных наблюдений на основе групп с нарушенной симметрией // Управленческое консультирование, 2020. № 5. С. 116–127.
9. Адамович И. М., Волков О. И. Устойчивость технологии поддержки конкретно-исторических исследований к попыткам искажения истории // Системы и средства информатики, 2021. Т. 31. № 2. С. 152–162. doi: 10.14357/08696527210214.

10. Chandola V., Banerjee A., Kumar V. Anomaly detection: A survey // ACM Comput. Surv., 2009. Vol. 41. Iss. 3. P. 1–58.
11. Шкодырев В. П., Ягафаров К. И., Баитовенко В. А., Ильина Е. Э. Обзор методов обнаружения аномалий в потоках данных // CEUR Workshop Procee., 2017. Vol. 1864. Paper 8. 7 p.
12. Кокорева Я. В., Макаров А. А. Поэтапный процесс кластерного анализа данных на основе алгоритма кластеризации k-means // Молодой ученый, 2015. № 13(93). С. 126–128.

Поступила в редакцию 03.06.22

AN APPROACH TO SEARCHING FOR ANOMALIES IN CONCRETE HISTORICAL DATA

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 119333, Russian Federation

Abstract: The article continues the series of works devoted to the technology of concrete historical research supporting. The technology is based on the principles of co-creation and crowdsourcing and is designed for a wide range of users which are not professional historians and biographers. The article is devoted to the further development of the technology by integrating the mechanism of automated search for anomalies in concrete-historical information. The analysis of the existing approach to the search for contradictions in historical and biographical facts has been carried out and its limitations and shortcomings have been revealed. The expediency of the transition from the search for obvious contradictions to the search for anomalies is justified within the development of this approach. The causes of anomalies have been analyzed, their classification has been carried out, and the features of anomalies in concrete-historical data have been determined. The analysis of the known methods for point anomalies detecting has been carried out and the impossibility of using the methods based on supervised learning as well as metric methods in the technology has been justified. The most promising method based on clustering has been found and further steps for its implementation have been determined. The necessary changes to the technology object model are described and justified.

Keywords: concrete historical investigation; distributed technology; anomaly; historical-biographical fact; automated procedure

DOI: 10.14357/08696527220313

References

1. Gribach, S. V. 2010. Issledovanie semeynykh krizisov posredstvom psikholingvisticheskogo eksperimenta [The study of family crises through a psycholinguistic experiment].

- Sborniki konferentsiy NITs Sotsiosfera* [Conference NIC Sociosfera Proceedings] 6:45–54.
2. Adamovich, I. M., and O. I. Volkov. 2016. Tekhnologiya raspredelenogo 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. doi: 10.14357/08696527160311.
 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. doi: 10.14357/08696527190116.
 4. Adamovich, I. M., and O. I. Volkov. 2019. Printsipy organizatsii dannykh dlya tekhnologii podderzhki konkretno-istoricheskikh issledovaniy [The principles of data organization for the technology of concrete historical research support]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(2):161–171. doi: 10.14357/08696527190214.
 5. Adamovich, I. M., and O. I. Volkov. 2016. Ierarkhicheskaya forma predstavleniya biograficheskogo fakta [Hierarchical format of a biographical fact]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 26(2):108–122. doi: 10.14357/08696527160207.
 6. Adamovich, I. M., and O. I. Volkov. 2020. Avtomatizirovannyy poisk protivorechiy v konkretno-istoricheskoy informatsii [Automated search for contradictions in concrete-historical information]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(3):145–153. doi: 10.14357/08696527200313.
 7. Adamovich, I. M., and O. I. Volkov. 2015. Sistema izvlecheniya biograficheskikh faktov iz tekstov istoricheskoy napravlenosti [The system of facts extraction from historical texts]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(3):235–250. doi: 10.14357/08696527150315.
 8. Kislyakov, A. N., and S. V. Polyakov. 2020. Iyerarkhicheskie metody klasterizatsii v zadache poiska anomal'nykh nablyudeniy na osnove grupp s narushennoy simmetriey [Hierarchical clustering methods in a task to find abnormal observations based on groups with broken symmetry]. *Upravlencheskoe konsul'tirovanie* [Administrative Consulting] 5:116–127.
 9. Adamovich, I. M., and O. I. Volkov. 2021. Ustoychivost' tekhnologii podderzhki konkretno-istoricheskikh issledovaniy k popytkam iskazheniya istorii [The resistance of technology of concrete historical investigation support to attempts of history distortion]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 31(2):152–162. doi: 10.14357/08696527210214.
 10. Chandola, V., A. Banerjee, and V. Kumar. 2009. Anomaly detection: A survey. *ACM Comput. Surv.* 41(3):1–58.
 11. Shkodyrev, V., K. Yagafarov, V. Bashtovenko, and E. Ilyina. 2017. Obzor metodov obnaruzheniya anomal'iy v potokakh dannykh [The overview of anomaly detection methods in data streams]. *CEUR Workshop Procee.* 1864:8. 7 p.
 12. Kokoreva, Ya. V., and A. A. Makarov. 2015. Poetapnyy protsess klasterного analiza dannykh na osnove algoritma klasterizatsii k-means [A phased process of cluster data analysis based on the k-means clustering algorithm]. *Molodoy uchenyy* [Young Scientist] 13(93):126–128.

Received June 3, 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

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

О СРАВНИТЕЛЬНОМ ГЕНЕЗИСЕ БАЗИСНЫХ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ И ОБЩЕСТВЕННО-ЭКОНОМИЧЕСКИХ ФОРМАЦИЙ ЧЕЛОВЕЧЕСТВА

С. Н. Гринченко¹

Аннотация: С позиций информатико-кибернетического моделирования процесса развития самоуправляющейся иерархосетевой системы Человечества рассмотрен параллелизм процессов генезиса базисных информационных технологий (БИТ) и общественно-экономических формаций (ОЭФ) Человечества. Периоду формирования БИТ сигнальных поз/звуков/движений («пред-пред-человеком» *Hominioidea*) соответствует ОЭФ с рабочим названием «пред-пред-первобытно-общинный строй», периоду формирования БИТ мимики/жестов («пред-человеком» *Homo erectus*) — ОЭФ «пред-первобытно-общинный строй», периоду формирования БИТ речи/языка (человеком *Homo sapiens-1*) — ОЭФ «первобытно-общинный строй-1» (ПОС), периоду формирования БИТ письменности/чтения (*Homo sapiens-2*) — ОЭФ «феодализм-1 + ПОС-2», периоду формирования БИТ тиражирования текстов/книгопечатания (*Homo sapiens-3*) — ОЭФ «капитализм-1 + феодализм-2 + ПОС-3», периоду формирования БИТ локальных компьютеров (*Homo sapiens-4*) — ОЭФ « α -цифрализм-1 + капитализм-2 + феодализм-3 + ПОС-4», периоду формирования БИТ телекоммуникаций/сетей (*Homo sapiens-5*) — ОЭФ « β -цифрализм-1 + α -цифрализм-2 + капитализм-3 + феодализм-4 + ПОС-5», периоду формирования перспективной нано-БИТ (*Homo sapiens-6*) — ОЭФ « γ -цифрализм-1 + β -цифрализм-2 + α -цифрализм-3 + капитализм-4 + феодализм-5 + ПОС-6» и т. д. Приводятся характеристики соответствующих управленческих структур: от «пред-пред-демократии» в рамках семьи до перспективной «демократо-6-автократо-5-олигархо-4- α -арифмократо-3- β -арифмократо-2- γ -арифмократии-1» в рамках Промежуточного Космоса.

Ключевые слова: базисные информационные технологии; общественно-экономические формации; информатико-кибернетическая модель; самоуправляющаяся иерархосетевая система Человечества; системный генезис; принцип системной кумуляции; принцип системной согласованности

DOI: 10.14357/08696527220314

Под *базисными информационными технологиями* будем понимать наиболее крупные группы исторически эволюционирующих информационных технологий

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, sgrinchenko@ipiran.ru

общения между людьми, существенно различающихся субстратом (средствами) своей реализации [1].

В свою очередь, энциклопедическая формулировка *общественно-экономической формации* определяет ее как «... исторически определенный тип общества, представляющий собой особую ступень в его развитии; "... общество, находящееся на определенной ступени исторического развития, общество с своеобразным отличительным характером" (Маркс К., см. Маркс К. и Энгельс Ф. Соч., 2-е изд., т. 6, с. 442)» [2]. В своих ранних работах К. Маркс выделял пять таких сменяющих друг друга формаций: первобытно-общинную, рабовладельческую, феодальную, капиталистическую и коммунистическую. На сегодня считается, что «рабовладельческая и феодальная "стадии" составляют в действительности лишь одну единую сословно-классовую общественную формацию» [3, с. 113].

Возникает вопрос: адекватно ли сопоставление системного генезиса БИТ и ОЭФ как в историческом прошлом Человечества, так и в будущем?

Как представляется, системный генезис БИТ, равно как и системный генезис ОЭФ, описывается в терминах этапов глобального эволюционизма самоуправляющейся иерархосетевой системы Человечества как единого саморазвивающегося целого [4, 5]. Его основа — теснейшая связка информационных, производственных и интеллектуальных технологий, используемых усложняющимися индивидами в ходе исторического развития от древнейших времен до наших дней и далее. Согласно информатико-кибернетической модели (ИКМ), критическими точками этих процессов считаются *системные перевороты* в истории Человечества, маркерами которых выступают соответствующие *информационные перевороты*, т. е. старты новых БИТ, с модельными расчетными датировками¹ ~ 28230–1860–123–8, 1 тыс. лет назад–1446–1946–1979–1981–1981 гг.–..., с системной сингулярностью этого ряда около 1981 г. [7, 8].

Здесь уместно отметить, что важнейшей особенностью ИКМ является *принцип системной кумуляции*, согласно которому возникновение в иерархических природных системах новых подсистем не означает элиминации ранее возникших (все они сосуществуют, активно взаимодействуют и коэволюционируют), а также *принцип системной согласованности*, гласящий, что возникновение в ходе эволюции системы Человечества новых подсистем сопровождается кардинальными переменами в структуре и характере приспособительного поведения подсистем, ранее возникших (хотя их вклад в общий ход системного приспособительного поведения уменьшается) [9].

Сопоставляя эти датировки с генезисом ОЭФ, следует указать, что работающие в рамках гуманитарно-философских наук авторы не приводят количественных временных оценок хода этого процесса, давая лишь качественные его характеристики. Модельный подход дает возможность ввести в представле-

¹Временные и пространственные количественные параметры ИКМ базируются на геометрической прогрессии со знаменателем $e^e = 15,15426...$, выявленной А. В. Жирмунским и В. И. Кузьминым при исследовании критических уровней в развитии биологических систем [6].

ние о генезисе ОЭФ соответствующие пространственно-временные координаты, а также целый ряд новых (рабочих) терминов, отражающих как содержательную сторону этого системного процесса, так и этапы его усложнения, маркируемые соответствующими возможностями ИКМ (см. таблицу и рисунок). В частности, такие термины, как:

- (1) **«Пред-пред-первобытно-общинный строй»** с его БИТ сигнальных поз/звучков/движений, для которого характерна превалирующая общественная управленческая структура — семейная «пред-пред-демократия» (на территориях «дворов»);
- (2) **«Пред-первобытно-общинный строй»** с его БИТ мимики/жестов, для которого характерна родоплеменная «пред-демократия» (на территориях «поселений»);
- (3) **«Первобытно-общинный строй-1»** (ПОС-1: единица обозначает факт первичности здесь этой формации, имея в виду усложнение ее проявлений в будущем) с его БИТ речи/языка и «демократией-1» (на территориях «округ»);
- (4) **«Феодализм-1»** с фрагментарными проявлениями **ПОС-2**, с его БИТ письменности/чтения и национальной «демократо-2-автократией-1» (на территориях «сверхрайонов» — национальных государств);
- (5) **«Капитализм-1»** с фрагментарными проявлениями **феодализма-2** и **ПОС-3**, с его БИТ тиражирования текстов/книгопечатания и имперской «демократо-3-автократо-2-олигархией-1» (на территориях «сверхстран» — держав, империй, др.);
- (6) **« α -Цифрализм-1»** (социализм, госкапитализм, постиндустриальное общество и др.) с фрагментарными проявлениями **капитализма-2**, **феодализма-3** и **ПОС-4**, с его БИТ локальных компьютеров, начального этапа развития создаваемого человеком искусственного интеллекта [10, 11] и планетарной (глобальной) «демократо-4-автократо-3-олигархо-2- α -арифмократией-1»;
- (7) **« β -Цифрализм-1»** («кастовый» социализм (?), «инклюзивный» капитализм (?) и т. п.) с фрагментарными проявлениями **α -цифрализма-2**, **капитализма-3**, **феодализма-4** и **ПОС-5**, с его БИТ телекоммуникаций/сетей, возрастающей ролью искусственного интеллекта и «демократо-5-автократо-4-олигархо-3- α -арифмократо-2- β -арифмократией-1» Околоземного Космоса;
- (8) **« γ -Цифрализм-1»** (коммунизм (?), капитализм стейкхолдеров (?) и т. п.) с фрагментарными проявлениями **β -цифрализма-2**, **α -цифрализма-3**, **капитализма-4**, **феодализма-5** и **ПОС-6**, с его перспективной наноаппаратно поддерживаемой БИТ, ведущей ролью искусственного интеллекта и «демократо-6-автократо-5-олигархо-4- α -арифмократо-3- β -арифмократо-2- γ -арифмократией-1» Промежуточного Космоса и т. д.

Сопоставление генезисов БИТ [4] и ОЭФ в системе Человечества

	БИТ. Субъект	Время старта и куль- минации	Ареал (радиус экви- валентного круга)	«Формация/строй» . Тип общественного хозяйства	Превалирующая территориальная управленческая структура
1	Сигнальные позы/звуки/ движения. <i>Hominoides</i>	~ 28,2 и ~ 9,26 млн лет назад	До 64 м, «двор»	«Пред-пред-первобытно- общинный строй» . Присваивающее «пред- пред-хозяйство»	Семейная «пред- пред-демократия» на территории «дво- ра»
2	Мимика/ жесты. <i>Homo erectus</i>	~ 1,86 и ~ 0,612 млн лет назад	До 1 км, «поселение»	«Пред-первобытно-об- щинный строй» . Присваивающее «пред- хозяйство»	Родо-племенная «пред-демократия» на территории «по- селения»
3	Речь/язык. <i>Homo sapiens-1</i>	~ 123 и ~ 40,3 тыс. лет назад	До 15 км, «округа»	Первобытно-общинный строй-1 (ПОС-1) . Присваивающее хозяйство	«Демократия-1» на территории «окру- ги»
4	Письмен- ность/чтение. <i>Homo sapiens-2</i>	~ 8,1 и ~ 2,7 тыс. лет назад	До 223 км, «сверхрайон»	ПОС-2 + феодализм-1 . Мелкое натуральное про- изводство	Национальная «де- мократо-2-автокра- тия-1» на террито- рии «сверхрайона» (национальные госу- дарства)
5	Тиражирова- ние текстов. <i>Homo sapiens-3</i>	~ 1446 и ~ 1806 гг.	До 3370 км, «сверхстрана»	ПОС-3 + феодализм-2 + капитализм-1 . Крупное товарное произ- водство	Имперская «демо- крато-3-автокра- то-2-олигархия-1» на территории «сверхстраны» (им- перии, державы, др.)
6	Локальный компьютер. <i>Homo sapiens-4</i>	~ 1946 и ~ 1970 гг.	До 51 тыс. км, планета Земля	ПОС-4 + феодализм-3 + капитализм-2 + α-цифрализм-1 : социа- лизм, госкапитализм и др. Производство трансна- циональных корпораций	Планетарная (гло- бальная) «демокра- то-4-автократо-3- олигархо-2-α-арифмо- кратия-1» на тер- ритории планеты
7	Телекомму- никации/ сети. <i>Homo sapiens-5</i>	~ 1979 и ~ 2003 гг.	До 773 тыс. км, Околосемный Космос	ПОС-5 + феодализм-4 + капитализм-3 + α-циф- рализм-2 + β-цифра- лизм-1 : «кастовый» со- циализм (?), «инклюзив- ный» капитализм (?) и др. Производство транскон- тинентальных корпораций	«Демократо-5-авто- крато-4-олигархо- 3-α-арифмокра- то-2-β-арифмокра- тия-1» в простран- ственном объеме Околосемного Кос- моса
8	Нано-ИТ. <i>Homo sapiens-6</i>	~ 1981 и ~ 2341 гг.	До 11,7 млн км, Промежуточ- ный Космос	ПОС-6 + феодализм-5 + капитализм-4 + α-циф- рализм-3 + β-цифра- лизм-2 + γ-цифра- лизм-1 : коммунизм (?), капитализм стейкхолде- ров (?) и др. Производство транспла- нетарных корпораций	«Демократо-6-авто- крато-5-олигар- хо-4-α-арифмокра- то-3-β-арифмокра- то-2-γ-арифмокра- тия-1» в простран- ственном объеме Промежуточного Космоса
...

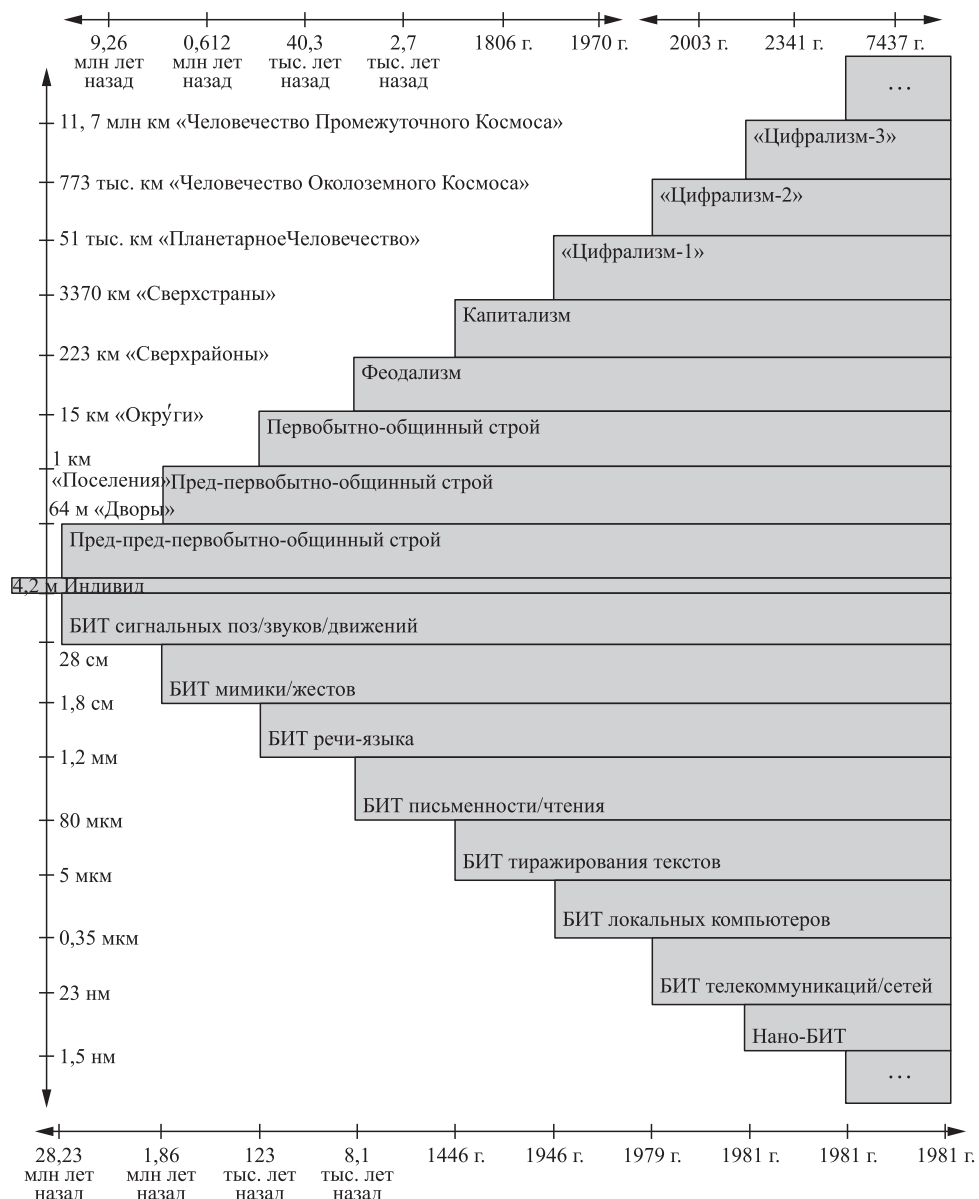


Диаграмма основных этапов генезисов БИТ и ОЭФ в самоуправляющейся системе Человечества (двойной логарифмический масштаб)

Таким образом, предлагаемая модель параллелизма в глобальной эволюции БИТ и ОЭФ системы Человечества, верификация основных положений которой проведена на имеющихся историко-археологических данных [12–14], может быть положена в основу новых подходов к изучению сущности Человека и Человечества как имманентных составляющих Мироздания.

Литература

1. Гринченко С. Н. Базисные информационные технологии и «промышленные революции»: какова их взаимосвязь? // Системы и средства информатики, 2020. Т. 30. № 3. С. 164–172.
2. Глезерман Г. Е. Формация общественно-экономическая // Большая советская энциклопедия. — М.: Советская энциклопедия, 1977. Т. 27. С. 343–345.
3. Илюшечкин В. П. Теория стадийного развития общества (история и проблемы). — М.: Восточная литература РАН, 1996. 406 с.
4. Гринченко С. Н. Метаэволюция (систем неживой, живой и социально-технологической природы). — М.: ИПИ РАН, 2007. 456 с. http://www.ipiran.ru/publications/publications/grinchenko/book_2.
5. Гринченко С. Н. Мировоззренческое значение современных концепций информатики // Открытое образование, 2010. № 6. С. 107–119.
6. Жирмунский А. В., Кузьмин В. И. Критические уровни в процессах развития биологических систем. — М.: Наука, 1982. 179 с.
7. Гринченко С. Н. Развитие базисных информационных технологий и системная сингулярность: кибернетическая точка зрения // Вестник Восточно-Сибирской открытой академии, 2020. № 39. Ст. 3. 10 с. <http://vsoa.esrae.ru/214-1244>.
8. Grinchenko S. N., Shchapova Y. L. The deductive approach to Big History's Singularity // The 21st century singularity and global futures. A Big History perspective / Eds. A. V. Korotayev, D. LePoire. — World-systems evolution and global futures ser. — Cham: Springer, 2020. P. 201–210.
9. Гринченко С. Н. Общение людей и информационные технологии: принципы системной кумуляции и системной согласованности // Мир психологии, 2020. № 3(103). С. 235–245.
10. Гринченко С. Н. Генезис искусственного интеллекта в системе человечества: кибернетическое представление // Современные информационные технологии и ИТ-образование, 2020. Т. 16. № 3. С. 643–652.
11. Гринченко С. Н. О системной иерархии искусственного интеллекта // Информатика и её применения, 2021. Т. 15. Вып. 1. С. 111–115.
12. Гринченко С. Н., Щапова Ю. Л. Информационные технологии в истории Человечества // Информационные технологии, 2013. № S8. С. 1–32.
13. Щапова Ю. Л., Гринченко С. Н. Введение в теорию археологической эпохи: числовое моделирование и логарифмические шкалы пространственно-временных координат. — М.: Ист. фак. Моск. ун-та, ФИЦ ИУ РАН, 2017. 236 с. <http://www.hist.msu.ru/upload/iblock/03f/45831.pdf>.

14. Щанова Ю. Л., Гринченко С. Н., Кокорина Ю. Г. Информатико-кибернетическое и математическое моделирование археологической эпохи: логико-понятийный аппарат. — М.: ФИЦ ИУ РАН, 2019. 136 с.

Поступила в редакцию 28.11.21

ON THE COMPARATIVE GENESIS OF BASIC INFORMATION TECHNOLOGIES AND SOCIOECONOMIC FORMATIONS OF THE HUMANKIND

S. N. Grinchenko

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

Abstract: From the standpoint of informatics-cybernetic modeling of the development process of the self-controlling hierarchical-network system of Humankind, the parallelism of the processes of the genesis of basic information technologies (BIT) and socioeconomic formations (SEF) of Humankind is considered. The period of the formation of BIT signal poses/sounds/movements (“pre-pre-human” Hominoidea) corresponds to the SEF with the working title “pre-pre-primitive communal formation,” the period of the formation of BIT mimics/gestures (“pre-human” Homo erectus) — SEF “pre-primitive communal formation,” the period of formation of BIT of speech/language (human Homo sapiens-1) — SEF “primitive communal formation-1” (PCF), the period of formation of BIT writing/reading (Homo sapiens-2) — SEF “Feudalism-1 + PCF-2,” the period of formation of BIT of replicating texts/printing (Homo sapiens-3) — SEF “Capitalism-1 + Feudalism-2 + PCF-3,” the period of formation of BIT local computers (Homo sapiens-4) — SEF “ α -Digitalism-1 + Capitalism-2 + Feudalism-3 + PCF-4,” the period of formation of BIT of telecommunications/networks (Homo sapiens-5) — SEF “ β -Digitalism-1 + α -Digitalism-2 + Capitalism-3 + Feudalism-4 + PCF-5,” the period of formation of a promising nano-BIT (Homo sapiens-6) — SEF “ γ -Digitalism-1 + β -Digitalism-2 + α -Digitalism-3 + Capitalism-4 + Feudalism-5 + PCF-6,” etc. The characteristics of the corresponding management structures are given: from “pre-pre-democracy” within the family to the promising “democrat–autocrat–oligarch– γ -arithmocrat–3– β -arithmocrat–2– α -arithmocracy-1” within the Intermediate Space.

Keywords: basic information technologies; socioeconomic formations; informatics-cybernetic model; self-controlling hierarchical-network system of Humankind; systemic genesis; systemic cumulation principle; systemic consistency principle

DOI: 10.14357/08696527220314

References

1. Grinchenko, S.N. 2020. Bazisnye informatsionnye tekhnologii i “promyshlennyye revolyutsii”: kakova ikh vzaimosvyaz’? [Basic information technologies and “industrial revolutions”: What is their relationship?] *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(3):164–172.
2. Glezerman, G. E. 1977. Formatsiya obshchestvenno-ekonomicheskaya [Socioeconomic formation]. *Bol’shaya sovetskaya entsiklopediya* [Great Soviet encyclopedia] 27:343–345.
3. Ilyushechkin, V. P. 1996. *Teoriya stadiynogo razvitiya obshchestva (istoriya i problema)* [The theory of staged development of society (history and problems)]. Moscow: Vostochnaya literatura RAN. 406 p.
4. Grinchenko, S.N. 2007. *Metaevolyutsiya (sistem nezhivoy, zhivoy i sotsial’no-tekhnologicheskoy prirody)* [Metaevolution (of inanimate, animate, and sociotechnological nature systems)]. Moscow: IPIRAN. 456 p. http://www.ipiran.ru/publications/publications/grinchenko/book_2,
5. Grinchenko, S.N. 2010. Mirovozzrencheskoe znachenie sovremennykh kontseptsiy informatiki [World-outlook meaning of current informatics conceptions]. *Otkrytoe obrazovanie* [Open Education] 6:107–119.
6. Zhirmunskiy, A. V., and V. I. Kuz’m. 1982. *Kriticheskie urovni v protsessakh razvitiya biologicheskikh sistem* [Critical levels in the development of biological systems]. Moscow: Nauka. 179 p.
7. Grinchenko, S.N. 2020. Razvitie bazisnykh informatsionnykh tekhnologiy i sistemnaya singulyarnost’: kiberneticheskaya tochka zreniya [Basic information technology development and system singularity: A cybernetic point of view]. *Vestnik Vostochno-Sibirskoy otkrytoy akademii* [Bulletin of the East Siberian Open Academy] 39:3. 10 p. Available at: <http://vsoa.esrae.ru/214-1244> (accessed August 11, 2022).
8. Grinchenko, S. N., and Y. L. Shchapova. 2020. The deductive approach to Big History’s Singularity. *The 21st century singularity and global futures*. Eds. A. V. Korotayev and D. LePoire. World-systems evolution and global futures ser. Cham: Springer. 201–210.
9. Grinchenko, S.N. 2020. Obshchenie lyudey i informatsionnye tekhnologii: printsipy sistemnoy kumulyatsii i sistemnoy soglasovannosti [Communication of people and information technology: Principles of systemic cumulation and systemic coordination]. *Mir psikhologii* [The World of Psychology] 3(103):235–245.
10. Grinchenko, S.N. 2020. Genezis iskusstvennogo intellekta v sisteme chelovechestva: Kiberneticheskoe predstavlenie [The genesis of artificial intelligence in the Humankind’s system: A cybernetic view]. *Sovremennyye informatsionnye tekhnologii i IT-obrazovanie* [Modern Information Technologies and IT-Education] 16(3):643–652.
11. Grinchenko, S.N. 2021. O sistemnoy iyerarkhii iskusstvennogo intellekta [On the system hierarchy of artificial intelligence]. *Informatika i ee Primeneniya — Inform. Appl.* 15(1):111–115.
12. Grinchenko, S.N., and Y. L. Shchapova. 2013. Informatsionnye tekhnologii v istorii Chelovechestva [Information technology in history of Humankind]. *Informatsionnye tekhnologii* [Information Technologies] S8:1–32.
13. Shchapova, Y. L., and S.N. Grinchenko. 2017. *Vvedenie v teoriyu arkheologicheskoy epokhi: chislovoe modelirovaniye i logarifmicheskie shkaly prostranstvenno-vremennykh*

koordinat [Introduction to the theory of the archaeological epoch: Numerical modeling and logarithmic scales of space–time coordinates]. Moscow: Faculty of History MSU, FRC CSC RAS. 236 p. <http://www.hist.msu.ru/upload/iblock/03f/45831.pdf>.

14. Shchapova, Yu. L., S. N. Grinchenko, and Yu. G. Kokorina. 2019. *Informatiko-kiberneticheskoe i matematicheskoe modelirovanie arkhelogicheskoy epokhi: logiko-ponyatiynyy apparat* [Informatics-cybernetic and mathematical modeling of the archaeological epoch: Logical-conceptual apparatus]. Moscow: FRC CSC RAS. 136 p.

Received November 28, 2021

Contributor

Grinchenko Sergey N. (b. 1946) — Doctor of Science in technology, 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; sgrinchenko@ipiran.ru

ОБ АВТОРАХ

Адамович Игорь Михайлович (р. 1934) — кандидат технических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Бегишев Вячеслав Олегович (р. 1988) — кандидат физико-математических наук, доцент кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов

Волков Олег Игоревич (р. 1964) — ведущий программист Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Гаврилов Виктор Евдокимович (р. 1950) — главный специалист Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Гончаренко Владислав Владимирович (р. 1994) — аспирант Московского физико-технического института

Гончаров Александр Анатольевич (р. 1994) — научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Григорян Рафаэль Каренович (р. 1992) — аспирант биологического факультета Московского государственного университета имени М. В. Ломоносова; нейрофизиолог лаборатории нейрофизиологии и нейрокомпьютерных интерфейсов ООО «Нейри»

Гринченко Сергей Николаевич (р. 1946) — доктор технических наук, профессор, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Дурново Александр Адрианович (р. 1949) — ведущий программист Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Дьяченко Денис Юрьевич (р. 1987) — инженер-исследователь Федерального исследовательского центра «Информатика и управление» Российской академии наук

Дьяченко Юрий Георгиевич (р. 1958) — кандидат технических наук, старший научный сотрудник 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

Инькова Ольга Юрьевна (р. 1965) — доктор филологических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук; сотрудник Университета Женева

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

Малашенко Юрий Евгеньевич (р. 1946) — доктор физико-математических наук, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Морозов Николай Викторович (р. 1956) — старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Назарова Ирина Александровна (р. 1966) — кандидат физико-математических наук, научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Нейчев Радослав Георгиев (р. 1994) — аспирант Московского физико-технического института

Никишин Дмитрий Александрович (р. 1976) — кандидат технических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Попкова Наталия Александровна (р. 1992) — младший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Рождественский Юрий Владимирович (р. 1952) — кандидат технических наук, ведущий научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Румовская София Борисовна (р. 1985) — кандидат технических наук, научный сотрудник Калининградского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Самохина Алина Максимовна (р. 1996) — аспирант Московского физико-технического института

Степанов Никита Вячеславович (р. 1996) — ассистент кафедры инфокоммуникационных технологий и систем связи Санкт-Петербургского государственного университета аэрокосмического приборостроения

Степченко Дмитрий Юрьевич (р. 1973) — старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Степченко Юрий Афанасьевич (р. 1951) — кандидат технических наук, ведущий научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Стрижов Вадим Викторович (р. 1967) — доктор физико-математических наук, ведущий научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Тюрликов Андрей Михайлович (р. 1957) — доктор технических наук, профессор, заведующий кафедрой инфокоммуникационных технологий и систем связи Санкт-Петербургского государственного университета аэрокосмического приборостроения

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 положения 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^AT_EX (.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 — А4; 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 и
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 povyshe-niya 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 kompaktny tor [Mathematical modeling of the plasma in the compact torus]. Moscow. D.Sc. Diss. 272 p.

Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticheskogo slovary informatzionnogo 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 Pubs. 10 p.

Описание патента:

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.

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 (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^AT_EX (.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 and
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. Mamikonians, 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 slovarya 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 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]. 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 (to Svetlana Strigina)

http://www.ipiran.ru/english/journal_systems.asp

SYSTEMS AND MEANS OF INFORMATICS (SISTEMY I SREDSTVA INFORMATIKI)

SCIENTIFIC JOURNAL

Volume 32 No.3 Year 2022

Editor-in-Chief and Chair of Editorial Council
Academician I. A. Sokolov

IN THIS ISSUE:

SOME METHODOLOGICAL ISSUES OF ASSESSING THE LEVEL OF TECHNOLOGICAL READINESS OF INFORMATION SYSTEMS PROJECTS <i>A. A. Zatsarinny and Yu. S. Ionenkov</i>	4
PROBLEMS AND THREATS OF SOME NEW DIGITAL TECHNOLOGIES IMPLEMENTATION <i>V. E. Gavrilov and A. A. Zatsarinny</i>	15
MODELING OF THE DEVELOPMENT OF A COMPLETE EARLY DIFFERENTIAL DIAGNOSIS AND PROGNOSIS OF THE STATE OF COMPLEX OBJECTS IN MEDICINE. PROBLEMS AND PROSPECTS <i>S. B. Rumovskaya and I. A. Kirikov</i>	26
CLASSIFICATION MODELS FOR P300 EVOKED POTENTIALS <i>A. M. Samokhina, R. G. Neychev, V. V. Goncharenko, R. K. Grigoryan, and V. V. Strijov</i>	36
OVERVIEW OF APPROACHES TO SPACE-TIME MODELING AND THE MAIN TRENDS IN THE DEVELOPMENT OF T-GIS <i>D. A. Nikishin</i>	50
OPTIMAL DEVICE ASSOCIATIONS IN MACHINE-TO-MACHINE NETWORKS OF NB-IoT AND LTE-M STANDARDS <i>N. V. Stepanov, A. M. Turlikov, and V. O. Begishev</i>	63
ANALYSIS OF THE DISTRIBUTION OF LOADS AND RESOURCES IN A MULTIUSER NETWORK WITH EQUAL-SHARE INTERNODAL FLOWS <i>Yu. E. Malashenko and I. A. Nazarova</i>	71
SELF-TIMED SHIFT REGISTER CASES <i>Yu. A. Stepchenkov, Yu. G. Diachenko, Yu. V. Rogdestvenski, N. V. Morozov, D. Yu. Stepchenkov, and D. Yu. Diachenko</i>	81